Dmitry Dolgopyat and Omri M. Sarig

Local Limit Theorems for Inhomogeneous Markov Chains

Springer Nature

Acknowledgements

The authors wish to thank Jon Aaronson, Manfred Denker, Yuri Kifer, Omri Solan, and Ofer Zeitouni for useful comments. Special thanks go to Yeor Hafouta for numerous pertinent comments on early drafts of this work.

The work on this monograph was partially supported by the U.S.-Israel Binational Science Foundation grant 201610 (D.D & O.S.), by the National Science Foundation grants DMS 1665046 and DMS 1956049 (D.D.), and by the Israel Science Foundation grants 1149/18 and 264/22 (O.S.).

Contents

Notation

$\nabla \mathsf{a}$	the additive functional $\{a_{n+1}^{(N)}(X_{n+1}^{(N)}) - a_n^{(N)}(X_n^{(N)})\}$ (a gradient)
$\mathscr{B}(\mathfrak{S})$	the Borel σ-algebra of a separable complete metric space $\mathfrak{S}$
$\mathfrak{c}_-, \mathfrak{c}_+$	large deviations thresholds, see §7.4
$C_c(\mathbb{R})$	the space of continuous $\varphi : \mathbb{R} \to \mathbb{R}$ with compact support
C_{mix}	the mixing constant from Proposition 2.13
Cov	the covariance
CVar	the circular variance, see §4.3.1
$d_n(\xi), d_n^{(N)}(\xi)$	structure constants, see §2.3
$D_N(\xi)$	structure constants, see §2.3
δ_x	point mass measure at x (Dirac's measure)
$\delta(\pi)$	the contraction coefficient of a Markov operator π, see §2.2.2
$\delta(\mathsf{f})$	the graininess constant of f, see chapter 4
ϵ_0	(usually) the uniform ellipticity constant, see §2.2.1
$\mathbb{E}, \mathbb{E}_x$	the expectation operator. $\mathbb{E}_x := \mathbb{E}(\cdot \mid X_1 = x)$
ess sup	the essential supremum, see chapter 2
$\mathsf{f}, \mathsf{g}, \mathsf{h}$	additive functionals
$f_n, f_n^{(N)}$	an entry of an additive functional f of a Markov chain or array
$\mathcal{F}_N(\xi)$	the normalized log-moment generating function, see chapter 7
$G_{alg}(\mathsf{X}, \mathsf{f})$	the algebraic range, see chapter 4
$G_{ess}(\mathsf{X}, \mathsf{f})$	the essential range, see chapter 4
Γ	the balance (of a hexagon), see §2.3.1
$H(\mathsf{X}, \mathsf{f})$	the co-range, see chapter 4
$\mathrm{Hex}(N, n)$	the space of level N hexagons at position n, see §2.3.1
$\mathcal{I}_N(\eta)$	the rate function, see chapter 7
k_N	(usually) the length of the N-th row of an array, minus one
log	the natural logarithm (same as ln)
$m_{\mathrm{Hex}}, m'_{\mathrm{Hex}}$	hexagon measures, see §2.3.1
$\mu(\mathrm{d}x)$	a measure with its integration variable
Osc	the oscillation, see §2.2.2
$(\Omega, \mathscr{F}, \mu, T)$	a measurable map $T : \Omega \to \Omega$ on a measure space $(\Omega, \mathscr{F}, \mu)$
$\mathbb{P}(A), \mathbb{P}_x(A)$	the probability of the event A. $\mathbb{P}_x(A) := \mathbb{P}(A \mid X_1 = x)$
$\pi_{n,n+1}(x, \mathrm{d}y)$	the n-th transition kernel of a Markov chain
$p_n(x, y)$	(usually) the density of $\pi_{n,n+1}(x, \mathrm{d}y)$
$\Phi_N(x, \xi)$	characteristic functions, see §5.2.2
$\mathfrak{r}_+, \mathfrak{r}_-$	positivity thresholds, see §7.4
S_N	$\sum_{i=1}^{N} f_i(X_i, X_{i+1})$ (chains), or $\sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$ (arrays)
$\mathfrak{S}_n, \mathfrak{S}_n^{(N)}$	the state space of X_n (chains) or of $X_n^{(N)}$ (arrays)

$u_n, u_n^{(N)}, U_N$	structure constants, see §2.3
Var	the variance
V_N	the variance of S_N
X_n, $X_n^{(N)}$	an entry of a Markov chain, or a Markov array
X	a Markov chain or a Markov array
$\mathsf{X}^{(N)}$	the N-th row of a Markov array
z_N	(usually) a real number not too far from $\mathbb{E}(S_N)$
a.e.; a.s.	almost everywhere; almost surely
CLT	Central Limit Theorem
iid	independent and identically distributed
LLT	Local Limit Theorem
LHS, RHS	left-hand-side, right-hand-side (of an equation)
TFAE	the following are equivalent
s.t.	such that
w.l.o.g.	without loss of generality
$\because$, $\therefore$	because, therefore
$\wedge$, $\vee$	$x \wedge y := \min\{x, y\}$, $x \vee y := \max\{x, y\}$
1_E	the indicator function, equal to 1 on E and to 0 elsewhere
$a \pm \varepsilon$	a quantity inside $[a - \varepsilon, a + \varepsilon]$
$\mathrm{e}^{\pm\varepsilon} a$	a quantity in $[\mathrm{e}^{-\varepsilon} a, \mathrm{e}^{\varepsilon} a]$
$\sim$	$a_n \sim b_n \Leftrightarrow a_n/b_n \xrightarrow[n\to\infty]{} 1$
$\asymp$	$a_n \asymp b_n \Leftrightarrow 0 < \liminf(a_n/b_n) \leq \limsup(a_n/b_n) < \infty$
$\lesssim$	$a_n \lesssim b_n \Leftrightarrow \limsup(a_n/b_n) < \infty$
$\ll$	for measures: $\mu \ll \nu$ means "$\nu(E) = 0 \Rightarrow \mu(E) = 0$ for all measurble E"; For numbers: non-rigorous "much smaller than"
$\approx$	non-rigorous shorthand for "approximately equal"
$:=$	is defined to be equal to
$\overset{!}{=}, \overset{!}{\leq}, \overset{!}{\sim}$	an equality, inequality, or asymptotic that will be justified later
$\overset{?}{=}, \overset{?}{\leq}, \overset{?}{\sim}$	an equality, inequality, or asymptotic whose veracity is unknown
$X_n \xrightarrow[n\to\infty]{\text{prob}} Y$	convergence in probability
$X_n \xrightarrow[n\to\infty]{\text{dist}} Y$	convergence in distribution
$X_n \xrightarrow[n\to\infty]{L^p} Y$	convergence in L^p
$[S_N > t]$	The event that the condition in brackets happens. For example, if $\varphi : \mathfrak{S} \to \mathbb{R}$, then $[\varphi(\omega) > t] := \{\omega \in \mathfrak{S} : \varphi(\omega) > t\}$
$\lfloor x \rfloor$, $\lceil x \rceil$	$\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$, $\lceil x \rceil := \min\{n \in \mathbb{Z} : n \geq x\}$
$\{x\}, \langle x \rangle$	$\{x\} := x - \lfloor x \rfloor$; $\langle x \rangle$ = the number in $[-\pi, \pi)$ s.t. $x - \langle x \rangle \in 2\pi\mathbb{Z}$
$\{x\}_{t\mathbb{Z}}, [x]_{t\mathbb{Z}}$	the numbers s.t. $x = [x]_{t\mathbb{Z}} + \{x\}_{t\mathbb{Z}}$, $[x]_{t\mathbb{Z}} \in t\mathbb{Z}$ and $\{x\}_{t\mathbb{Z}} \in [0, t)$

The **Fourier Transform** of an L^1-function $\phi : \mathbb{R} \to \mathbb{R}$: $\widehat{\phi}(\xi) := \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi u} \phi(u) \mathrm{d}u$.

The **Legendre-Fenchel transform** of a convex $\phi : \mathbb{R} \to \mathbb{R}$: $\phi^*(\eta) := \sup\limits_{\xi}[\xi\eta - \varphi(\xi)]$.

Chapter 1
Overview

Abstract We give an overview of the main results of this work.

1.1 Setup and Aim

Our aim is to describe the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (a,b)]$ as $N \to \infty$, where $S_N = \sum_{n=1}^{N} f_n(X_n, X_{n+1})$, X_n is a Markov chain, and z_N are real numbers not too far from $\mathbb{E}(S_N)$. Such results are called **local limit theorems (LLT)**. For the history of the problem, see the end of the chapter. The novelty of this work is that we allow the Markov chain to be **inhomogeneous**: The set of states of X_n, the transition probabilities, and f_n may all depend on n.

We will usually assume that f_n are uniformly bounded real-valued functions, and that $\{X_n\}$ is uniformly elliptic, a technical condition which we will state in Chapter 2, and that implies uniform exponential mixing. These assumptions place us in the Gaussian domain of attraction. The analogy with classical results for sums of independent identically distributed (iid) random variables suggests that in the best of all situations, we should expect the behavior below (in what follows $V_N := \mathrm{Var}(S_N)$, $A_N \sim B_N \Leftrightarrow A_N/B_N \xrightarrow[N\to\infty]{} 1$, and the question marks are there to emphasize that at this point of the discussion, these are conjectures, not assertions):

(1) **Local Deviations:** If $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, then

$$\mathbb{P}[S_N - z_N \in (a,b)] \overset{?}{\sim} \frac{|a-b|}{\sqrt{2\pi V_N}} \exp[-z^2/2].$$

(2) **Moderate Deviations:** If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then

$$\mathbb{P}[S_N - z_N \in (a,b)] \overset{?}{\sim} \frac{|a-b|}{\sqrt{2\pi V_N}} \exp\Big[-\frac{1+o(1)}{2}\Big(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\Big)^2\Big].$$

(3) **Large Deviations:** If $\Big|\frac{z_N - \mathbb{E}(S_N)}{V_N}\Big| < c$ with $c > 0$ sufficiently small, then

$$\mathbb{P}[S_N - z_N \in (a,b)] \overset{?}{\sim} \frac{|a-b|}{\sqrt{2\pi V_N}} \exp\Big[-V_N \mathcal{I}_N(z_N/V_N)\Big] \times$$
$$\times \rho_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right) \times \frac{1}{|a-b|}\int_a^b \mathrm{e}^{-t\xi_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)}\,\mathrm{d}t,$$

where $\mathcal{I}_N(\cdot)$ are the Legendre transforms of $\mathcal{F}_N(\xi) := \frac{1}{V_N}\log \mathbb{E}(\mathrm{e}^{\xi S_N})$, and

- $\rho_N(t) \xrightarrow[t\to 0]{} 1$ uniformly in N, and $\rho_n(\cdot)$ are uniformly bounded away from zero and infinity on $[-c, c]$;
- $C^{-1}|\eta| \le |\xi_N(\eta)| \le C|\eta|$ for all $|\eta| < c$ and N, with C independent of N;
- c, ξ_N, ρ_N are independent of z_N and (a,b).

(The asymptotic results in the large deviation regime are more precise than in the moderate deviation case, but less universal. See Chapter 7 for more details.)

Although the asymptotic formulas (1)–(3) above are true in many cases, they sometimes fail — even when S_N is a sum of iid's.

The aim of this work is to give general sufficient conditions for (1)–(3), and to provide the necessary asymptotic corrections when some of these conditions fail. To do this we first identify all the obstructions to (1)–(3), and then we analyze S_N when these obstructions happen.

1.2 The Obstructions to the Local Limit Theorems

The **algebraic range** is the smallest closed additive subgroup $G \le \mathbb{R}$ for which there are $\alpha_n \in \mathbb{R}$ so that $f_n(X_n, X_{n+1}) - \alpha_n \in G$ almost surely for all n. We show that the following list is a complete set of obstructions to (1)–(3):

(I) **Lattice Behavior**: The algebraic range is $t\mathbb{Z}$ for some $t > 0$.

(II) **Center-Tightness**: There are centering constants m_N such that $\{S_N - m_N\}$ is tight. In Chapter 3 we will see that in this case $\mathrm{Var}(S_N)$ must be bounded. We will also see that center-tightness is equivalent to $\mathrm{Var}(S_N) \not\to \infty$.

(III) **Reducibility**: $f_n = g_n + c_n$ where the algebraic range of $\{g_n(X_n, X_{n+1})\}$ is strictly smaller than the algebraic range of $\{f_n(X_n, X_{n+1})\}$, and where $S_n(\mathsf{c}) := \sum_{n=1}^{N} c_n(X_n, X_{n+1})$ is center-tight (equivalently, its variance does not tend to infinity as $N \to \infty$).

One of our main results is that if these three obstructions do not occur, then the asymptotic expansions (1)–(3) above hold.

1.3 How to Show that the Obstructions Do Not Occur

While it is usually easy to rule out the lattice obstruction (I), it is often not clear how to rule out (II) and (III). What we need is a tool that determines from the data of f_n and X_n whether $\{f_n(X_n, X_{n+1})\}$ is center-tight or reducible.

In Chapter 2, we introduce numerical constants $d_n(\xi)$ $(n \ge 3, \xi \in \mathbb{R})$, which are defined purely in terms of the functions f_n and the transition probabilities

$$\pi_{n,n+1}(x, E) := \mathbb{P}(X_{n+1} \in E | X_n = x),$$

and which can be used to determine which obstructions occur and which do not:

- If $\sum d_n^2(\xi) = \infty$ for all $\xi \neq 0$, then the obstructions (I),(II),(III) do not occur, and the asymptotic expansions (1)–(3) hold.
- If $\sum d_n^2(\xi) < \infty$ for all $\xi \neq 0$, then $\mathrm{Var}(S_N)$ is bounded (obstruction II).
- If $\sum d_n^2(\xi) = \infty$ for some but not all $\xi \neq 0$, then $\mathrm{Var}(S_N) \to \infty$ but we are either lattice or reducible: (II) fails, but at least one of (I),(III) occurs.

We call $d_n(\xi)$ the **structure constants** of $\mathsf{X} = \{X_n\}$ and $\mathsf{f} = \{f_n\}$.

1.4 What Happens When the Obstructions Do Occur

1.4.1 Lattice Case

The lattice obstruction (I) already happens for sums of iid random variables, and the classical approach how to adjust (1)–(3) to this setup extends without much difficulty to the Markov case.

Suppose the algebraic range is $t\mathbb{Z}$ with $t > 0$, i.e. there are constants α_n such that $f_n(X_n, X_{n+1}) - \alpha_n \in t\mathbb{Z}$ almost surely for all n. Assume further that $t\mathbb{Z}$ is the smallest group with this property. In this case $S_N \in \gamma_N + t\mathbb{Z}$ a.s. for all N, where $\gamma_N = \sum_{i=1}^N \alpha_i \mod t\mathbb{Z}$. Instead of analyzing $\mathbb{P}[S_N - z_N \in (a,b)]$, which may be equal to zero, we study $\mathbb{P}[S_N = z_N]$, with $z_N \in \gamma_N + t\mathbb{Z}$.

We show that in case (I), if the algebraic range is $t\mathbb{Z}$, and obstructions (II) and (III) do not occur, then (as in the case of iid's):

(1') If $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$ and $z_N \in \gamma_N + t\mathbb{Z}$, then $\mathbb{P}[S_N = z_N] \sim \frac{t}{\sqrt{2\pi V_N}} e^{-z^2/2}$.

(2') If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$ and $z_N \in \gamma_N + t\mathbb{Z}$, then $\mathbb{P}[S_N = z_N] \sim \frac{t}{\sqrt{2\pi V_N}} \exp\left[-\frac{1+o(1)}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right]$.

(3') If $\left|\frac{z_N - \mathbb{E}(S_N)}{V_N}\right| < c$ with $c > 0$ sufficiently small, and $z_N \in \gamma_N + t\mathbb{Z}$, then

$$\mathbb{P}[S_N = z_N] \sim \frac{t}{\sqrt{2\pi V_N}} \exp[-V_N \mathcal{I}_N(z_N/V_N)] \times \rho_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right),$$

where $\mathcal{I}_N$ and ρ_N have the properties listed in the non-lattice case (3).

The previous results hold for lattice valued, irreducible, non-center-tight additive functionals, that is, when (I) holds and (II),(III) fail. Here is an equivalent condition in terms of the data of X_n and f_n:

$$\exists t > 0 \text{ such that } \sum d_n^2(\xi) < \infty \text{ exactly when } \xi \in \frac{2\pi}{t}\mathbb{Z}.$$

Under this condition, (1')–(3') hold with the parameter t.

1.4.2 Center-Tight Case

We show that obstruction (II) happens iff $f_n(X_n, X_{n+1})$ can be put in the form

$$f_n(X_n, X_{n+1}) = a_{n+1}(X_{n+1}) - a_n(X_n) + h_n(X_n, X_{n+1}) + c_n \tag{1.1}$$

where $a_n(X_n)$ are uniformly bounded, c_n are constants, $h_n(X_n, X_{n+1})$ have mean zero, and

$$\sum \mathrm{Var}[h_n(X_n, X_{n+1})] < \infty.$$

The freedom in choosing $a_n(X_n)$ is too great to allow general statements on the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (a,b)]$, see Example 3.1. But as we shall we see in Chapter 3, (1.1) does provide us with some almost sure control. It implies that

$$S_N = a_{N+1}(X_{N+1}) - a_1(X_1) + \sum_{n=1}^N h_n(X_n, X_{n+1}) + \gamma_N,$$

where $\gamma_N = \sum_{i=1}^N c_i$, and where $\sum_{n=1}^\infty h_n(X_n, X_{n+1})$ *converges almost surely.*

This means that in the center-tight scenario, $S_N - \mathbb{E}(S_N)$ can be decomposed into the sum of two terms: A bounded term, possibly oscillatory, that depends only on X_{N+1}, and a term which depends on the entire past $X_1, \ldots, X_{N+1}$, that converges almost surely.

1.4.3 Reducible Case

In the reducible case, we can decompose

$$f_n(X_n, X_{n+1}) = g_n(X_n, X_{n+1}) + c_n(X_n, X_{n+1}) \tag{1.2}$$

with center-tight $\mathsf{c} = \{c_n(X_n, X_{n+1})\}$, and where the algebraic range of $\mathsf{g} = \{g_n(X_n, X_{n+1})\}$ is strictly smaller than the algebraic range of $\mathsf{f} = \{f_n(X_n, X_{n+1})\}$.

In principle, it is possible that g is reducible too, but in Chapter 6 we show that one can find an "optimal" decomposition where g is irreducible, and cannot be decomposed further. The algebraic range of the "optimal" g is the "infimum" of all possible reduced ranges:

$$G_{ess} := \bigcap \left\{ G : \begin{array}{l} G \text{ is the algebraic range of some } \mathsf{g} \\ \text{which satisfies (1.2) with a center-tight } \mathsf{c} \end{array} \right\}.$$

We call G_{ess} the **essential range** of f. It can be calculated explicitly from the data of f and the Markov chain X, in terms of the structure constants, see Theorem 4.4.

It follows from the definitions that G_{ess} is a proper closed subgroup of $\mathbb{R}$, so $G_{ess} = \{0\}$ or $t\mathbb{Z}$ or $\mathbb{R}$. In the reducible case, $G_{ess} = \{0\}$ or $t\mathbb{Z}$, because if $G_{ess} = \mathbb{R}$, then the algebraic range (which contains G_{ess}) is also equal to $\mathbb{R}$.

If $G_{ess} = \{0\}$, then the optimal g has algebraic range $\{0\}$, and $g_n(X_n, X_{n+1})$ are a.s. constant. In this case f is center-tight, and we are back to case (II).

If $G_{ess} = t\mathbb{Z}$ with $t > 0$, then g is lattice, non-center-tight, and irreducible. Split

$$S_N = S_N(\mathsf{g}) + S_N(\mathsf{c}), \text{ with } S_N(\mathsf{g}) = \sum_{n=1}^{N} g_n(X_n, X_{n+1}) \,,\, S_n(\mathsf{c}) = \sum_{n=1}^{N} c_n(X_n, X_{n+1}). \tag{1.3}$$

Then $S_n(\mathsf{g})$ satisfies the lattice LLT (1')–(3') with parameter t, and $\text{Var}[S_N(\mathsf{c})] = O(1)$. Trading constants between g and c, we can also arrange $\mathbb{E}[S_N(\mathsf{c})] = O(1)$.

Unfortunately, even though $\text{Var}[S_n(\mathsf{f})] \to \infty$ and $\text{Var}[S_N(\mathsf{c})] = O(1)$, examples show that $S_N(\mathsf{c})$ is still powerful enough to disrupt the local limit theorem for S_N, lattice or non-lattice (Example 6.1). Heuristically, what could happen is that the mass of $S_N(\mathsf{g})$ concentrates on cosets of $t\mathbb{Z}$ according to (1')–(3'), but $S_N(\mathsf{c})$ spreads this mass to the vicinity of the lattice in a non-universal but tight manner.

This suggests that (1)–(3) should be approximately true for intervals (a, b) of length $|a - b| \gg t$, but false for intervals of length $|a - b| \ll t$. In Chapter 6 we prove results in this direction.

For intervals with size $|a - b| > 2t$, we show that for all $z_N \in \mathbb{R}$ such that $(z_N - \mathbb{E}(S_N))/\sqrt{V_N} \to z$, for all N large enough,

$$\frac{1}{3}\left(\frac{\mathrm{e}^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\right) \le \mathbb{P}[S_N - z_N \in (a,b)] \le 3\left(\frac{\mathrm{e}^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\right).$$

If $|a - b| > L > t$, we can replace 3 by a constant $C(L,t)$ such that $C(L,t) \xrightarrow[L/t\to\infty]{} 1$.

For general intervals, possibly with length less than t, we show the following: There is a random variable $\mathfrak{H} = \mathfrak{H}(X_1, X_2, X_3, \ldots)$ and uniformly bounded random variables $b_N = b_N(X_1, X_{N+1})$ so that for every $z_N \in t\mathbb{Z}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, and for every $\phi : \mathbb{R} \to \mathbb{R}$ continuous with compact support,

$$\lim_{N\to\infty}\sqrt{V_N}\mathbb{E}[\phi(S_N - z_N - b_N(X_1, X_{N+1}))] = \frac{t\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}}\mathbb{E}[\phi(mt+\mathfrak{H})]. \tag{1.4}$$

For $\phi \approx 1_{[a,b]}$ with $|a-b| \gg t$, the right-hand-side of (1.4) is approximately equal to $\frac{1}{\sqrt{2\pi}}|a-b|\mathrm{e}^{-z^2/2}$, in accordance with (1), see Lemma 6.10. But for $|a-b| \ll t$, (1.4) depends on the detailed structure of f through t, $b_N(X_1, X_{N+1})$ and $\mathfrak{H}$.

What are $b_N(X_N, X_{N+1})$ and $\mathfrak{H}$? Recall that c on the right-hand-side of (1.3) is center-tight. As such, it can be put in the form (1.1). Namely, $c_n(X_n, X_{n+1}) = a_{n+1}(X_{n+1}) - a_n(X_n) + h_n(X_n, X_{n+1}) + c_n^*$, where $\sup_n(\operatorname{ess\,sup}|a_n|) < \infty$, c_n^* are constants, $\mathbb{E}(h_n(X_n, X_{n+1})) = 0$, and $\sum h_n$ converges almost surely. Let $\gamma_N := \sum_{n=1}^{N} c_n^* = \mathbb{E}(S_N(\mathsf{c})) + O(1) = O(1)$. The proof of (1.4) shows that

- $b_N = a_{N+1}(X_{N+1}) - a_1(X_1) + \{\gamma_N\}_{t\mathbb{Z}}$, where $\{x\}_{t\mathbb{Z}} \in [0,t)$, $\{x\}_{t\mathbb{Z}} = x \bmod t\mathbb{Z}$;
- $\mathfrak{H} = \sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$. (It is possible to replace $\mathfrak{H}$ by a different random variable $\mathfrak{F}$ which is bounded, see Chapter 6.)

This works as follows. Let $z_N^* := z_N - [\gamma_N]_{t\mathbb{Z}}$, where $[x]_{t\mathbb{Z}} := x - \{x\}_{t\mathbb{Z}} \in t\mathbb{Z}$. Then $z_N^* \in t\mathbb{Z}$, $\frac{z_N^* - \mathbb{E}(S_N)}{V_N} = \frac{z_N - \mathbb{E}(S_N) + O(1)}{V_N} \to z$, and

$$S_N - b_N - z_N = [S_N(\mathsf{g}) - z_N^*] + S_N(\mathsf{h}). \tag{1.5}$$

By subtracting b_N from S_N, we are shifting the distribution of S_N to the distribution of the sum of two terms: The first, $S_N(\mathsf{g})$, is an *irreducible* $t\mathbb{Z}$-valued additive functional; the second, $S_N(\mathsf{h})$, converges almost surely to $\mathfrak{H}$.

Suppose for the sake of discussion that $S_N(\mathsf{g})$, $S_N(\mathsf{h})$ were independent. Then (1.5), the identity $\mathfrak{H} = \sum_{n\geq 1} h_n(X_n, X_{n+1})$, and the lattice LLT for $S_N(g)$ say that

$$\lim_{N\to\infty}\sqrt{V_N}\mathbb{E}[\phi(S_N - b_N - z_N)] = \int_{\mathbb{R}}\phi(x)m(\mathrm{d}x), \tag{1.6}$$

where $m := \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}} m_{t\mathbb{Z}} * m_{\mathfrak{H}}$, $m_{\mathfrak{H}}(E) := \mathbb{P}[\mathfrak{H} \in E]$, $m_{t\mathbb{Z}} := t\cdot$counting measure of $t\mathbb{Z}$, and $*$ denotes the convolution. (See §5.2.3.) Calculating, we find that $\int_{\mathbb{R}}\phi \mathrm{d}m =$ right-hand-side of (1.4).

In general, $S_N(\mathsf{g})$ and $S_N(\mathsf{h})$ are not independent. But in Chapter 6 we show that (1.4) and (1.6) remain valid. There we also discuss other consequences of (1.4), including the asymptotic distributional behavior of S_N modulo $t\mathbb{Z}$.

1.5 Some Final Words on the Setup of this Work

We would like to comment on a choice we made when we wrote this work, specifically, our focus on additive functionals of the form $f_n = f_n(X_n, X_{n+1})$.

This choice is somewhat unorthodox: The theory of Markov processes is mostly concerned with the case $f_n = f_n(X_n)$ (see e.g. [50, 149, 181]), and the theory of stochastic processes is mostly concerned with the case $f_n = f_n(X_n, X_{n+1}, \ldots)$, under assumptions on the weak dependence of X_k and X_n when $|k-n| \gg 1$ (see e.g. [103, 16]). We decided to study $f_n = f_n(X_n, X_{n+1})$ for the following reasons:

- The case $f_n = f_n(X_n, X_{n+1})$ is richer than the case $f_n = f_n(X_n)$ because it contains gradients $a_{n+1}(X_{n+1}) - a_n(X_n)$. Two additive functionals which differ by a gradient with uniformly bounded $\operatorname{ess\,sup}|a_n|$ will have the same CLT behavior, but they may have different LLT behavior, because their algebraic ranges can be different. This leads to an interesting reduction theory which we would have missed had we only considered the case $f_n = f_n(X_n)$.[1]

- The case $f_n(X_n, \dots, X_{n+m})$ with $m > 1$ is similar to the case $m = 2$, and does not require new ideas, see Example 2.3 and the discussion in §2.4. We decided to keep $m = 1$ and leave the (routine) extension to $m > 1$ to the reader.
- The case $f_n = f_n(X_n, X_{n+1}, \dots)$ is of great interest, and we hope to address it in the future. At the moment, our results do not cover it.

We hope to stimulate research into the local limit theorem of additive functionals of general non-stationary stochastic processes with mixing conditions. Such work will have applications outside the theory of stochastic processes, such as the theory of dynamical systems. Our work here is a step in this direction.

1.6 Prerequisites

We made an attempt to make this text self-contained and accessible to readers with standard background in analysis and probability theory. A familiarity with the material of Rudin's book *Real and complex analysis* [170, Ch. 1-9] and Breiman's book *Probability* [17, Ch. 1-8] should be sufficient. Appendices A-C supply additional background material, not in these textbooks.

A few sections marked by $(^*)$ contain topics which are slightly off the main path of the book. Some of these sections require additional background, which we recall, but sometimes without proofs. The material in the starred sections is not used in other parts of the book, and they can be skipped at first reading.

1.7 Notes and References

The local limit theorem has a very long history. To describe it, let us distinguish the following three lines of development:

(1) LLT for identically distributed independent (iid) random variables,
(2) LLT for other stationary stochastic processes,
(3) LLT for non-stationary stochastic processes.

Local Limit Theorems for Sums of IID Random Variables. The first LLT dates to de Moivre's 1738 book [38], and provides approximations for $\mathbb{P}[a \le S_n \le b]$ when $S_n = X_1 + \cdots + X_n$, and X_i are iid, equal to zero or one with equal probabilities. Laplace extended de Moivre's results to the case when X_i are equal to zero or one with non-equal probabilities [124, 125].

In 1921, Pólya extended these results to the vector valued iid which generate the simple random walk on $\mathbb{Z}^d$, and deduced his famous criterion for the recurrence of simple random walks [162].

The next historical landmark is Gnedenko's 1948 work [78, 79] which initiated the study of the LLT for sums of iid with *general* lattice distributions. He asked for the weakest possible assumptions on the distribution of iid's X_i which lead to LLT with Gaussian or stable limits. Khinchin popularized the problem by emphasizing its importance to the foundations of quantum statistical physics [109], and it was studied intensively by the Russian school, with important contributions by Linnik, Ibragimov, Prokhorov, Richter, Saulis, Petrov and others. We will comment on some of these contributions in later chapters. For the moment, we refer the reader to the excellent books by Gnendenko & Kolmogorov [80], Ibragimov & Linnik [103], Petrov [156], and to the many references they contain.

[1] We cannot reduce the case $f_n(X_n, X_{n+1})$ to the case $f_n(Y_n)$ by working with the Markov chain $Y_n = (X_n, X_{n+1})$ because $\{Y_n\}$ will no longer satisfy some of our standing assumptions, specifically the uniform ellipticity condition (see Chapter 2).

The early works on the local limit theorem all focused on the lattice case. The first result we are aware of which could be considered to be a non-lattice local limit theorem is in [80]: Suppose that each of the iid's X_i have mean zero, finite variance σ^2, and a probability density function $p(x) \in L^r$ with $1 < r \le 2$, then the density function $p_n(x)$ of $X_1 + \cdots + X_n$ satisfies $\sigma\sqrt{n}p_n(\sigma\sqrt{n}x) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

There could be non-lattice iid's without density functions, for example the iid's X_i equal to (-1), 0, or $\sqrt{2}$ with equal probabilities (the algebraic range is $\mathbb{R}$, because the group generated by (-1) and $\sqrt{2}$ is dense). Shepp [183] was the first to consider non-lattice LLT in such cases. His approach was to provide asymptotic formulas for $\mathbb{P}[a \le S_n - \mathbb{E}(S_N) \le b]$ for arbitrary intervals $[a, b]$, or for $\sqrt{2\pi \mathrm{Var}(S_N)}\mathbb{E}[\phi(S_N - \mathbb{E}(S_N))]$ for all test functions $\phi : \mathbb{R} \to \mathbb{R}$ which are continuous with compact support. In this monograph, we use a slight modification of Shepp's formulation of the LLT. Instead of working with $S_N - \mathbb{E}(S_N)$, we work with $S_N - z_N$ subject to the assumptions that z_N is "not too far" from $\mathbb{E}(S_N)$, and that $S_N - z_N \in$ algebraic range.

Stone proved non-lattice LLT in Shepp's sense for sums of vector valued iid in [191], extending earlier work of Rvačeva [174], who treated the lattice case. These works are important not only because of the intrinsic interest in the vector valued case, but also because of technical innovations which became tools of the trade, see e.g. §5.2.1 and [17].

Local Limit Theorems for Other Stationary Stochastic Processes. The earliest LLT for non-iid sequences $\{X_i\}$ is due to Kolmogorov [116]. He considered stationary homogeneous Markov chains $\{X_i\}$ with a finite set of states $\mathfrak{S} = \{a_1, \ldots, a_n\}$, and proved a local limit theorem for the occupation times $S_N = \sum_{i=1}^{N} \mathbf{f}(X_i)$, where $\mathbf{f}(x) = (1_{a_1}(x), \ldots, 1_{a_n}(x))$.

Following further developments for finite state Markov chains by Sirazhdinov [185], Nagaev [149] proved very general CLT and LLT for $S_N = \sum_{i=1}^{N} f(X_i)$ for a large class of stationary homogeneous countable Markov chains $\{X_i\}$, and for a variety of unbounded functions f, both in the gaussian and in the stable case. See Chapter 8.

Nagaev's paper introduced the method of characteristic function operators (which we call in this work "Nagaev perturbation operators"), and opened the way for proving LLT for other weakly dependent stationary stochastic processes, and in particular to time series of probability preserving dynamical systems. Guivarc'h & Hardy [88] proved gaussian LLT for Birkhoff sums $S_N = \sum_{n=1}^{N} f(T^n x)$ for Anosov diffeomorphisms $T : X \to X$ with an invariant Gibbs measure, and Hölder continuous functions f. Rosseau-Egele [168] and Broise [20] proved such theorems for piecewise expanding interval map possessing an absolutely continuous invariant measure, $X = [0, 1]$, and $f \in BV$. Aaronson & Denker [5] gave general LLT for stationary processes generated by Gibbs-Markov maps both in the gaussian and in the non-gaussian domain of attraction.

These results have found many applications in infinite ergodic theory, dynamical systems and hyperbolic geometry, see for example [2], [3], [6].

The influence of Nagaev's method can also be recognized in other works on other asymptotic problems in dynamics and geometry, see for example [93], [107], [123], [126], [127], [159], [160], [182].

For the connection between the LLT and the behavior of local times for stationary stochastic processes, see [44, 61].

Kosloff and Volný showed that every ergodic and aperiodic probability preserving system has an observable whose Birkhoff sums satisfy the lattice LLT [121].

Local Limit Theorems for Non-Stationary Stochastic Processes. The interest in limit theorems for sums of *non*-identically distributed, independent, random variables goes back to the works of Chebyshev [196], Lyapunov [136], and Lindeberg [132] who considered the central limit theorem for such sums.

The study of LLT for sums of independent non-identically distributed random variables started later, in the works of Prokhorov [163] and Rozanov [169]. A common theme in these works is to *assume* an asymptotic formula for $\mathbb{P}[a \le \frac{S_N - A_N}{B_N} \le b]$ for suitable normalizing constants A_N, B_N, and then to ask what extra conditions imply an asymptotic for $\mathbb{P}[a \le S_N - A_N \le b]$.

An important counterexample by Gamkrelidze [76] pointed the way towards the following phenomenon: The distribution of S_N may lie close to a proper sub-group of its algebraic range without actually charging it, and a variety of sufficient conditions which rule this out were developed over the years. We mention especially the Prokhorov condition in the lattice case [169] (see §8.2), the Mineka-Silverman condition in the non-lattice case [144], Statulevičius's condition [190], and conditions motivated by additive number theory such as those appearing in [146] and [147]. For a discussion of these conditions, see [148].

Dolgopyat proved a LLT for sums of non-identically distributed, independent random variables which also applies to the reducible case [56].

Dobrushin proved a general central limit theorem for inhomogeneous Markov chains in [50] (see Chapter 3). Local limit theorems for inhomogeneous Markov chains are considered in [189]. Merlevède, M. Peligrad and C. Peligrad proved local limit theorems for sums $\sum_{i=1}^{N} f_i(X_i)$ where $\{X_i\}$ is a ψ-mixing inhomogeneous Markov chain, under the irreducibility condition of Mineka & Silverman [142]. Hafouta obtained local limit theorems for a class of inhomogeneous Markov chains in [90]. In a different direction, central limit theorems for time-series of inhomogeneous sequences of Anosov diffeomorphisms are proved in [12] and [30].

An important source of examples of inhomogeneous Markov chains is a Markov chain in random environment, when considered for a specific ("quenched") realizations of the environment (see Chapter 9). Hafouta & Kifer proved local limit theorems for non-conventional ergodic sums in [92], and local limit theorems for random dynamical systems including Markov chains in random environment in [93]. Demers, Péne & Zhang [41] prove a LLT for an integer valued observable for a random dynamical system.

Comparing the theory of inhomogeneous Markov chains to theory of Markov chains in random environment studied in [93], we note the following differences (see Chapter 9 for more discussion of this subject):

(a) The theory of inhomogeneous Markov chains applies to fixed realizations of noise and not just to all realizations in an unspecified set of full measure.
(b) In the random environment setup, a center–tight additive functional must be a coboundary plus a constant, while in the general case it can also have a component with summable variances.
(c) In the non center-tight random environment setup, the variance grows linearly for a.e. realization of noise. But for a general inhomogeneous Markov chain it can grow arbitrarily slowly.

The Contribution of This Work. The novelty of this work is in providing optimal sufficient conditions for the classical asymptotic formulas for $\mathbb{P}[S_N - z_N \in (a,b)]$, and in the analysis of $\mathbb{P}[S_N - z_N \in (a,b)]$ when these conditions fail.

In particular, we provide simple way to see when the obstructions to the LLT occur (based on structure constants $d_n(\xi)$), we derive a new asymptotic formula for $\mathbb{P}[S_N - z_N \in (a,b)]$ in the reducible case, when $\mathrm{Var}(S_N) \to \infty$, and we prove a structure theorem for S_N in case $\mathrm{Var}(S_N) \not\to \infty$. Unlike previous works, our analysis does not require any assumptions on the rate of growth of $\mathrm{Var}(S_N)$, beyond convergence to infinity.

Chapter 2
Markov Arrays, Additive Functionals, and Uniform Ellipticity

Abstract This chapter presents the main objects of our study. We define Markov arrays and additive functionals, discuss the uniform ellipticity condition, and introduce the structure constants.

2.1 The Basic Setup

2.1.1 Inhomogeneous Markov Chains

A **Markov chain** is given by the following data:

- **State Spaces:** Borel spaces $(\mathfrak{S}_n, \mathscr{B}(\mathfrak{S}_n))$ $(n \geq 1)$, where $\mathfrak{S}_n$ is a complete separable metric space, and $\mathscr{B}(\mathfrak{S}_n)$ is the Borel σ-algebra of $\mathfrak{S}_n$. $\mathfrak{S}_n$ is the set of "the possible states of the Markov chain at time n."
- **Transition Probabilities** (or **Transition Kernels**): a family of Borel probability measures $\pi_{n,n+1}(x, \mathrm{d}y)$ on $\mathfrak{S}_{n+1}$ $(x \in \mathfrak{S}_n, n \geq 1)$, so that for every Borel $E \subset \mathfrak{S}_{n+1}$, the function $x \mapsto \pi_{n,n+1}(x, E)$ is measurable. The measure $\pi_{n,n+1}(x, E)$ is "the probability of the event E at time $n+1$, given that the state at time n was x."
- **Initial Distribution:** $\pi(\mathrm{d}x)$, a Borel probability measure on $\mathfrak{S}_1$. $\pi(E)$ is "the probability that the state x at time 1 belongs to E."

The **Markov chain** associated with this data is the Markov process $\mathsf{X} := \{X_n\}_{n\geq 1}$ such that $X_n \in \mathfrak{S}_n$ for all n, and so that for all Borel $E_i \subset \mathfrak{S}_i$, $\mathbb{P}(X_1 \in E_1) = \pi(E_1)$, $\mathbb{P}(X_{n+1} \in E_{n+1}|X_n = x_n) = \pi_{n,n+1}(x_n, E_{n+1})$.

X is uniquely defined, and its joint distribution is given by

$$\mathbb{P}(X_1 \in E_1, \cdots, X_n \in E_n) := \int_{E_1} \pi(\mathrm{d}x_1) \int_{E_2} \pi_{1,2}(x_1, \mathrm{d}x_2) \cdots \int_{E_n} \pi_{n-1,n}(x_{n-1}, \mathrm{d}x_n). \tag{2.1}$$

Let $\mathbb{P}, \mathbb{E}$ and Var denote the probability, expectation, and variance calculated using this joint distribution. If π is the point mass at x, we write $\mathbb{P}_x$ and $\mathbb{E}_x$.

X satisfies the following important **Markov property** (see e.g. [17, Ch. 7]):

$$\mathbb{P}(X_{k+1} \in E|X_k, X_{k-1}, \cdots, X_1) = \mathbb{P}(X_{k+1} \in E|X_k) = \pi_{k,k+1}(X_k, E), \tag{2.2}$$

$$\begin{aligned}\mathbb{P}(X_n \in E_n, \cdots, X_{k+1} \in E_{k+1}|X_k, \cdots, X_1) &= \mathbb{P}(X_n \in E_n, \cdots, X_{k+1} \in E_{k+1}|X_k) \\ &= \int_{E_{k+1}} \pi_{k,k+1}(X_k, \mathrm{d}x_{k+1}) \cdots \int_{E_n} \pi_{n-1,n}(x_{n-1}, \mathrm{d}x_n) \text{ for all } n \geq k+1.\end{aligned} \tag{2.3}$$

The proofs are a direct calculation. Let $\mathscr{F}_{k,\infty}$ denote the σ-algebra generated by X_i with $i \geq k$. Then an approximation argument shows that for each $A \in \mathscr{F}_{k,\infty}$,

$$\mathbb{P}\big((X_{k+1}, X_{k+2}, \cdots) \in A|X_k, \cdots, X_1\big) = \mathbb{P}\big((X_{k+1}, X_{k+2}, \cdots) \in A|X_k\big). \tag{2.4}$$

If the state spaces and the transition probabilities do not depend on n, that is, $\mathfrak{S}_n = \mathfrak{S}$ and $\pi_{n,n+1}(x, \mathrm{d}y) = \pi(x, \mathrm{d}y)$, then we call X a **homogeneous** Markov chain. Otherwise, X is called an **inhomogeneous** Markov chain. In this work, we are mainly interested in the inhomogeneous case.

Example 2.1 (Markov Chains with Finite State Spaces) These are Markov chains X with state spaces $\mathfrak{S}_n = \{1, \ldots, d_n\}$, $\mathscr{B}(\mathfrak{S}_n) = \{\text{all subsets of } \mathfrak{S}_n\}$. In this case the transition probabilities are completely characterized by the rectangular stochastic matrices with entries $\pi^n_{xy} := \pi_{n,n+1}(x, \{y\})$ $(x = 1, \ldots, d_n\ ;\ y = 1, \ldots, d_{n+1})$. The initial distribution is completely characterized by the probability vector $\pi_x := \pi(\{x\})$ $(x = 1, \ldots, d_n)$.

Then $\mathbb{P}(X_1{=}x_1, \cdots, X_n = x_n){=}\pi_{x_1}\pi^1_{x_1x_2}\pi^2_{x_2x_3}\cdots\pi^{n-1}_{x_{n-1}x_n}$. This leads to the following discrete version of (2.1):

$$\mathbb{P}(X_1 \in E_1, \cdots, X_n \in E_n) = \sum_{x_1\in E_1} \pi_{x_1} \sum_{x_2\in E_2} \pi^1_{x_1x_2} \cdots \sum_{x_n\in E_n} \pi^{n-1}_{x_{n-1}x_n}.$$

Example 2.2 (Markov Chains in Random Environment) Let X denote a homogeneous Markov chain with state space $\mathfrak{S}$, transition probability $\pi(x, \mathrm{d}y)$, and initial distribution concentrated at a point x_1. It is possible to view X as a model for the motion of a particle on $\mathfrak{S}$ as follows. At time 1, the particle is located at x_1, and a particle at position x will jump after one time step to a random location y, distributed like $\pi(x, \mathrm{d}y)$: $\mathbb{P}(y \in E) = \pi(x, E)$. With this interpretation, $X_n =$ the position of the particle at time n. The homogeneity of X is reflected in the fact that the law of motion which governs the jumps does not change in time.

Let us now refine the model, and add a dependence of the transition probabilities on an external parameter ω, which we think of as "the environment." For example, ω could represent a external force field which affects the likelihood of various movements, and which can be modified by God or some other experimentalist. The transition probabilities become $\pi(x, \omega, \mathrm{d}y)$.

Suppose the environment ω changes in time according to some deterministic rule. This is modeled by a map $T : \Omega \to \Omega$, where Ω is the collection of all possible states of the environment, and T is a deterministic law of motion which says that an environment at state ω will evolve after one unit of time to the state $T(\omega)$. Iterating we see that if the initial state of the environment at time zero was ω, then its state at time n will be $\omega_n = T^n(\omega) = (T \circ \cdots \circ T)(\omega)$.

Returning to our particle, we see that if the initial condition of the environment at time zero is ω, then the transition probabilities at time n are $\pi^{\omega}_{n,n+1}(x, \mathrm{d}y) = \pi(x, T^n(\omega), \mathrm{d}y)$.

Thus each $\omega \in \Omega$ gives rise to an inhomogeneous Markov chain X^{ω}, which describes the Markovian dynamics of a particle, coupled to a changing environment.

If $T(\omega) = \omega$, the environment stays fixed, and the Markov chain is homogeneous, otherwise the Markov chain is inhomogeneous. We will return to Markov chains in random environment in chapter 9.

Example 2.3 (Markov Chains with Finite Memory) We can weaken the Markov property (2.2) by specifying that for some fixed $k_0 \geq 1$, for all $E \in \mathscr{B}(\mathfrak{S}_{n+1})$, $\mathbb{P}(X_{n+1} \in E|X_n, \ldots, X_1) = \begin{cases} \mathbb{P}(X_{n+1} \in E|X_n, \ldots, X_{n-k_0+1}) & n > k_0; \\ \mathbb{P}(X_{n+1} \in E|X_n, \ldots, X_1) & n \leq k_0. \end{cases}$

Stochastic processes like that are called "Markov chains with finite memory" (of length k_0). Markov chains with memory of length 1 are ordinary Markov chains. Markov chains with memory of length $k_0 > 1$ can be recast as ordinary Markov chains by considering the stochastic process $\widetilde{\mathsf{X}} = \{(X_n, \ldots, X_{n+k_0-1})\}_{n\geq 1}$ with its natural state spaces, initial distribution, and transition kernels.

Example 2.4 (A Non-Example) Every inhomogeneous Markov chain X can be presented as a homogeneous Markov chain Y, but this is not very useful.

To obtain such a representation, recall that the state spaces of X are complete separable metric spaces $\mathfrak{S}_i$. As such, they are Borel isomorphic to $\mathbb{R}$, or to $\mathbb{Z}$, or to a finite set, or to a union of the above sets (see e.g. [187], §3). In any case the state spaces can be embedded in a Borel way into $\mathbb{R}$. Fix some Borel bi-measurable injections $\varphi_i : \mathfrak{S}_i \hookrightarrow \mathbb{R}$. Let $Y_n = (\varphi_n(X_n), n)$. This is a new presentation of X.

We claim that Y is a homogeneous Markov chain.

Let δ_ξ denote the Dirac measure at ξ, defined by $\delta_\xi(E) := 1$ when $E \ni \xi$ and $\delta_\xi(E) := 0$ otherwise. Let $\mathfrak{S}_n, \pi_{n,n+1}$ and π denote the states spaces, transition probabilities, and initial distribution of X. Let Z be the *homogeneous* Markov chain with state space $\mathfrak{S} := \mathbb{R} \times \mathbb{N}$, initial distribution $\widehat{\pi} := (\pi \circ \varphi_1^{-1}) \times \delta_1$ (a measure on $\mathfrak{S}_1 \times \{1\}$), and transition probabilities $\widehat{\pi}\big((x, n), A \times B\big) = \pi_{n,n+1}\big(\varphi_n^{-1}(x), \varphi_{n+1}^{-1}(A)\big)\delta_{n+1}(B)$, for $x \in \varphi_n(\mathfrak{S}_n)$. A direct calculation shows that the joint distribution Z is equal to the joint distribution of $\mathsf{Y} = \{(\varphi_n(X_n), n)\}_{n\geq 1}$. So Y is a homogeneous Markov chain.

Such presentations will not be useful to us, because they destroy useful structures which are essential for our work on the local limit theorem. For example, they destroy the uniform ellipticity property, that we will discuss in §2.2 below.

2.1.2 Inhomogeneous Markov Arrays

For technical reasons that we will explain later, it is useful to consider a generalization of a Markov chain, called a **Markov array**. To define a Markov array, we need the following data:

- **Row Lengths:** $k_N + 1$ where $k_N \geq 1$ and $(k_N)_{N\geq 1}$ is strictly increasing.
- **State Spaces:** $(\mathfrak{S}_n^{(N)}, \mathscr{B}(\mathfrak{S}_n^{(N)}))$, $(1 \leq n \leq k_N + 1)$, where $\mathfrak{S}_n^{(N)}$ is a complete separable metric space, and $\mathscr{B}(\mathfrak{S}_n^{(N)})$ is its Borel σ-algebra.
- **Transition Probabilities** (or **Transition Kernels):** $\{\pi_{n,n+1}^{(N)}(x, \mathrm{d}y)\}_{x\in\mathfrak{S}_n^{(N)}}$ $(1 \leq n \leq k_N)$, where $\pi_{n,n+1}^{(N)}$ are Borel probability measures on $\mathfrak{S}_{n+1}^{(N)}$, so that for every Borel $E \subset \mathfrak{S}_{n+1}^{(N)}$, the function $x \mapsto \pi_{n,n+1}^{(N)}(x, E)$ is measurable.
- **Initial Distributions:** Borel probability measures $\pi^{(N)}(\mathrm{d}x)$ on $\mathfrak{S}_1^{(N)}$.

For each $N \geq 1$, this data determines a finite Markov chain of length $k_N + 1$: $\mathsf{X}^{(N)} = (X_1^{(N)}, X_2^{(N)}, \dots, X_{k_N+1}^{(N)})$ called the ***N*-th row** of the array. These rows can be arranged in a triangular array

$$\mathsf{X} = \begin{cases} X_1^{(1)}, \dots, X_{k_1+1}^{(1)} \\ X_1^{(2)}, \dots, X_{k_1+1}^{(2)}, \dots, X_{k_2+1}^{(2)} \\ X_1^{(3)}, \dots, X_{k_1+1}^{(3)}, \dots, X_{k_2+1}^{(3)}, \dots, X_{k_3+1}^{(3)} \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{cases}$$

Each row $\mathsf{X}^{(N)}$ comes equipped with a joint distribution, which depends on N. But *no joint distribution on elements of different rows is specified.*

We will continue to denote the joint probability distribution, expectation, and variance of $\mathsf{X}^{(N)}$ by $\mathbb{P}, \mathbb{E}$, and Var. These objects depend on N, but the index N can be suppressed, because it is always obvious from the context. As always, in cases when we wish to condition on the initial state $X_1^{(N)} = x$, we will write $\mathbb{P}_x$ and $\mathbb{E}_x$.

Example 2.5 (Markov Chains as Markov Arrays) Every Markov chain $\{X_n\}$ gives rise to a Markov array with row lengths $k_N = N + 1$ and rows $\mathsf{X}^{(N)} = (X_1, \dots, X_{N+1})$. In this case $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$, $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$, and $\pi^{(N)} = \pi$.

Conversely, any Markov array so that $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$, $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$, and $\pi^{(N)} = \pi$ determines a Markov chain with state spaces $\mathfrak{S}_n$, transition probabilities $\pi_{n,n+1}^{(N)} = \pi_{n,n+1}$, and initial distributions $\pi^{(N)} = \pi$.

Example 2.6 (Change of Measure) Suppose $\{X_n\}_{n\geq 1}$ is a Markov chain with data $\mathfrak{S}_n, \pi_{n,n+1}, \pi$, and let $\varphi_n^{(N)}(x, y)$ be a family of positive measurable functions on $\mathfrak{S}_n \times \mathfrak{S}_{n+1}$ so that $\int \varphi_n^{(N)}(x, y)\pi_{n,n+1}(x, \mathrm{d}y) < \infty$ for all x, n and N. Let

$$\pi_n^{(N)}(x, \mathrm{d}y) := \frac{\varphi_n^{(N)}(x, y)}{\int \varphi_n^{(N)}(x, y)\pi_{n,n+1}(x, \mathrm{d}y)}\pi_{n,n+1}(x, \mathrm{d}y).$$

Then the data $k_N = N + 1$, $\mathfrak{S}_n^{(N)} := \mathfrak{S}_n$, $\pi^{(N)} := \pi$ and $\pi_{n,n+1}^{(N)}$ determines a Markov array called the **change of measure of $\{X_n\}$ with weights $\varphi_n^{(N)}$**.

Why study Markov arrays? There are several reasons, and the one most pertinent to this work is the following: The theory of large deviations for Markov *chains*, relies on a change of measure which results in Markov *arrays*. Thus, readers who are only interested in local limit theorems for Markov chains in the local regime $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\mathrm{Var}(S_N)}} \to z$, may ignore the theory of arrays and limit their attention to Markov chains. But those who are also interested in the large deviations regime, where $\left|\frac{z_N - \mathbb{E}(S_N)}{\mathrm{Var}(S_N)}\right|$ is of order 1, will need the theory for Markov arrays.

2.1.3 Additive Functionals

An **additive functional of a Markov chain** with state spaces $\mathfrak{S}_n$ is a sequence of measurable functions $f_n : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \to \mathbb{R}$. The pair $\mathsf{X} = \{X_n\}, \mathsf{f} = \{f_n\}$ determines a stochastic process

$$S_N = f_1(X_1, X_2) + f_2(X_2, X_3) + \cdots + f_N(X_n, X_{N+1}) \quad (N \geq 1).$$

We will often abuse terminology and call $\{S_N\}_{N\geq 1}$ an "additive functional" of X.

An **additive functional of a Markov array** X with row lengths $k_N + 1$ and state spaces $\mathfrak{S}_n^{(N)}$ is an array of measurable functions $f_n^{(N)} : \mathfrak{S}_n^{(N)} \times \mathfrak{S}_{n+1}^{(N)} \to \mathbb{R}$ with row lengths k_N:

$$\mathsf{f} = \begin{cases} f_1^{(1)}, \ldots, f_{k_1}^{(1)} \\ f_1^{(2)}, \ldots, f_{k_1}^{(2)}, \ldots, f_{k_2}^{(2)} \\ f_1^{(3)}, \ldots, f_{k_1}^{(3)}, \ldots, f_{k_2}^{(3)}, \ldots, f_{k_3}^{(3)} \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

This determines a sequence of random variables $S_N = f_1^{(N)}(X_1^{(N)}, X_2^{(N)}) + f_2^{(N)}(X_2^{(N)}, X_3^{(N)}) + \cdots + f_{k_N}^{(N)}(X_{k_N}^{(N)}, X_{k_N+1}^{(N)})$, $N \geq 1$, which we also refer to as an "additive functional." But be careful! *This is not a stochastic process, because no joint distribution of $S_1, S_2, \ldots$ is specified.*

Suppose f, g are two additive functionals on X. For Markov chains X, we define, $\mathsf{f} + \mathsf{g} := \{f_n + g_n\}$, $c\mathsf{f} := \{cf_n\}$, $\quad |\mathsf{f}| := \sup_n \left(\sup_{x,y} |f_n(x, y)| \right)$ and $\text{ess sup} |\mathsf{f}| := \sup_n (\text{ess sup} |f_n(X_n, X_{n+1})|)$.

Similarly, if X is a Markov array with row lengths $k_N + 1$, then we set

$$\mathsf{f} + \mathsf{g} := \{f_n^{(N)} + g_n^{(N)}\}, \quad c\mathsf{f} := \{cf_n^{(N)}\}, \quad |\mathsf{f}| := \sup_N \sup_{1\leq n\leq k_N} \left(\sup_{x,y} |f_n^{(N)}(x, y)| \right),$$

and $\text{ess sup} |\mathsf{f}| := \sup_N \sup_{1\leq n\leq k_N} \left(\text{ess sup} |f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})| \right)$.

The notation $|\mathsf{f}| \leq K$ a.s. will mean that $\text{ess sup} |\mathsf{f}| \leq K$ ("a.s." stands for "almost surely"). An additive functional is called **uniformly bounded** if there is a constant K such that $|\mathsf{f}| \leq K$, and **uniformly bounded a.s.** if $\exists K$ such that $|\mathsf{f}| \leq K$ a.s.

2.2 Uniform Ellipticity

2.2.1 The Definition

A Markov chain X with state spaces $\mathfrak{S}_n$ and transition probabilities $\pi_{n,n+1}(x, dy)$ is called **uniformly elliptic**, if there are Borel probability measures μ_n on $\mathfrak{S}_n$, Borel measurable functions $p_n : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \to [0, \infty)$, and an **ellipticity constant** $0 < \epsilon_0 < 1$ such that for all $n \geq 1$,

(a) $\pi_{n,n+1}(x, dy) = p_n(x, y)\mu_{n+1}(dy)$;
(b) $0 \leq p_n \leq 1/\epsilon_0$;
(c) $\int_{\mathfrak{S}_{n+1}} p_n(x, y)p_{n+1}(y, z)\mu_{n+1}(dy) > \epsilon_0$.

We call $\mu_{n+1}(dy)$ **background measures**. Corollary 2.9 below says that if X is uniformly elliptic with some background measures, then it is uniformly elliptic with respect to the "natural" background measures $\mu_n(E) = \mathbb{P}(X_n \in E)$ $(n \geq 3)$.

The integral in (c) is the two-step transition probability $\mathbb{P}(X_{n+2} = z|X_n = x)$, and we will sometime call (c) a **two-step ellipticity condition**. For more general **γ-step ellipticity conditions**, see §2.4.

Example 2.7 (Doeblin Chains) These are Markov chains X with state spaces $\mathfrak{S}_n$ of bounded cardinality $|\mathfrak{S}_n| \leq M < \infty$, and with transition probabilities $\pi^n_{xy} := \pi_{n,n+1}(x, \{y\})$ such that

(1) $\exists \epsilon'_0 > 0$ s.t. for all $n \geq 1$ and $(x, y) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$, either $\pi^n_{xy} = 0$ or $\pi^n_{xy} > \epsilon'_0$;
(2) for all n, for all $(x, z) \in \mathfrak{S}_n \times \mathfrak{S}_{n+2}$, there exists $y \in \mathfrak{S}_{n+1}$ such that $\pi^n_{xy}\pi^{n+1}_{yz} > 0$.

Doeblin chains are uniformly elliptic: Take μ_n to be the uniform measure on $\mathfrak{S}_n$ and $p_n(x, y) := \pi^n_{xy}/|\mathfrak{S}_{n+1}|$. Then (a) is clear, (b) holds with any $\epsilon_0 < 1/M$, and (c) holds with $\epsilon_0 := (\epsilon'_0/M)^2$. Doeblin chains are named after W. Doeblin, who studied homogeneous countable Markov chains satisfying similar conditions.

Here is the formulation of the uniform ellipticity conditions for Markov arrays. A Markov array X with state spaces $\mathfrak{S}^{(N)}_n$, transition probabilities $\pi^{(N)}_{n,n+1}(x, \mathrm{d}y)$, and row lengths $k_N + 1$ is called **uniformly elliptic**, if there exist Borel probability measures $\mu^{(N)}_n$ on $\mathfrak{S}^{(N)}_n$, Borel measurable functions $p^{(N)}_n : \mathfrak{S}^{(N)}_n \times \mathfrak{S}^{(N)}_{n+1} \to [0, \infty)$, and a constant $0 < \epsilon_0 < 1$ as follows: For all $N \geq 1$ and $1 \leq n \leq k_N$,
(a) $\pi^{(N)}_{n,n+1}(x, \mathrm{d}y) = p^{(N)}_n(x, y)\mu^{(N)}_{n+1}(\mathrm{d}y)$; (b) $0 \leq p^{(N)}_n \leq 1/\epsilon_0$; (c) $\int_{\mathfrak{S}_{n+1}} p^{(N)}_n(x, y)p^{(N)}_{n+1}(y, z)\mu^{(N)}_{n+1}(\mathrm{d}y) > \epsilon_0$.

Proposition 2.8 *If a Markov array X is uniformly elliptic with background measures $\mu^{(N)}_n$ and ellipticity constant ϵ_0, then for every $3 \leq n \leq k_N + 1 < \infty$,*

$$\epsilon_0 \leq \frac{\mathbb{P}(X^{(N)}_n \in E)}{\mu^{(N)}_n(E)} \leq \epsilon_0^{-1} \qquad (E \in \mathscr{B}(\mathfrak{S}^{(N)}_N)).$$

Proof We fix a row N, and drop the superscripts $^{(N)}$. Define a probability measure on $\mathfrak{S}_n$ by $P_n(E) = \mathbb{P}(X_n \in E)$, then for every $1 \leq n < k_N$, for every bounded measurable $\varphi : \mathfrak{S}_{n+2} \to \mathbb{R}$,

$$\int \varphi \,\mathrm{d}P_{n+2} = \mathbb{E}(\varphi(X_{n+2})) = \mathbb{E}\Big[\mathbb{E}\Big(\mathbb{E}(\varphi(X_{n+2})|X_{n+1}, X_n)\Big|X_n\Big)\Big] = \mathbb{E}[\mathbb{E}(\mathbb{E}(\varphi(X_{n+2})|X_{n+1})|X_n)] \ \ (\because \text{Markov property})$$
$$= \iiint \varphi(z)\, \pi_{n+1,n+2}(y, \mathrm{d}z)\, \pi_{n,n+1}(x, \mathrm{d}y)P_n(\mathrm{d}x).$$

So $\int \varphi \mathrm{d}P_{n+2} = \int \varphi(z) \left[\int\!\!\left(\int p_{n+1}(y, z)p_n(x, y)\mu_{n+1}(\mathrm{d}y)\right) P_n(\mathrm{d}x)\right] \mu_{n+2}(\mathrm{d}z)$. The quantity in the square brackets is bounded below by ϵ_0 and bounded above by ϵ_0^{-1}. So the measures P_{n+2}, μ_{n+2} are equivalent, and $\epsilon_0 \leq \frac{\mathrm{d}P_{n+2}}{\mathrm{d}\mu_{n+2}} \leq \epsilon_0^{-1}$. □

Corollary 2.9 *If a Markov array X is uniformly elliptic, then there are $\epsilon_0 > 0$ and $p^{(N)}_n(x, y)$ so that the uniform ellipticity conditions (a),(b) and (c) hold with the background measures $\mu^{(N)}_n(E) := \mathbb{P}[X^{(N)}_n \in E]$ for $n \geq 3$.*

Proof If X is uniformly elliptic with background measures $\mu^{(N)}_n$, then it is also uniformly elliptic with any other choice of background measures $\widehat{\mu}^{(N)}_n$ so that $C^{-1} \leq \mathrm{d}\widehat{\mu}^{(N)}_n/\mathrm{d}\mu^{(N)}_n \leq C$, with C positive and independent of n and N. The corollary follows from the previous proposition. □

Corollary 2.10 *Let X be a uniformly elliptic Markov chain with ellipticity constant ϵ_0, and suppose Y is a Markov array obtained from X by the change of measure construction described in Example 2.6. If the weights satisfy $C^{-1} \leq \varphi^{(N)}_n(x, y) \leq C$ for all n and N, then Y is uniformly elliptic with ellipticity constant ϵ_0/C^4, and $\exists M = M(\epsilon_0, C) > 1$ such that for all N and $3 \leq n \leq N$,*

$$M^{-1} \leq \frac{\mathbb{P}[Y^{(N)}_n \in E]}{\mathbb{P}[X_n \in E]} \leq M. \tag{2.5}$$

Proof Let $\pi_{n,n+1}(x, \mathrm{d}y)$ be the transition probabilities of X. By assumption, $\pi_{n,n+1}(x, \mathrm{d}y) = p_n(x, y)\mu_{n+1}(\mathrm{d}y)$ where $p_n(x, y)$ satisfies the uniform ellipticity conditions. Then the transition probabilities of Y are given by

$$q_n^{(N)}(x,y)\mu_{n+1}(\mathrm{d}y),\ \text{ where } q_n^{(N)}(x,y) := \frac{p_n(x,y)\varphi_n^{(N)}(x,y)}{\int \varphi_n^{(N)}(x,z)\pi_{n,n+1}(x,\mathrm{d}z)}. \tag{2.6}$$

Since $q_n^{(N)}(x,y) = C^{\pm 2} p_n(x,y)$, Y is uniformly elliptic with ellipticity constant ϵ_0/C^4. (2.5) follows from (2.6) and Proposition 2.8. □

Caution! The Radon-Nikodym derivative of the *joint distributions* of $(X_3,\dots,X_N)$ and $(Y_3^{(N)},\dots,Y_N^{(N)})$ need not be uniformly bounded as $N\to\infty$.

2.2.2 Contraction Estimates and Exponential Mixing

Suppose $\mathfrak{X},\mathfrak{Y}$ are complete and separable metric spaces.

A **transition kernel** from $\mathfrak{X}$ to $\mathfrak{Y}$ is a family $\{\pi(x,\mathrm{d}y)\}_{x\in\mathfrak{X}}$ of Borel probability measures on $\mathfrak{Y}$ so that $x\mapsto \pi(x,E)$ is measurable for all $E\subset\mathfrak{Y}$ Borel. A transition kernel $\{\pi(x,\mathrm{d}y)\}_{x\in\mathfrak{X}}$ determines two **Markov operators**, one acting on measures and the other acting on functions. The action on measures takes a probability measure μ on $\mathfrak{X}$ and maps it to a probability measure on $\mathfrak{Y}$ via $\pi(\mu)(E) := \int_{\mathfrak{X}} \pi(x,E)\mu(\mathrm{d}x)$. The action on functions takes a bounded Borel function $u:\mathfrak{Y}\to\mathbb{R}$ and maps it to a bounded Borel function on $\mathfrak{X}$ via $\pi(u)(x) = \int_{\mathfrak{Y}} u(y)\pi(x,\mathrm{d}y)$. We have a duality:

$$\int u(y)\,\pi(\mu)(\mathrm{d}y) = \int \pi(u)(x)\,\mu(\mathrm{d}x). \tag{2.7}$$

Define the **oscillation** of a function $u:\mathfrak{Y}\to\mathbb{R}$ to be

$$\mathrm{Osc}(u) := \sup_{y_1,y_2\in\mathfrak{Y}} |u(y_1)-u(y_2)|. \tag{2.8}$$

The **contraction coefficient** of $\{\pi(x,\mathrm{d}y)\}_{x\in\mathfrak{X}}$ is

$$\delta(\pi) := \sup\{|\pi(x_1,E)-\pi(x_2,E)| : x_1,x_2\in\mathfrak{X},\quad E\in\mathscr{B}(\mathfrak{Y})\}.$$

The **total variation distance** between two probability measures μ_1,μ_2 on $\mathfrak{X}$ is

$$\|\mu_1-\mu_2\|_{\mathrm{Var}} := \sup\{|\mu_1(A)-\mu_2(A)| : A\subset\mathfrak{X} \text{ is measurable}\} \equiv \frac{1}{2}\sup\left\{\int w(x)(\mu_1-\mu_2)(\mathrm{d}x) \middle| w:\mathfrak{X}\to[-1,1] \text{ is measurable}\right\}.$$

Caution! $\|\mu_1-\mu_2\|_{\mathrm{Var}}$ is actually *one-half* of the total variation of $\mu_1-\mu_2$, because it is equal to $(\mu_1-\mu_2)^+(\mathfrak{X})$ and to $(\mu_1-\mu_2)^-(\mathfrak{X})$, but not to

$$|\mu|(\mathfrak{X}) = (\mu_1-\mu_2)^+(\mathfrak{X}) + (\mu_1-\mu_2)^-(\mathfrak{X}).$$

Lemma 2.11 (Contraction Lemma) *Suppose $\mathfrak{X},\mathfrak{Y}$ are complete and separable metric spaces, and $\{\pi(x,\mathrm{d}y)\}_{x\in\mathfrak{X}}$ is a transition kernel from $\mathfrak{X}$ to $\mathfrak{Y}$. Then:*

(a) $0\le\delta(\pi)\le 1$.
(b) $\delta(\pi) = \sup\{\mathrm{Osc}[\pi(u)] \mid u:\mathfrak{Y}\to\mathbb{R}$ *measurable, and* $\mathrm{Osc}(u)\le 1\}$.
(c) If $\mathfrak{Z}$ is a complete separable metric space, π_1 is a transition kernel from $\mathfrak{X}$ to $\mathfrak{Y}$, and π_2 is a transition kernel from $\mathfrak{Y}$ to $\mathfrak{Z}$, then $\delta(\pi_1\circ\pi_2)\le\delta(\pi_1)\delta(\pi_2)$.
(d) $\mathrm{Osc}[\pi(u)]\le\delta(\pi)\,\mathrm{Osc}(u)$ *for every $u:\mathfrak{Y}\to\mathbb{R}$ bounded and measurable.*
(e) $\|\pi(\mu_1)-\pi(\mu_2)\|_{\mathrm{Var}}\le\delta(\pi)\|\mu_1-\mu_2\|_{\mathrm{Var}}$ *for all Borel probability measures μ_1,μ_2 on $\mathfrak{X}$.*

(f) *Suppose λ is a probability measure on $\mathfrak{X}\times\mathfrak{Y}$ with marginals $\mu_{\mathfrak{X}}$, $\mu_{\mathfrak{Y}}$, and transition kernel $\{\pi(x,\mathrm{d}y)\}$, i.e. $\lambda(E\times\mathfrak{Y})=\mu_{\mathfrak{X}}(E)$, $\lambda(\mathfrak{X}\times E)=\mu_{\mathfrak{Y}}(E)$, and $\lambda(\mathrm{d}x,\mathrm{d}y)=\pi(x,\mathrm{d}y)\mu_{\mathfrak{X}}(\mathrm{d}x)$. Let $f\in L^2(\mu_{\mathfrak{X}})$, $g\in L^2(\mu_{\mathfrak{Y}})$ be two elements with zero integral. Then* $\left|\int_{\mathfrak{X}\times\mathfrak{Y}} f(x)g(y)\lambda(\mathrm{d}x,\mathrm{d}y)\right|\leq\sqrt{\delta(\pi)}\|f\|_{L^2(\mu_{\mathfrak{X}})}\|g\|_{L^2(\mu_{\mathfrak{Y}})}$.

(g) *Let λ, $\mu_{\mathfrak{X}}$, $\mu_{\mathfrak{Y}}$ and π be as in (f), and suppose $g\in L^2(\mu_{\mathfrak{Y}})$ has integral zero, then $\pi(g)\in L^2(\mu_{\mathfrak{X}})$ has integral zero, and* $\|\pi(g)\|_{L^2(\mu_{\mathfrak{X}})}\leq\sqrt{\delta(\pi)}\|g\|_{L^2(\mu_{\mathfrak{Y}})}$.

Proof (a) is trivial. The inequality $\leq$ in (b) is because for every $E\in\mathscr{B}(\mathfrak{Y})$, $u:=1_E$ satisfies $\mathrm{Osc}(u)\leq 1$. To see $\geq$, fix some $u:\mathfrak{Y}\to\mathbb{R}$ measurable such that $\mathrm{Osc}(u)\leq 1$. Suppose first that u is a "simple function:" a measurable function with finitely many values. Then we can write $u=c+\sum_{i=1}^{m}\alpha_i 1_{A_i}$ where $c\in\mathbb{R}$, $|\alpha_i|\leq\frac{1}{2}\mathrm{Osc}(u)$, and A_i measurable and pairwise disjoint. For every pair of points $x_1,x_2\in\mathfrak{X}$,

$$
\begin{aligned}
|\pi(u)(x_1)-\pi(u)(x_2)| &= \left|\sum_{i=1}^{m}\alpha_i[\pi(x_1,A_i)-\pi(x_2,A_i)]\right| \\
&\leq \left|\sum_{\pi(x_1,A_i)>\pi(x_2,A_i)}\alpha_i[\pi(x_1,A_i)-\pi(x_2,A_i)]\right|+\left|\sum_{\pi(x_1,A_i)<\pi(x_2,A_i)}\alpha_i[\pi(x_1,A_i)-\pi(x_2,A_i)]\right| \\
&\overset{!}{\leq} \frac{1}{2}\mathrm{Osc}(u)\delta(\pi)+\frac{1}{2}\mathrm{Osc}(u)\delta(\pi)=\delta(\pi)\,\mathrm{Osc}(u)\leq\delta(\pi)\quad(\because A_i \text{ are disjoint}).
\end{aligned}
$$

So $\mathrm{Osc}[\pi(u)]\leq\delta(\pi)$ for all simple functions u with $\mathrm{Osc}(u)\leq 1$.

It follows that $\mathrm{Osc}[\pi(u)]\leq\delta(\pi)$ for all measurable u s.t. $\mathrm{Osc}(u)\leq 1$. This proves (b).

Clearly, $(b)\Rightarrow(d)\Rightarrow(c)$. To see (e), we restrict to the non-trivial case $\mu_1\neq\mu_2$. Let $\mu:=\mu_1-\mu_2$, and decompose $\mu=\mu^+-\mu^-$ where $\mu^\pm$ are singular positive measures (this is the Jordan decomposition). Since $\mu(\mathfrak{X})=0$, μ^+ and μ^- have equal total mass, and $\mu^\pm(\mathfrak{X})=\frac{1}{2}(\mu^+(\mathfrak{X})+\mu^-(\mathfrak{X}))=\frac{1}{2}|\mu|(\mathfrak{X})\equiv\|\mu_1-\mu_2\|_{\mathrm{Var}}$. Let

$$
\widehat{\mu}_1:=\mu^+/\|\mu_1-\mu_2\|_{\mathrm{Var}}\,,\ \widehat{\mu}_2:=\mu^-/\|\mu_1-\mu_2\|_{\mathrm{Var}}\,,\ \widehat{\mu}:=\widehat{\mu}_1-\widehat{\mu}_2=\frac{\mu_1-\mu_2}{\|\mu_1-\mu_2\|_{\mathrm{Var}}}.
$$

Note that $\widehat{\mu}_1$ and $\widehat{\mu}_2$ are probability measures.

For every non-constant measurable function $w:\mathfrak{Y}\to[-1,1]$,

$$
\begin{aligned}
&\frac{\frac{1}{2}\int_{\mathfrak{Y}} w(y)\pi(\mu)(\mathrm{d}y)}{\|\mu_1-\mu_2\|_{\mathrm{Var}}}=\frac{1}{2}\int_{\mathfrak{Y}} w(y_1)\pi(\widehat{\mu}_1)(\mathrm{d}y_1)-\int_{\mathfrak{Y}} w(y_2)\pi(\widehat{\mu}_2)(\mathrm{d}y_2) \\
&=\frac{1}{2}\int_{\mathfrak{X}}\pi(w)(x_1)\widehat{\mu}_1(\mathrm{d}x_1)-\int_{\mathfrak{X}}\pi(w)(x_2)\widehat{\mu}_2(\mathrm{d}x_2),\ \text{see (2.7)} \\
&=\frac{1}{2}\int_{\mathfrak{X}}\int_{\mathfrak{X}}[\pi(w)(x_1)-\pi(w)(x_2)]\widehat{\mu}_1(\mathrm{d}x_1)\widehat{\mu}_2(\mathrm{d}x_2),\ \text{because } \widehat{\mu}_i(\mathfrak{X})=1, \\
&\leq\frac{1}{2}\delta(\pi)\,\mathrm{Osc}(w)\leq\delta(\pi),\ \text{by (b) and because } \mathrm{Osc}(w)\leq 2\|w\|_\infty\leq 2.
\end{aligned}
$$

Passing to the supremum over all $w(y)$ gives part (e).

Part (f) is the content of Lemma 4.1 in [181]. We reproduce the proof given there.

Consider the σ-algebra $\mathscr{G}:=\{\mathfrak{X}\times E : E\subset\mathfrak{Y}\text{ is measurable}\}$, then $\mathscr{G}$ represents the information on the $\mathfrak{Y}$–coordinate of $(x,y)\in\mathfrak{X}\times\mathfrak{Y}$.

Let $\widetilde{\pi}_y$ be a measurable family of conditional probabilities given $\mathscr{G}$, i.e. $\widetilde{\pi}_y$ is a probability measure on $\mathfrak{X}\times\{y\}$, $y\mapsto\int f\mathrm{d}\widetilde{\pi}_y$ is Borel for every Borel function $f:\mathfrak{X}\times\mathfrak{Y}\to[0,1]$, $\lambda=\int_{\mathfrak{X}\times\mathfrak{Y}}\widetilde{\pi}_y\mathrm{d}\lambda$, and for every λ–absolutely integrable $f(x,y)$, $\mathbb{E}_\lambda(f(x,y)|\mathscr{G})(y)=\int_{\mathfrak{X}} f\mathrm{d}\widetilde{\pi}_y$ λ-a.e.

We may identify $\widetilde{\pi}_y$ with a probability measure $\widehat{\pi}(y,\mathrm{d}x)$ on $\mathfrak{X}$ defined by $\widehat{\pi}(y,E)=\widetilde{\pi}_y(E\times\{y\})$ $(E\subset\mathfrak{X}$ Borel$)$.

It is useful to think of $\widehat{\pi}(y, \mathrm{d}x)$ as the transition kernel "which goes the opposite way" to $\pi(x, \mathrm{d}y)$. Indeed, if $\pi(x, \mathrm{d}y)$ is the transition probability of a Markov chain $\{X_n\}$ from n to $n+1$, and λ is the joint distribution of (X_n, X_{n+1}), then $\widehat{\pi}(y, \mathrm{d}x)$ is the transition probability from $n+1$ to n, i.e. $\widehat{\pi}(y, E) = \mathbb{P}(X_n \in E | X_{n+1} = y)$.

The operators $\pi : L^2(\mu_{\mathfrak{Y}}) \to L^2(\mu_{\mathfrak{X}})$ and $\widehat{\pi} : L^2(\mu_{\mathfrak{X}}) \to L^2(\mu_{\mathfrak{Y}})$ are dual to one another, because $\int_{\mathfrak{X}} f(x)\pi(g)(x)\mathrm{d}\mu_{\mathfrak{X}}(x)$ and $\int_{\mathfrak{Y}} \widehat{\pi}(f)(y)g(y)\mathrm{d}\mu_{\mathfrak{Y}}(y)$ are both equal to $\iint f(x)g(y)\lambda(\mathrm{d}x, \mathrm{d}y)$.

Claim: *$Q := \pi \circ \widehat{\pi} : L^2(\mu_{\mathfrak{X}}) \to L^2(\mu_{\mathfrak{X}})$ is self-adjoint, Q preserves the linear subspace $L_0^2(\mu_{\mathfrak{X}}) := \{f \in L^2(\mu_{\mathfrak{X}}) : \int f \mathrm{d}\mu_{\mathfrak{X}} = 0\}$, and the spectral radius of $Q : L_0^2 \to L_0^2$ is at most $\delta(Q)$.*

Proof of the Claim: Q is self adjoint, because $Q^* = (\pi\widehat{\pi})^* = \widehat{\pi}^*\pi^* = \pi\widehat{\pi} = Q$.

It is useful to notice that Q is given by $(Qf)(x) = \int_{\mathfrak{X}} f(x')Q(x, \mathrm{d}x')$ where $Q(x, E)$ is the probability measure on $\mathfrak{X}$ given by $Q(x, E) = \int \widehat{\pi}(y, E)\pi(x, \mathrm{d}y)$. $Q(x, \mathrm{d}x')$ is a transition probability from $\mathfrak{X}$ to $\mathfrak{X}$. Notice that $Q(\mu_{\mathfrak{X}}) = \mu_{\mathfrak{X}}$:

$$
\begin{aligned}
(Q\mu_{\mathfrak{X}})(E) &= \int_{\mathfrak{X}} Q(x, E)\mu_{\mathfrak{X}}(\mathrm{d}x) = \int_{\mathfrak{X}}\int_{\mathfrak{Y}} \mu_{\mathfrak{X}}(\mathrm{d}x)\pi(x, \mathrm{d}y)\widetilde{\pi}_y(E \times \{y\}) \\
&= \int_{\mathfrak{X}\times\mathfrak{Y}} \widetilde{\pi}_y(E \times \{y\})\lambda(\mathrm{d}x, \mathrm{d}y) = \int_{\mathfrak{X}\times\mathfrak{Y}} \widetilde{\pi}_y(E \times \mathfrak{Y})\mathrm{d}\lambda = \lambda(E \times \mathfrak{Y}) = \mu_{\mathfrak{X}}(E).
\end{aligned}
$$

Thus, for all $f \in L^2(\mu_{\mathfrak{X}})$, $\int Qf \mathrm{d}\mu_{\mathfrak{X}} = \int f \mathrm{d}(Q\mu_{\mathfrak{X}}) = \int f \mathrm{d}\mu_{\mathfrak{X}}$. It follows that $Q : L^2(\mu_{\mathfrak{X}}) \to L^2(\mu_{\mathfrak{X}})$ preserves the linear space L_0^2.

For every $\varphi \in L_0^2 \cap L^\infty$, $\|\varphi\|_\infty \leq \mathrm{Osc}(\varphi)$. Since Q preserves $L_0^2 \cap L^\infty$, for every f in this space, we have by parts (c) and (d) that

$$
\|Q^n f\|_2 \leq \|Q^n f\|_\infty \leq \mathrm{Osc}(Q^n f) \leq \delta(Q)^n \,\mathrm{Osc}(f). \tag{2.9}
$$

This implies that the spectral radius of $Q : L_0^2 \to L_0^2$ is less than or equal to $\delta(Q)$. Otherwise there is an L_0^2-function, part of whose spectral decomposition corresponds to the part of the spectrum outside $\{\lambda \in \mathbb{R} : |\lambda| \leq \delta(Q) + \epsilon\}$ (self-adjoint operators have real spectrum). Any sufficiently close $L_0^2 \cap L^\infty$–function would have components with similar properties; but the existence of such components is inconsistent with (2.9). The proof of the claim is complete.

$Q : L_0^2 \to L_0^2$ is a self-adjoint, with spectral radius at most $\delta(Q)$, so for all $f \in L_0^2(\mu_{\mathfrak{X}})$, $\langle Q(f), f\rangle_{L_0^2} \leq \delta(Q)\|f\|_{L_0^2}^2$, and

$$
\|\widehat{\pi}(f)\|^2_{L_0^2(\mu_{\mathfrak{Y}})} = \langle \widehat{\pi}(f), \widehat{\pi}(f)\rangle_{L_0^2(\mu_{\mathfrak{Y}})} = \langle Q(f), f\rangle_{L_0^2(\mu_{\mathfrak{X}})} \leq \delta(Q)\|f\|^2_{L_0^2(\mu_{\mathfrak{X}})}.
$$

We can now prove (f). Fix $f \in L_0^2(\mu_{\mathfrak{X}})$, $g \in L_0^2(\mu_{\mathfrak{Y}})$, then

$$
\begin{aligned}
&\left|\int_{\mathfrak{X}\times\mathfrak{Y}} f(x)g(y)\lambda(\mathrm{d}x, \mathrm{d}y)\right| = \left|\int_{\mathfrak{Y}} \mu_{\mathfrak{Y}}(\mathrm{d}y)\int_{\mathfrak{X}} \widehat{\pi}(y, \mathrm{d}x)f(x)g(y)\right| = \langle \widehat{\pi}(f), g\rangle_{L^2(\mu_{\mathfrak{Y}})} \\
&\leq \|\widehat{\pi}(f)\|_2\|g\|_2 \leq \sqrt{\delta(Q)}\|f\|_2\|g\|_2, \text{ as stated in (f).}
\end{aligned}
$$

If $g \in L_0^2(\mu_{\mathfrak{Y}})$, then $\displaystyle\int \pi(g)\mathrm{d}\mu_{\mathfrak{X}} = \iiint g(y)\mu_{\mathfrak{X}}(\mathrm{d}x)\pi(x, \mathrm{d}y) = \iiint g(y)\lambda(\mathrm{d}x, \mathrm{d}y) = \int g(y)\mu_{\mathfrak{Y}}(\mathrm{d}y) = 0$. Substituting $f := \pi(g)$ in (f) and noting that $\displaystyle\int \pi(g)(x)g(y)\lambda(dx, dy) = \int (\pi(g))^2(x)d\mu_{\mathfrak{X}}(dx) = \|\pi(g)\|_2^2$, we obtain (g). □

We now return to the setup of Markov arrays and consider the following **two-step transition probabilities**

$$
\pi^{(N)}_{n,n+2}(x, E) := \int \pi^{(N)}_{n+1,n+2}(y, E)\, \pi^{(N)}_{n,n+1}(x, \mathrm{d}y),
$$

defined for $1 \le n \le k_N - 1$, $x \in \mathfrak{S}_n^{(N)}$, and $E \in \mathscr{B}(\mathfrak{S}_{n+2}^{(N)})$. The uniform ellipticity condition gives the following uniform bound for $\delta(\pi_{n,n+2}^{(N)})$:

Lemma 2.12 *Let* X *be a uniformly elliptic Markov array with ellipticity coefficient* ϵ_0. *Then* $\sup\limits_N \sup\limits_{1\le n<k_N} \delta(\pi_{n,n+2}^{(N)}) \le 1-\epsilon_0$. *Similarly for Markov chains.*

Proof We fix N and drop the superscripts $^{(N)}$.

Uniform ellipticity implies that $\pi_{n,n+2}(x, \mathrm{d}y) \ll \mu_{n+2}(\mathrm{d}y)$ and that the Radon-Nikodym density is bounded from below by ϵ_0. This allows us to find a transition kernel $\widehat{\pi}_{n,n+1}(x, \mathrm{d}y)$ such that

$$\pi_{n,n+2}(x, \mathrm{d}y) = \epsilon_0 \mu_{n+2}(\mathrm{d}y) + (1-\epsilon_0)\widehat{\pi}_{n,n+2}(x, \mathrm{d}y). \tag{2.10}$$

Note that the first term does not depend on x.

Let $u : \mathfrak{S}_{n+2} \to \mathbb{R}$ be a measurable function with $\mathrm{Osc}(u) \le 1$, then we can write $u(\cdot) = c + w(\cdot)$, where c is a constant and $\|w\|_\infty \le \frac{1}{2}$. A direct calculation shows that

$$\left|\int_{\mathfrak{S}_n} u(z)\pi_{n,n+2}(x_1, \mathrm{d}z) - \int_{\mathfrak{S}_n} u(z)\pi_{n,n+2}(x_2, \mathrm{d}z)\right| = \left|\int_{\mathfrak{S}_n} w(z)\pi_{n,n+2}(x_1, \mathrm{d}z) - \int_{\mathfrak{S}_n} w(z)\pi_{n,n+2}(x_2, \mathrm{d}z)\right| =$$

$$(1-\epsilon_0)\left|\int_{\mathfrak{S}_n} w(z)\widehat{\pi}_{n,n+2}(x_1, \mathrm{d}z) - \int_{\mathfrak{S}_n} w(z)\widehat{\pi}_{n,n+2}(x_2, \mathrm{d}z)\right| \le (1-\epsilon_0)\|w\|_\infty \left[\pi_{n,n+2}(x_1, \mathfrak{S}_{n+2}) + \pi_{n,n+2}(x_2, \mathfrak{S}_{n+2})\right] \le 1-\epsilon_0. \square$$

Proposition 2.13 *Let* X *be a uniformly elliptic Markov array with row lengths* $k_N + 1$. *Then there exist* $\theta \in (0,1)$ *and* $C_{mix} > 0$, *which only depend on the ellipticity constant* ϵ_0, *as follows. Suppose* $h_n^{(N)}(x, y)$ *are measurable functions on* $\mathfrak{S}_n^{(N)} \times \mathfrak{S}_{n+1}^{(N)}$, *and let* $h_n^{(N)} := h_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$, *then:*

(1) If $h_n^{(N)}$ *is bounded and* $\mathbb{E}(h_n^{(N)}) = 0$, *then for all* $1 \le m < n \le k_N$,

$$\|\mathbb{E}(h_n^{(N)}|X_m^{(N)})\|_\infty \le C_{mix}\theta^{n-m}\|h_n^{(N)}\|_\infty. \tag{2.11}$$

(2) If $\mathrm{Var}(h_n^{(N)}), \mathrm{Var}(h_m^{(N)}) < \infty$ *and* $\mathbb{E}(h_n^{(N)}), \mathbb{E}(h_m^{(N)}) = 0$, *then for all* $1 \le m < n \le k_N$,

$$\|\mathbb{E}(h_n^{(N)}|X_m^{(N)})\|_2 \le C_{mix}\theta^{n-m}\|h_n^{(N)}\|_2; \tag{2.12}$$

$$|\mathbb{E}(h_m^{(N)}h_n^{(N)})| \le C_{mix}\theta^{n-m}\|h_m^{(N)}\|_2\|h_n^{(N)}\|_2. \tag{2.13}$$

The analogous statements hold for Markov chains.

Proof We fix N and drop the superscripts $^{(N)}$. Let $w_{n,k}(X_k) := \mathbb{E}(h_n|X_k)$ $(k < n)$. Then $\pi_{k-1,k}(w_{n,k}) = w_{n,k-1}$, because

$$\pi_{k-1,k}(w_{n,k})(X_{k-1}) = \int w_{n,k}(y)\pi_{k-1,k}(X_{k-1}, \mathrm{d}y) = \mathbb{E}[w_{n,k}(X_k)|X_{k-1}] = \mathbb{E}[\mathbb{E}(h_n|X_k)|X_{k-1}]$$

$$\stackrel{!}{=} \mathbb{E}[\mathbb{E}(h_n|X_k, X_{k-1}, \ldots, X_1)|X_{k-1}], = \mathbb{E}(h_n|X_{k-1}) = w_{n,k-1}(X_{k-1}) \quad \text{(! follows by the Markov property)}$$

Hence $w_{n,m}(X_m) = (\pi_{m,m+1} \circ \cdots \circ \pi_{n-1,n})(w_{n,n})(X_m)$.

By the previous lemmas, $\mathrm{Osc}[w_{n,m}] \le (1-\epsilon_0)^{\lfloor \frac{n-m}{2} \rfloor}\mathrm{Osc}[w_{n,n}]$. Notice that for every bounded measurable function v, $\|v\|_\infty \le |\mathbb{E}(v)| + \mathrm{Osc}(v)$. Since by assumption $\mathbb{E}(w_{n,m}(X_m)) = \mathbb{E}(h_n) = 0$,

$$\|\mathbb{E}(h_n|X_m)\|_\infty = \|w_{n,m}(X_m)\|_\infty \le (1-\epsilon_0)^{\lfloor \frac{n-m}{2} \rfloor}\mathrm{Osc}[w_{n,n}].$$

Since $\mathrm{Osc}[w_{n,n}] \le 2\|w_{n,n}\|_\infty \le 2\|h_n\|_\infty$, (2.11) follows.

(2.12) can be proved in the same way, using Lemma 2.11(g).

By the Markov property, $\mathbb{E}(h_m h_n) = \mathbb{E}[\mathbb{E}(h_m h_n|X_m, X_{m+1})] = \mathbb{E}[h_m\mathbb{E}(h_n|X_{m+1})]$. So $|\mathbb{E}(h_m h_n)| \le \|h_m\|_2\|\mathbb{E}(h_n|X_{m+1})\|_2$, and (2.12)$\Rightarrow$(2.13), perhaps with a bigger C_{mix}. $\square$

Lemma 2.14 *Suppose* f *is a uniformly bounded additive functional of a uniformly elliptic Markov array* X. *There is a constant* C *which only depends on the ellipticity constant of* X *such that* $\forall\{\ell_N\}$, *if* $h^{(N)}_{\ell_N}$ *are uniformly bounded measurable functions on* $\mathfrak{S}^{(N)}_{\ell_N} \times \mathfrak{S}^{(N)}_{\ell_{N+1}}$, *and* $\operatorname{ess\,sup}|h^{(N)}_{\ell_N}| \le L$, *then*
$\operatorname{Cov}\left(S_N, h^{(N)}_{\ell_N}(X^{(N)}_{\ell_N}, X^{(N)}_{\ell_{N+1}})\right) \le CL \operatorname{ess\,sup}|\mathsf{f}|.$

Proof Write $\operatorname{Cov}(S_N, h^{(N)}_{\ell_N}) = \sum_{n=1}^{k_N} \operatorname{Cov}(f^{(N)}_n, h^{(N)}_{\ell_N})$ and use (2.13). □

2.2.3 Bridge Probabilities

We would like to define "the distribution of $X^{(N)}_n$ given $X^{(N)}_{n-1} = x$ and $X^{(N)}_{n+1} = z$" for *every* (not just almost every) x, z.

Suppose X is uniformly elliptic, and write $\pi^{(N)}_{n,n+1}(x, dy) = p^{(N)}_n(x, y)\mu^{(N)}_{n+1}(dy)$, with $p^{(N)}_n$ and $\mu^{(N)}_n$ as in the definition of uniform ellipticity.

Then $Z^{(N)}_n(x, z) := \int_{\mathfrak{S}^{(N)}_{n+1}} p^{(N)}_n(x, y)p^{(N)}_{n+1}(y, z)\mu^{(N)}_{n+1}(dy) > \epsilon_0 > 0$, and we can *define* a measure on $\mathfrak{S}^{(N)}_{n+1}$ by

$$\mathbb{P}\left(E \;\middle|\; \begin{matrix} X^{(N)}_n = x \\ X^{(N)}_{n+2} = z \end{matrix}\right) := \frac{1}{Z^{(N)}_n(x, z)} \int_E p^{(N)}_n(x, y)p^{(N)}_{n+1}(y, z)\mu^{(N)}_{n+1}(dy). \tag{2.14}$$

Lemma 2.15 *Let* $\psi_E(x, z)$ *denote the right-hand-side of* (2.14), *then*

$$\psi_E(X^{(N)}_n, X^{(N)}_{n+2}) = \mathbb{P}(X^{(N)}_{n+1} \in E | X^{(N)}_n, X^{(N)}_{n+2}) \;\; \mathbb{P}\text{-almost everywhere.}$$

Proof We fix N and drop the superscripts ${}^{(N)}$.

Clearly $\psi_E(X_n, X_{n+2})$ is measurable with respect to the σ-algebra generated by X_n, X_{n+2}. To prove the lemma, we need to check that for every bounded measurable function φ on $\mathfrak{S}_n \times \mathfrak{S}_{n+2}$, $\mathbb{E}[(\varphi\psi_E)(X_n, X_{n+2})] = \mathbb{E}[\varphi(X_n, X_{n+2})1_E(X_{n+1})]$.

Let P_n denote the measure $P_n(E') = \mathbb{P}(X_n \in E')$, then

$$\mathbb{E}[(\varphi\psi_E)(X_n, X_{n+2})] = \iiint \varphi(x, z)\frac{\int p_n(x, y')p_n(y', z)1_E(y')\mu_{n+1}(dy')}{Z_n(x, z)} p_n(x, y)p_{n+1}(y, z)P_n(dx)\mu_{n+1}(dy)\mu_{n+2}(dz).$$

After integrating out y, we are left with the triple integral $\iiint \varphi(x, z)1_E(y')p(x, y')p_n(y', z)P_n(dx)\mu_{n+1}(dy')\mu_{n+2}(dz)$, which equals $\mathbb{E}[\varphi(X_n, X_{n+2})1_E(X_{n+1})]$. □

The lemma does not "prove" (2.14): Conditional probabilities are only defined almost everywhere; but the point of (2.14) is that it makes sense for *all* x, z.

Motivated by Lemma 2.15, we call (2.14) the **bridge distribution** of $X^{(N)}_{n+1}$ given that $X^{(N)}_n = x$ and $X^{(N)}_{n+2} = z$.

The "bridge" $x \to E \to z$ in (2.14) has length 2. It is easy to extend the definition to bridges of length $m \ge 3$. Suppose $1 \le n \le n + m \le k_N + 1$, and let

$$\mathfrak{S}^{(N)}_{n,m} := \mathfrak{S}^{(N)}_n \times \cdots \times \mathfrak{S}^{(N)}_{n+m}, \tag{2.15}$$

$$\mu^{(N)}_{n,m} := \mu^{(N)}_n \times \cdots \times \mu^{(N)}_{n+m}, \tag{2.16}$$

$$p_{n,m}^{(N)}(x_n, \dots, x_{n+m}) := \prod_{i=0}^{m-1} p_{n+i}^{(N)}(x_{n+i}, x_{n+i+1}), \tag{2.17}$$

$$p_n^{(N)}(x_n \to x_{n+m}) := \int_{\mathfrak{S}_{n+1,m-2}^{(N)}} p_{n,m}^{(N)}(x_n, \xi_{n+1}, \dots, \xi_{n+m-1}, x_{n+m}) \mathrm{d}\mu_{n+1,m-2}^{(N)}, \tag{2.18}$$

where the integration is over $(\xi_{n+1}, \dots, \xi_{n+m-1})$.

Note that $p_n^{(N)}(x_n \to x_{n+m})$ is the density of $\mathbb{P}[X_{n+m}^{(N)} = \mathrm{d}x_{n+m} | X_n^{(N)} = x_n]$ with respect to $\mu_{n+m}^{(N)}$. By uniform ellipticity, $p_n^{(N)}(x_n \to x_{n+m}) \le \epsilon_0^{-m}$, and

$$p_n^{(N)}(x_n \to x_{n+m}) = \int_{\mathfrak{S}_{n+1,m-2}^{(N)}} \mathrm{d}\mu_{n+1,m-3}^{(N)} \Big[p_{n,m-2}^{(N)}(x_n, \xi_{n+1}, \dots, \xi_{n+m-2}) \times$$

$$\times \int_{\mathfrak{S}_{n+m-1}^{(N)}} p(\xi_{n+m-2}, \xi_{n+m-1}) p(\xi_{n+m-1}, x_{n+m}) \mu_{n+m-1}^{(N)}(\mathrm{d}\xi_{n+m-1}) \Big] \ge \epsilon_0. \tag{2.19}$$

Since $p_n^{(N)}(x_n \to x_{n+m}) \neq 0$, we can define the bridge distribution of $X_{n+1}^{(N)}$ given that $X_n^{(N)} = x_n$ and $X_{n+m}^{(N)} = z_{n+m}$ to be the measure on $\mathfrak{S}_{n+1,m-2}^{(N)}$, given by

$$\mathbb{P}\left(E \,\middle|\, \begin{matrix} X_n^{(N)} = x_n \\ X_{n+m}^{(N)} = z_{n+m} \end{matrix} \right) := \frac{1}{p_n^{(N)}(x_n \to z_{n+m})} \int_E p_{n,m}^{(N)}(x_n, \xi_{n+1}, \dots, \xi_{n+m-1}, z_{n+m}) \mathrm{d}\mu_{n+1,m-2}^{(N)}, \tag{2.20}$$

where the integration is over $(\xi_{n+1}, \dots \xi_{n+m-1})$.

Again, this agrees a.s. with $\mathbb{E}(1_E | X_n^{(N)}, X_{n+m}^{(N)})(x_n, z_{n+m})$. But unlike the conditional expectation, (2.20) makes sense globally, and pointwise, and is not just an L^1-equivalence class.

2.3 Structure Constants

Throughout this section we assume that f is an additive functional on a uniformly elliptic Markov array X with row lengths $k_N + 1$, state spaces $\mathfrak{S}_n^{(N)}$, and transition probabilities $\pi_{n,n+1}^{(N)}(x, \mathrm{d}y) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(\mathrm{d}y)$, as in the ellipticity condition. By Corollary 2.9, we may assume that $\mu_n^{(N)}(E) = P_n(E) := \mathbb{P}(X_n^{(N)} \in E)$ for $n \ge 3$.

2.3.1 Hexagons

A **Level N hexagon at position $3 \le n \le k_N$** is 6-tuple $\left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1} \right)$ where $x_i, y_i \in \mathfrak{S}_i^{(N)}$. A hexagon is **admissible** if

$$p_{n-2}^{(N)}(x_{n-2}, x_{n-1}) p_{n-1}^{(N)}(x_{n-1}, x_n) p_n^{(N)}(x_n, y_{n+1}) \neq 0, \quad p_{n-2}^{(N)}(x_{n-2}, y_{n-1}) p_{n-1}^{(N)}(y_{n-1}, y_n) p_n^{(N)}(y_n, y_{n+1}) \neq 0.$$

Admissible hexagons exist because of uniform ellipticity. The **hexagon spaces** are the spaces of level N admissible hexagons at position n. We denote them by $\mathrm{Hex}(N, n)$ or, in the case of Markov chains, by $\mathrm{Hex}(n)$.

The **hexagon measure** $m_{\mathrm{Hex}} = m_{\mathrm{Hex}}^{N,n}$ is the probability measure on $\mathrm{Hex}(N, n)$ arising from the following sampling procedure for $\left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1} \right)$:

- Let $\{X_n^{(N)}\}$ and $\{Y_n^{(N)}\}$ be two independent copies of X;
- (x_{n-2}, x_{n-1}) is sampled from the distribution of $(X_{n-2}^{(N)}, X_{n-1}^{(N)})$;

- (y_n, y_{n+1}) is sampled from the distribution of $(Y_n^{(N)}, Y_{n+1}^{(N)})$ (so it is independent of (x_n, x_{n+1}));
- x_n and y_{n-1} are conditionally independent given the previous choices, and are sampled using the bridge distributions

$$\mathbb{P}(x_n \in E|x_{n-1}, y_{n+1}) = \mathbb{P}\left(X_n^{(N)} \in E|X_{n-1}^{(N)} = x_{n-1}, X_{n+1}^{(N)} = y_{n+1}\right)$$
$$\mathbb{P}(y_{n-1} \in E|x_{n-2}, y_n) = \mathbb{P}\left(Y_{n-1}^{(N)} \in E|Y_{n-2}^{(N)} = x_{n-2}Y_n^{(N)} = y_n\right).$$

To write this explicitly in coordinates, we let $P = \left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1}\right)$, and set:

$$\begin{aligned}
&m_{\mathrm{Prod}}(\mathrm{d}P) = m_{\mathrm{Prod}}^{N,n}(\mathrm{d}P) := P_{n-2}(\mathrm{d}x_{n-2}) \prod_{i=n-1}^{n} \mu_i^{(N)}(\mathrm{d}x_i) \prod_{i=n-1}^{n+1} \mu_i^{(N)}(\mathrm{d}y_i) \\
&\varphi^+(P) = \varphi^+_{N,n}(P) := p_{n-2}^{(N)}(x_{n-2}, x_{n-1}) p_{n-1}^{(N)}(x_{n-1}, x_n) p_n^{(N)}(x_n, y_{n+1}) \\
&\varphi^-(P) = \varphi^-_{N,n}(P) := p_{n-2}^{(N)}(x_{n-2}, y_{n-1}) p_{n-1}^{(N)}(y_{n-1}, y_n) p_n^{(N)}(y_n, y_{n+1}) \\
&Z^+(P) = Z^+_{N,n}(P) := \int p_{n-1}^{(N)}(x_{n-1}, \xi) p_n^{(N)}(\xi, y_{n+1}) \mu_n^{(N)}(\mathrm{d}\xi) \\
&Z^-(P) = Z^-_{N,n}(P) := \int p_{n-2}^{(N)}(x_{n-2}, \xi) p_{n-1}^{(N)}(\xi, y_n) \mu_{n-1}^{(N)}(\mathrm{d}\xi).
\end{aligned} \tag{2.21}$$

We will drop the indices N, n when they are clear from the context. It is not difficult to see that the following identity holds:

$$m_{\mathrm{Hex}}(\mathrm{d}P) = \frac{\varphi^+(P)\varphi^-(P)}{Z^+(P)Z^-(P)} m_{\mathrm{Prod}}(\mathrm{d}P). \tag{2.22}$$

The hexagon measure is asymmetric in the following sense: $m_{\mathrm{Hex}} \circ \iota \neq m_{\mathrm{Hex}}$, where ι is the involution $\iota : \left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1}\right) \mapsto \left(x_{n-2}; \begin{matrix} y_{n-1} \\ x_{n-1} \end{matrix}; \begin{matrix} y_n \\ x_n \end{matrix}; y_{n+1}\right)$. There is another natural measure on $\mathrm{Hex}(N, n)$ which is invariant under ι.

This measure, which we will denote by m'_{Hex}, is the result of the following sampling procedure for $P := \left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1}\right)$:

- Let $\{X_n^{(N)}\}$ and $\{Y_n^{(N)}\}$ be two independent copies of X;
- x_{n-2} is sampled from the distribution of $X_{n-2}^{(N)}$;
- y_{n+1} is sampled from the distribution of $Y_{n+1}^{(N)}$ (so it is independent of x_n);
- conditioned on x_{n-2}, x_{n+1}, the pairs (x_{n-1}, x_n) and (y_{n-1}, y_n) are independent, and identically distributed like the bridge distribution of $(X_{n-1}^{(N)}, X_n^{(N)})$ given $X_{n-2}^{(N)} = x_{n-2}, X_{n+1}^{(N)} = x_{n+1}$.

Equivalently, if $Z(P) = Z_{N,n}(P) = p_n^{(N)}(x_{n-2} \to y_{n+1})$ (see (2.18)), then

$$m'_{\mathrm{Hex}}(\mathrm{d}P) = \frac{\varphi^+(P)\varphi^-(P)}{Z(P)^2} m_{\mathrm{Prod}}(\mathrm{d}P). \tag{2.23}$$

Recall that $Z^\pm(P) \in [\epsilon_0, \epsilon_0^{-1}]$ and $Z(P) \in [\epsilon_0, \epsilon_0^{-3}]$, see (2.18) and (2.19). Therefore:

$$\epsilon_0^4 \le \frac{\mathrm{d}m'_{\mathrm{Hex}}}{\mathrm{d}m_{\mathrm{Hex}}} \le \epsilon_0^{-8}. \tag{2.24}$$

Because of (2.24), we could have chosen either m_{Hex} or m'_{Hex} as the basis for our work. The reason we prefer the asymmetric m_{Hex} to the symmetric m'_{Hex} will become apparent in §2.3.3. There we will see that $\{m_{\mathrm{Hex}}^{N,n}\}$ can be coupled in a natural way.

2.3.2 Balance and Structure Constants

The **balance** of a hexagon $P := \left(x_{n-2}; \begin{matrix} x_{n-1} \\ y_{n-1} \end{matrix}; \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1} \right) \in \text{Hex}\,(N, n)$ is

$$\Gamma(P) := f_{n-2}^{(N)}(x_{n-2}, x_{n-1}) + f_{n-1}^{(N)}(x_{n-1}, x_n) + f_n^{(N)}(x_n, y_{n+1}) - f_{n-2}^{(N)}(x_{n-2}, y_{n-1}) - f_{n-1}^{(N)}(y_{n-1}, y_n) - f_n^{(N)}(y_n, y_{n+1}). \tag{2.25}$$

The **structure constants** of $\mathsf{f} = \{f_n^{(N)}\}$ are

$$\begin{aligned} u_n^{(N)} &:= u_n^{(N)}(\mathsf{f}) := \mathbb{E}_{m_{\text{Hex}}^{N,n}}(\Gamma(P)^2)^{1/2}, \quad d_n^{(N)}(\xi) := d_n^{(N)}(\xi, \mathsf{f}) := \mathbb{E}_{m_{\text{Hex}}^{N,n}}(|e^{i\xi\Gamma(P)} - 1|^2)^{1/2}, \\ U_N &:= U_N(\mathsf{f}) := \sum_{n=3}^{k_N} (u_n^{(N)})^2, \; D_N(\xi) := \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2, \end{aligned} \tag{2.26}$$

where the expectation is over random $P \in \text{Hex}\,(N, n)$. If X is a Markov chain, we write $u_n = u_n^{(N)}$ and $d_n(\xi) = d_n^{(N)}(\xi)$.

The significance of these quantities was mentioned briefly in §1.3, and will be explained in later chapters. At this point we can only say that the behavior of U_N determines if $\text{Var}(S_N) \to \infty$, and the behavior of $D_N(\xi)$ determines "how close" f is to an additive functional whose values all belong to a coset of the lattice $(2\pi/\xi)\mathbb{Z}$.

Lemma 2.16 *Suppose* f, g *are two additive functionals of on a uniformly elliptic Markov array* X*, and let* ξ, η *be real numbers, then*

(a) $d_n^{(N)}(\xi + \eta, \mathsf{f})^2 \le 8(d_n^{(N)}(\xi, \mathsf{f})^2 + d_n^{(N)}(\eta, \mathsf{f})^2)$;
(b) $d_n^{(N)}(\xi, \mathsf{f} + \mathsf{g})^2 \le 8(d_n^{(N)}(\xi, \mathsf{f})^2 + d_n^{(N)}(\xi, \mathsf{g})^2)$;
(c) $d_n^{(N)}(\xi, \mathsf{f}) \le |\xi| u_n^{(N)}(\mathsf{f})$;
(d) $u_n^{(N)}(\mathsf{f} + \mathsf{g})^2 \le 2[u_n^{(N)}(\mathsf{f})^2 + u_n^{(N)}(\mathsf{g})^2]$.

Proof For any $z, w \in \mathbb{C}$ such that $|z|, |w| \le 2$, we have [1] $|zw + z + w|^2 \le 8(|z|^2 + |w|^2)$. So if $P \in \text{Hex}\,(N, n)$ and $\xi_P := \xi\Gamma(P)$, $\eta_P := \eta\Gamma(P)$, then

$$|e^{i(\xi_P + \eta_P)} - 1|^2 = |(e^{i\xi_P} - 1)(e^{i\eta_P} - 1) + (e^{i\xi_P} - 1) + (e^{i\eta_P} - 1)|^2 \le 8(|e^{i\xi_P} - 1|^2 + |e^{i\eta_P} - 1|^2). \tag{2.27}$$

Part (a) follows by integrating over all $P \in \text{Hex}(n, N)$. Part (b) has a similar proof which we omit. Part (c) is follows from the inequality $|e^{i\theta} - 1|^2 = 4\sin^2\frac{\theta}{2} \le |\theta|^2$. Part (d) follows from the general inequality $(a + b)^2 \le 2(a^2 + b^2)$. □

Lemma 2.17 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*. Fix* $x \in \mathfrak{S}_1$*, and let* $d_n(\xi, x)$ *denote the structure constants of* f *on* X *conditioned on* $X_1 = x$*. Then there exists* $0 < \theta < 1$ *such that* $|d_n^2(\xi, x) - d_n^2(\xi)| = O(\theta^n)$ *for all* $\xi \in \mathbb{R}$.

The proof is given in §2.3.3.

Example 2.18 (Vanishing Structure Constants) Suppose $f_n(x, y) = a_{n+1}(y) - a_n(x) + c_n$ for all n, then the balance of each hexagon is zero and $u_n, d_n(\xi)$ are all zero. For more on this, see §3.2.1.

Suppose $f_n(x, y) = a_{n+1}(y) - a_n(x) + c_n \mod \frac{2\pi}{\xi}\mathbb{Z}$ for all n. Then $e^{i\xi\Gamma(P)} = 1$ for all hexagons P, and $d_n(\xi)$ are all zero. For more on this, see §4.3.1.

Example 2.19 (Sums of Independent Random Variables) Let $S_N = X_1 + \cdots + X_N$. where X_i are independent real valued random variables with non-zero variance. Let us see what u_n measures in this case.

Proposition 2.20 $u_n^2 = 2(\text{Var}(X_{n-1}) + \text{Var}(X_n))$ *and* $U_N := \sum_{n=3}^N u_n^2 \asymp \text{Var}(S_N)$ *(i.e* $\exists N_0$ *such that the ratio of the two sides is uniformly bounded for* $N \ge N_0$*).*

---
[1] $(zw+z+w)^2 = z^2w^2+z^2+w^2+2(z^2w+zw^2+zw)$, and $|z^2w^2| \le 4|zw| \le 2|z|^2+2|w|^2$, $|z^2w| \le 2|z|^2$, $2|zw| \le |z|^2+|w|^2$, $|zw^2| \le 2|w|^2$.

Proof Let $\{Y_n\}$ be an independent copy of $\{X_n\}$, and let $X_i^* := X_i - Y_i$ (the symmetrization of X_i). A simple calculation shows that the balance of a position n hexagon is equal in distribution to $X_{n-1}^* + X_n^*$. Clearly $\mathbb{E}[X_i^*] = 0$ and $\mathbb{E}[(X_i^*)^2] = 2\mathrm{Var}(X_i)$. Consequently,

$$u_n^2(\xi) = \mathbb{E}[(X_{n-1}^*)^2 + (X_n^*)^2] = 2\mathrm{Var}(X_{n-1}) + 2\mathrm{Var}(X_n).$$

Summing over n we obtain $\sum_{n=3}^N u_n^2 \asymp \mathrm{Var}(S_N)$. □

We remark that the proposition also holds for Markov arrays satisfying the *one-step* ellipticity condition (see §2.4 and §3.1).

Now we describe the meaning of $d_n(\xi)$ for sums of independent random variables, and relate it to the distance of X_{n-1} and X_n from cosets of $\frac{2\pi}{\xi}\mathbb{Z}$.

Given a real-valued random variable X, let

$$\mathfrak{D}(X,\xi) := \min_{\theta\in\mathbb{R}} \mathbb{E}\left[\mathrm{dist}^2\left(X, \theta + \frac{2\pi}{\xi}\mathbb{Z}\right)\right]^{1/2}. \tag{2.28}$$

The minimum exists because the quantity we are minimizing is a periodic and continuous function of θ.

Proposition 2.21 *For every $\xi \neq 0$, $d_n(\xi) = 0$ iff there are constants θ_i such that $X_i \in \theta_i + \frac{2\pi}{\xi}\mathbb{Z}$ a.s. ($i = n-1, n$). In addition, there exists $C(\xi) > 1$ such that if $d_n(\xi)$ and $\mathfrak{D}(X_{n-1},\xi)^2 + \mathfrak{D}(X_n,\xi)^2$ are not both zero, then*

$$C(\xi)^{-1} \le \frac{d_n^2(\xi)}{\mathfrak{D}(X_{n-1},\xi)^2 + \mathfrak{D}(X_n,\xi)^2} \le C(\xi).$$

Proof Choose $\theta_i \in [0, \frac{2\pi}{\xi}]$ such that $\mathfrak{D}(X_i,\xi) = \mathbb{E}[\mathrm{dist}^2(X_i, \theta_i + \frac{2\pi}{\xi}\mathbb{Z})]^{1/2}$. There is no loss of generality in assuming that $\theta_i = 0$, because the structure constants of $f_i(x) = x$ and $g_i(x) = x - \theta_i$ are the same. Henceforth we assume that

$$\mathfrak{D}(X_i,\xi) = \mathbb{E}\left[\mathrm{dist}^2\left(X_i, \frac{2\pi}{\xi}\mathbb{Z}\right)\right]^{1/2}. \tag{2.29}$$

As in the proof of the previous proposition, the balance of a position n hexagon is equal in distribution to $X_{n-1}^* + X_n^*$, where $X_i^* := X_i - Y_i$ and $\{Y_i\}$ is an independent copy of $\{X_i\}$. So $d_n^2(\xi) = \mathbb{E}(|e^{i(X_{n-1}^*+X_n^*)} - 1|^2)$.

We need the following elementary facts:

$$|e^{i(x+y)} - 1|^2 = 4\sin^2\tfrac{x+y}{2} = 4(\sin\tfrac{x}{2}\cos\tfrac{y}{2} + \sin\tfrac{y}{2}\cos\tfrac{x}{2})^2 \quad (x,y\in\mathbb{R}) \tag{2.30}$$

$$\tfrac{4}{\pi^2}\mathrm{dist}^2(t,\pi\mathbb{Z}) \le \sin^2 t \le \mathrm{dist}^2(t,\pi\mathbb{Z}) \quad (t\in\mathbb{R}) \tag{2.31}$$

$$\mathbb{P}[X_i^* \in (-\tfrac{\pi}{2\xi}, \tfrac{\pi}{2\xi}) + \tfrac{2\pi}{\xi}\mathbb{Z}] \ge \tfrac{1}{16} \quad (i \ge 1) \tag{2.32}$$

(2.30) is trivial; (2.31) is because of the inequality[2] $2t/\pi \le \sin t \le t$ on $[0, \frac{\pi}{2}]$. To see (2.32), we decompose $\mathbb{R} = \biguplus_{k=0}^3 \left([0, \frac{\pi}{2\xi}) + \frac{k\pi}{2\xi} + \frac{2\pi}{\xi}\mathbb{Z}\right)$, and note that there must be some $k \in \{0,1,2,3\}$ such that $\mathbb{P}[X_i \in [0, \frac{\pi}{2\xi}) + \frac{k\pi}{2\xi} + \frac{2\pi}{\xi}\mathbb{Z}] \ge \frac{1}{4}$. Since Y_i is an independent copy of X_i, $\mathbb{P}[X_i, Y_i \in [0, \frac{\pi}{2\xi}) + \frac{k\pi}{2\xi} + \frac{2\pi}{\xi}\mathbb{Z}] \ge \frac{1}{16}$. This event is a subset of the event $\left[X_i^* \in (-\frac{\pi}{2\xi}, \frac{\pi}{2\xi}) + \frac{2\pi}{\xi}\mathbb{Z}\right]$, and (2.32) follows.

Returning to the identity $d_n^2(\xi) = \mathbb{E}(|e^{i(X_{n-1}^*+X_n^*)} - 1|^2)$, we see that by (2.30)

$$d_n^2(\xi) = \mathbb{E}(|e^{i\xi(X_{n-1}^*+X_n^*)} - 1|^2) = 4\mathbb{E}\left(\sin^2\tfrac{\xi X_{n-1}^*}{2}\cos^2\tfrac{\xi X_n^*}{2} + \sin^2\tfrac{\xi X_n^*}{2}\cos^2\tfrac{\xi X_{n-1}^*}{2} + \tfrac{1}{2}\sin(\xi X_{n-1}^*)\sin(\xi X_n^*)\right).$$

Since X_i^* is symmetric, $\mathbb{E}[\sin(\xi X_i^*)] = 0$, and so

$$d_n^2(\xi) = 4\mathbb{E}\left(\sin^2\tfrac{\xi X_{n-1}^*}{2}\right)\mathbb{E}\left(\cos^2\tfrac{\xi X_n^*}{2}\right) + 4\mathbb{E}\left(\sin^2\tfrac{\xi X_n^*}{2}\right)\mathbb{E}\left(\cos^2\tfrac{\xi X_{n-1}^*}{2}\right). \tag{2.33}$$

[2] $\sin t \ge t/(\pi/2)$ by convexity. The inequality $\sin t \le t$ follows by integrating the inequality $\cos s \le 1$ between 0 and t.

By (2.32), $\mathbb{E}\left(\cos^2 \frac{\xi X_i^*}{2}\right) \geq \cos^2(\frac{\pi}{4})\mathbb{P}[X_i^* \in (-\frac{\pi}{2\xi}, \frac{\pi}{2\xi}) + \frac{2\pi}{\xi}\mathbb{Z}] \geq \frac{1}{32}$. Therefore,

$$d_n^2(\xi) = C_n(\xi)\left[\mathbb{E}\left(\sin^2 \tfrac{\xi X_{n-1}^*}{2}\right) + \mathbb{E}\left(\sin^2 \tfrac{\xi X_n^*}{2}\right)\right] \text{ with } C_n(\xi) \in [\tfrac{1}{8}, 4]. \tag{2.34}$$

It remains to bound $\mathbb{E}(\sin^2(\frac{\xi X_i^*}{2}))$ in terms of $\mathfrak{D}(X_i,\xi)^2$. By (2.30), $\mathbb{E}\left(\sin^2 \frac{\xi X_i^*}{2}\right) = \mathbb{E}\left[\left(\sin \frac{\xi X_i}{2} \cos \frac{\xi Y_i}{2} - \sin \frac{\xi Y_i}{2} \cos \frac{\xi X_i}{2}\right)^2\right]$;

$$\text{So } \mathbb{E}\left(\sin^2 \tfrac{\xi X_i^*}{2}\right) = 2\mathbb{E}\left(\sin^2 \tfrac{\xi X_i}{2}\right)\mathbb{E}\left(\cos^2 \tfrac{\xi X_i}{2}\right) - \tfrac{1}{2}\mathbb{E}\left(\sin(\xi X_i)\right)^2 \leq 2\mathbb{E}\left(\sin^2 \tfrac{\xi X_i}{2}\right)$$
$$\leq 2\mathbb{E}\left(\operatorname{dist}^2\left(\tfrac{\xi X_i}{2}, \pi\mathbb{Z}\right)\right) \equiv \tfrac{\xi^2}{2}\mathbb{E}\left(\operatorname{dist}^2\left(X_i, \tfrac{2\pi}{\xi}\mathbb{Z}\right)\right) = \tfrac{\xi^2}{2}\mathfrak{D}(X_i,\xi)^2, \text{ see (2.29), (2.31).}$$

Next by (2.31) and the definition of $\mathfrak{D}(X_i,\xi)$,

$$\mathbb{E}\left(\sin^2 \tfrac{\xi X_i^*}{2}\right) \geq \tfrac{4}{\pi^2}\mathbb{E}\left(\operatorname{dist}^2(\tfrac{\xi X_i^*}{2}, \pi\mathbb{Z})\right) = \frac{\xi^2}{\pi^2}\mathbb{E}\left(\operatorname{dist}^2(X_i - Y_i, \frac{2\pi}{\xi}\mathbb{Z})\right)$$
$$= \frac{\xi^2}{\pi^2}\mathbb{E}\left[\mathbb{E}\left(\operatorname{dist}^2(X_i, Y_i + \frac{2\pi}{\xi}\mathbb{Z})\Big|Y_i\right)\right] \geq \frac{\xi^2}{\pi^2}\mathbb{E}\left[\mathfrak{D}(X_i,\xi)^2\right] = \frac{\xi^2}{\pi^2}\mathfrak{D}(X_i,\xi)^2.$$

The proposition follows from (2.34). □

2.3.3 The Ladder Process

Lemma 2.22 *Let* X *be a uniformly elliptic Markov array with row lengths* $k_N + 1$. *Then there exists a uniformly elliptic Markov array* L *with the following structure:*

(a) Each row has entries $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)}) \quad (3 \leq n \leq k_N + 1)$.
(b) $\{Z_i^{(N)}\}$ *and* $\{X_i^{(N)}\}$ *are independent copies of* X.
(c) $Y_{n-1}^{(N)} \in \mathfrak{S}_{n-1}^{(N)}$ *are independent given* $\{X_i^{(N)}\}.\{Z_i^{(N)}\}$, *and* $\mathbb{P}\left(Y_{n-1}^{(N)} \in E \,\middle|\, \begin{matrix}\{X_i^{(N)}\} = \{x_i\} \\ \{Z_i^{(N)}\} = \{z_i\}\end{matrix}\right) = \mathbb{P}\left(X_{n-1}^{(N)} \in E \,\middle|\, \begin{matrix}X_{n-2}^{(N)} = z_{n-2} \\ X_n^{(N)} = x_n\end{matrix}\right)$.
(d) The hexagon $\begin{pmatrix} Z_{n-2}^{(N)} & Z_{n-1}^{(N)} & Y_n^{(N)} & X_{n+1}^{(N)} \\ & Y_{n-1}^{(N)} & X_n^{(N)} & \end{pmatrix}$ *is distributed like a random hexagon in* $(\mathrm{Hex}(N,n), m_N^{N,n})$.

Proof We denote the state spaces of X by $\mathfrak{S}_n^{(N)}$, and its transition probabilities by $\pi_n^{(N)}(x, dy) = p_n^{(N)}(x,y)\mu_{n+1}(dy)$, and assume that $p_n^{(N)}$ and $\mu_n^{(N)}$ satisfy the uniform ellipticity condition with ellipticity constant ϵ_0. We also assume without loss of generality that for every $n \geq 3$, $\mu_n^{(N)} = P_n^{(N)}$, where $P_n(E) = \mathbb{P}[X_n^{(N)} \in E]$ (see Corollary 2.9).

Let $\mathbb{P}\left(dy_n \,\middle|\, \begin{smallmatrix}X_{n-1}^{(N)} = z_{n-1} \\ X_{n+1}^{(N)} = x_{n+1}\end{smallmatrix}\right)$ denote the bridge measure on $\mathfrak{S}_n^{(N)}$. Define the Markov array L with

- State spaces $\underline{\mathfrak{S}}_n^{(N)} := \mathfrak{S}_{n-2}^{(N)} \times \mathfrak{S}_{n-1}^{(N)} \times \mathfrak{S}_n^{(N)}$.
- Rows $\underline{L}_n^{(N)} = (z_{n-2}, y_{n-1}, x_n)$ $(3 \leq n \leq k_N + 1,\ N \geq 1)$.
- Initial distribution $\pi^{(N)}(dz_1, dy_2, dx_3) = \int_{\mathfrak{S}_1^{(N)} \times \mathfrak{S}_3^{(N)}} P_1^{(N)}(dz)P_3^{(N)}(dx)\mathbb{P}\left(dy \,\middle|\, \begin{smallmatrix}X_1^{(N)} = z \\ X_3^{(N)} = x\end{smallmatrix}\right)$.
- Transition probabilities $\pi_n^{(N)}((z_{n-2}, y_{n-1}, x_n), E_{n-1} \times E_n \times E_{n+1}) =$

$$= \int_{E_{n-1}\times E_n \times E_{n+1}} p_{n-2}^{(N)}(z_{n-2}, z_{n-1})p_n^{(N)}(x_n, x_{n+1})\mathbb{P}(dy_n \,\Big|\, \begin{smallmatrix}X_{n-1}^{(N)} = z_{n-1} \\ X_{n+1}^{(N)} = x_{n+1}\end{smallmatrix})\mu_{n-1}^{(N)}(dz_{n-1})\mu_{n+1}^{(N)}(dx_{n+1}).$$

The definition of the transition probabilities encodes the following sampling procedure: We evolve $z_{n-2} \to z_{n-1}$ and $x_n \to x_{n+1}$ independently according to $\pi_{n-2}^{(N)}(z_{n-2}, dz)$, $\pi_n^{(N)}(x_n, dx)$, and then we sample y_n using the relevant bridge distribution. Properties (a)–(c) are immediate consequences, and property (d) follows from (a)–(c) and the definition of the hexagon measure $m_N^{N,n}$.

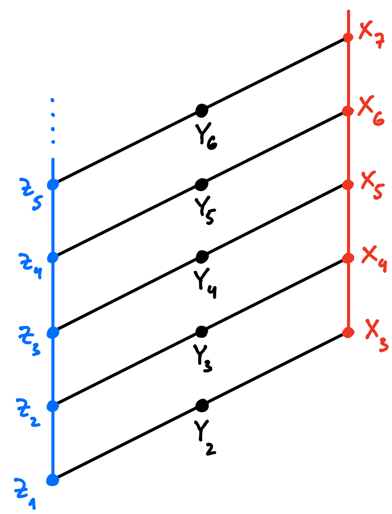

Fig. 2.1 The ladder process (in the case of Markov chains): $\{Z_i\}$, $\{X_i\}$ are independent copies of X, and Y_n are conditionally independent given $\{X_i\}$, $\{Z_i\}$. The marginal distribution of the position n hexagon $\left(Z_{n-2}; {Z_{n-1} \atop Y_{n-1}}; {Y_n \atop X_n}; X_{n+1}\right)$ is m^n_{Hex}.

Note that for $n \geq 5$, $\mu_k^{(N)} = P_k^{(N)}$ for $k = n-2, n$, and the marginal distribution of $\underline{L}_n^{(N)}$ is the measure $m_n^{(N)}(\mathrm{d}\underline{L}_n^{(N)})$ given by:

$$\frac{p_{n-2}^{(N)}(z_{n-2}, y_{n-1})p_{n-1}^{(N)}(y_{n-1}, x_n)}{\int_{\mathfrak{S}_{n-1}^{(N)}} p_{n-2}^{(N)}(z_{n-2}, \eta)p_{n-1}^{(N)}(\eta, x_n)\mu_{n-1}^{(N)}(\mathrm{d}\eta)}\mu_{n-2}^{(N)}(\mathrm{d}z_{n-2})\mu_{n-1}^{(N)}(\mathrm{d}y_{n-1})\mu_n^{(N)}(\mathrm{d}x_n).$$

We claim that L is uniformly elliptic with background measures $m_n^{(N)}$. In what follows we fix N, suppose $x_i, y_i, z_i \in \mathfrak{S}_i$, and write $p_n^{(N)} = p$ whenever the subscript is clear from the variables. Let $P(\underline{L}_n, \underline{L}_{n+1}) := p(z_{n-2}, z_{n-1})p(x_n, x_{n+1})$. Then $\pi_n^{(N)}(\underline{L}_n, \mathrm{d}\underline{L}_{n+1}) = P(\underline{L}_n, \underline{L}_{n+1})m_{n+1}(\mathrm{d}\underline{L}_{n+1})$,

By the ellipticity assumption on X, $P(\underline{L}_n, \underline{L}_{n+1}) \leq \epsilon_0^{-2}$. In addition,

$$\begin{aligned}
&\int P(\underline{L}_n, \underline{L}_{n+1})P(\underline{L}_{n+1}, \underline{L}_{n+2})m_{n+1}(\mathrm{d}\underline{L}_{n+1}) \\
&= \iiint p(z_{n-2}, z_{n-1})p(x_n, x_{n+1})p(z_{n-1}, z_n)p(x_{n+1}, x_{n+2})\times \\
&\qquad\qquad \times \frac{p(z_{n-1}, y_n)p(y_n, x_{n+1})}{\int p(z_{n-1}, \eta)p(\eta, x_{n+1})\mu_n(\mathrm{d}\eta)}\mu_{n-1}(\mathrm{d}z_{n-1})\mu_n(\mathrm{d}y_n)\mu_{n+1}(\mathrm{d}x_{n+1}) \\
&= \iint p(z_{n-2}, z_{n-1})p(x_n, x_{n+1})p(z_{n-1}, z_n)p(x_{n+1}, x_{n+2})\mu_{n-1}(\mathrm{d}z_{n-1})\mu_{n+1}(\mathrm{d}x_{n+1}) \\
&= \int p(z_{n-2}, z_{n-1})p(z_{n-1}, z_n)\mu_{n-1}(\mathrm{d}z_{n-1}) \int p(x_n, x_{n+1})p(x_{n+1}, x_{n+2})\mu_{n+1}(\mathrm{d}x_{n+1}).
\end{aligned}$$

The last expression is bounded below by ϵ_0^2. So the ladder process is uniformly elliptic with ellipticity constant ϵ_0^2. □

We can now prove Lemma 2.17: Suppose f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X. Fix some x in the state space of X_1, and let $d_n^2(\xi, x)$ denote the structure constants of f on X, conditioned on $X_1 = x$. We are asked to show that $|d_n^2(\xi) - d_n(\xi, x)|$ decays exponentially in n.

Let L denote the ladder process. By Lemma 2.22 (d), there exists an additive functional Γ on L so that $d_n^2(\xi) = \mathbb{E}(|e^{i\xi\Gamma(\underline{L}_n, \underline{L}_{n+1})} - 1|^2)$. By the uniform ellipticity of the ladder process and (2.11), there are constants $C > 0, \theta \in (0, 1)$ such that

$$\operatorname*{ess\ sup}_{(z',y',x')} \left| \mathbb{E}\left(|e^{i\xi\Gamma(\underline{L}_n, \underline{L}_{n+1})} - 1|^2 \middle| \underline{L}_1 = (z', y', x') \right) - d_n^2(\xi) \right| < C\theta^n.$$

To complete the proof, we construct a probability measure λ such that

$$d_n^2(\xi, x) = \iiint \mathbb{E}\left(|e^{i\xi\Gamma(\underline{L}_n, \underline{L}_{n+1})} - 1|^2 \middle| \underline{L}_1 = (z', y', x') \right) \lambda(dx', dy', dz'),$$

and integrate both sides of the previous inequality. The measure $\lambda(dx', dy', dz')$ is the measure such that x' is equal to x, z' is sampled from the distribution of X_3 conditioned on $X_1 = x$, and y' conditioned on x', z' has the bridge distribution. □

2.4 γ-Step Ellipticity Conditions

In this section, we discuss some useful variants of the uniform ellipticity condition. Suppose X is a Markov array with row lengths $k_N + 1$ and transition probabilities $\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(dy)$.

The **one-step ellipticity condition** is that for some $\epsilon_0 > 0$, for all $N \geq 1$ and $1 \leq n \leq k_N$, and for every $x \in \mathfrak{S}_n^{(N)}$ and $y \in \mathfrak{S}_{n+1}^{(N+1)}$, $\epsilon_0 < p_n^{(N)}(x, y) \leq \epsilon_0^{-1}$.

The **γ-step ellipticity condition** ($\gamma = 2, 3, \dots$) is that for some $\epsilon_0 > 0$, for all $N \geq 1$ and $n \leq k_N$, $0 \leq p_n^{(N)} \leq 1/\epsilon_0$; And for all $n \leq k_N - \gamma + 1$ and every $x \in \mathfrak{S}_n^{(N)}$ and $z \in \mathfrak{S}_{n+\gamma}^{(N)}$, the iterated integral

$$\int_{\mathfrak{S}_{n+1}^{(N)}} \cdots \int_{\mathfrak{S}_{n+\gamma-1}^{(N)}} p_n^{(N)}(x, y_1) \prod_{i=1}^{\gamma-2} p_{n+i}^{(N)}(y_i, y_{i+1}) p_{n+\gamma-1}^{(N)}(y_{\gamma-1}, z)\, \mu_{n+1}(dy_1) \cdots \mu_{n+\gamma-1}(dy_{\gamma-1})$$

is bigger than ϵ_0 (with the convention that $\prod_{i=1}^{0} := 1$).

The ellipticity condition we use in this work corresponds to $\gamma = 2$. This is weaker than the one-step condition, but stronger than the γ-step condition for $\gamma \geq 3$.

The results of this work could in principle be proved assuming only a γ-step condition with $\gamma \geq 2$. To do this, one needs to replace the space of hexagons by the space of $2(\gamma + 1)$-gons $\left(x_{n-\gamma}; \begin{matrix} x_{n-\gamma+1} \\ y_{n-\gamma+1} \end{matrix} \cdots \begin{matrix} x_n \\ y_n \end{matrix}; y_{n+1} \right)$ with its associated structure constants, and its associated **γ-ladder process** $\underline{L}_n^{(N)} = (Z_{n-\gamma}^{(N)}, Y_{n-\gamma+1}^{(N)}, \dots, Y_{n-1}^{(N)}, X_n^{(N)})$. Since no new ideas are needed, and since our notation is already heavy enough as it is, we will only treat the case $\gamma = 2$.

*2.5 Uniform Ellipticity and Strong Mixing Conditions

Suppose $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, and let $\mathscr{A}, \mathscr{B}$ be two sub σ-algebras of $\mathcal{F}$. There are several standard **measures of dependence** between $\mathscr{A}$ and $\mathscr{B}$:

$$\alpha(\mathscr{A},\mathscr{B}) := \sup\{|\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)| : A\in\mathscr{A}, B\in\mathscr{B}\};$$

$$\rho(\mathscr{A},\mathscr{B}) := \sup\left\{|\mathbb{E}(fg)-\mathbb{E}(f)\mathbb{E}(g)| : \begin{array}{l} f\in L^2(\mathscr{A}), g\in L^2(\mathscr{B}); \\ \|f-\mathbb{E}(f)\|_2 = 1, \|g-\mathbb{E}(g)\|_2 = 1\end{array}\right\};$$

$$\phi(\mathscr{A},\mathscr{B}) := \sup\left\{\left|\mathbb{P}(B|A)-\mathbb{P}(B)\right| : A\in\mathscr{A}, B\in\mathscr{B}, \mathbb{P}(A)\neq 0\right\};$$

$$\psi(\mathscr{A},\mathscr{B}) := \sup\left\{\left|\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)\mathbb{P}(B)}-1\right| : A\in\mathscr{A}, B\in\mathscr{B} \text{ with non-zero probabilities}\right\}.$$

If one of these quantities vanishes then they all vanish, and this happens iff $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$ for all $A\in\mathscr{A}$, $B\in\mathscr{B}$. In this case we say that $\mathscr{A},\mathscr{B}$ are **independent**. In the dependent case, α,ρ,ϕ,ψ can be used to bound the covariance between (certain) $\mathscr{A}$-measurable and $\mathscr{B}$-measurable random variables:

Theorem 2.23 *Suppose X is $\mathscr{A}$-measurable and Y is $\mathscr{B}$-measurable, then*

(1) $|\mathrm{Cov}(X,Y)|\le 8\alpha(\mathscr{A},\mathscr{B})^{1-\frac{1}{p}-\frac{1}{q}}\|X\|_p\|Y\|_q$ *whenever* $p\in(1,\infty]$, $q\in(1,\infty]$, $\frac{1}{p}+\frac{1}{q}<1$, $X\in L^p$, $Y\in L^q$.
(2) $|\mathrm{Cov}(X,Y)|\le \rho(\mathscr{A},\mathscr{B})\|X-\mathbb{E}Y\|_2\|Y-\mathbb{E}Y\|_2$ *whenever* $X,Y\in L^2$.
(3) $|\mathrm{Cov}(X,Y)|\le 2\phi(\mathscr{A},\mathscr{B})\|X\|_1\|Y\|_\infty$ *whenever* $X\in L^1, Y\in L^\infty$.
(4) $|\mathrm{Cov}(X,Y)|\le \psi(\mathscr{A},\mathscr{B})\|X\|_1\|Y\|_1$ *whenever* $X\in L^1, Y\in L^1$.

See [16, vol 1, ch. 3]. Here are some useful inequalities [16, vol 1, Prop. 3.11]):

Theorem 2.24 *If $(\Omega,\mathscr{F},\mathbb{P})$ is a probability space, and $\mathscr{A},\mathscr{B}$ are sub-σ-algebras of $\mathscr{F}$, then $\alpha:=\alpha(\mathscr{A},\mathscr{B})$, $\rho:=\rho(\mathscr{A},\mathscr{B})$, $\phi:=\phi(\mathscr{A},\mathscr{B})$, $\psi:=\psi(\mathscr{A},\mathscr{B})$ satisfy*

$$2\alpha\le\phi\le\frac{1}{2}\psi\ ,\ 4\alpha\le\rho\le 2\sqrt{\phi}. \tag{2.35}$$

We can use the measures of dependence to define various **mixing conditions**. Let $\mathsf{X}:=\{X_n\}_{n\ge1}$ be a general stochastic process, not necessarily stationary or Markov.

Let $\mathscr{F}_1^n$ denote the σ-algebra generated by $X_1,\dots,X_n$, and let $\mathscr{F}_m^\infty$ denote the σ-algebra generated by X_k for $k\ge m$.

(1) X is called ***α*-mixing**, if $\alpha(n):=\sup_{k\ge1}\alpha(\mathscr{F}_1^k,\mathscr{F}_{k+n}^\infty)\xrightarrow[n\to\infty]{}0$.
(2) X is called ***ρ*-mixing**, if $\rho(n):=\sup_{k\ge1}\rho(\mathscr{F}_1^k,\mathscr{F}_{k+n}^\infty)\xrightarrow[n\to\infty]{}0$.
(3) X is called ***ϕ*-mixing**, if $\phi(n):=\sup_{k\ge1}\phi(\mathscr{F}_1^k,\mathscr{F}_{k+n}^\infty)\xrightarrow[n\to\infty]{}0$.
(4) X is called ***ψ*-mixing**, if $\psi(n):=\sup_{k\ge1}\psi(\mathscr{F}_1^k,\mathscr{F}_{k+n}^\infty)\xrightarrow[n\to\infty]{}0$.

By (2.35), ψ-mixing $\Rightarrow$ ϕ-mixing $\Rightarrow$ ρ-mixing $\Rightarrow$ α-mixing. These implications are strict, see [16, vol 1, §5.23].

Let us see what is the relation between uniform ellipticity and these conditions. First we will show that uniform ellipticity implies exponential ψ-mixing, and then we will give a weak converse of this statement for finite state Markov chains.

Proposition 2.25 *Let X be a uniformly elliptic Markov chain, then for every $x\in\mathfrak{S}_1$, X conditioned on $X_1=x$ is ψ-mixing. Moreover, $\alpha(n),\rho(n),\phi(n),\psi(n)\xrightarrow[n\to\infty]{}0$ exponentially fast, uniformly in x.*

Proof We need the following fact:

Claim. *There exists a constant K which only depends on the ellipticity constant of X, as follows. For every $x\in\mathfrak{S}_1$, $k\ge2$, and for every bounded measurable function $h_k:\mathfrak{S}_k\to\mathbb{R}$, we have the inequality $\|\mathbb{E}_x(h_k(X_k)|X_{k-2})\|_\infty\le K\mathbb{E}_x(|h_k(X_k)|)$.*

Proof of the Claim. By the uniform ellipticity of X, the transition kernels of X can be put in the form $\pi_{n,n+1}(x,dy)=p_n(x,y)\mu_{n+1}(dy)$, where $0\le p_n\le\epsilon_0^{-1}$ and $\int p_n(x,y)p_{n+1}(y,z)\mu_{n+1}(dy)>\epsilon_0$. In addition, Prop. 2.8, applied to X with the initial distribution $\pi(dx)=$ point mass at x, tells us that the Radon-Nikodym derivative of μ_{n+1} with respect to the measure $\mathbb{P}_x(X_{n+1}\in E)$ is a.e. in $[\epsilon_0,\epsilon_0^{-1}]$.

It follows that for all ξ,

$$|\mathbb{E}_x(h_{k+2}(X_{k+2})|X_k = \xi)| \leq \iint p_k(\xi, y)p_{k+1}(y, z)|h_{k+2}(z)|\mu_{k+1}(\mathrm{d}y)\mu_{k+2}(\mathrm{d}z)$$
$$\leq \epsilon_0^{-2} \int |h_{k+2}(z)|\mu_{k+2}(\mathrm{d}z) \leq \epsilon_0^{-3}\mathbb{E}_x(|h_{k+2}(X_{k+2})|).$$

We now prove the proposition. Fix $x \in \mathfrak{S}_1$, and let ψ_x denote the ψ measure of dependence for X conditioned on $X_1 = x$. Let $\mathscr{F}_k$ denote the σ-algebra generated by X_k. Using the Markov property, it is not difficult to see that $\psi_x(n) = \sup_{k\geq 1} \psi_x(\mathscr{F}_k, \mathscr{F}_{k+n})$, see [16, vol 1, pp. 206–7].

Suppose now that $n > 2$, and fix some $x \in \mathfrak{S}_1$, and $A \in \mathscr{F}_k, B \in \mathscr{F}_{k+n}$ with positive $\mathbb{P}_x$-measure. Let $h_k := 1_A(X_k)$ and $h_{k+n} := 1_B(X_{k+n}) - \mathbb{P}_x(B)$. Then

$$|\mathbb{P}_x(A \cap B) - \mathbb{P}_x(A)\mathbb{P}_x(B)| = |\mathbb{E}_x(h_k h_{k+n})| = |\mathbb{E}_x(\mathbb{E}_x(h_k h_{k+n}|\mathscr{F}_k))|$$
$$= |\mathbb{E}_x(h_k\mathbb{E}_x(h_{k+n}|X_k))| \leq \mathbb{E}_x(|h_k|)\|\mathbb{E}_x(h_{k+n}|X_k)\|_\infty = \mathbb{P}_x(A)\|\mathbb{E}_x(\mathbb{E}_x(h_{k+n}|X_{k+n-2}, \ldots, X_1)|X_k)\|_\infty$$
$$= \mathbb{P}_x(A)\|\mathbb{E}_x(\mathbb{E}_x(h_{k+n}|X_{k+n-2})|X_k)\|_\infty, \text{ by the Markov property}$$
$$\leq \mathbb{P}_x(A) \cdot C_{mix}\theta^{n-2}\|\mathbb{E}_x(h_{k+n}|X_{k+n-2})\|_\infty, \text{ by uniform ellipticity and (2.11).}$$

The constants C_{mix} and θ are independent of x, because the Markov chains $\mathsf{X}|X_1 = x$ all have the same transition probabilities, and therefore the same ellipticity constant.

Invoking the claim, we find that

$$|\mathbb{P}_x(A \cap B) - \mathbb{P}_x(A)\mathbb{P}_x(B)| \leq \mathbb{P}_x(A) \cdot C_{mix}\theta^{n-2} \cdot K\mathbb{E}_x(|h_{k+n}|) \leq 2KC_{mix}\theta^{n-2}\mathbb{P}_x(A)\mathbb{P}_x(B).$$

Dividing by $\mathbb{P}_x(A)\mathbb{P}_x(B)$ and passing to the supremum over $A \in \mathscr{F}_k, B \in \mathscr{F}_{k+n}$, gives $\psi_x(n) \leq 2KC_{mix}\theta^{n-2}$. So $\psi_x(n) \to 0$ exponentially fast, uniformly in x. By (2.35), $\alpha_x(n), \rho_x(n), \phi_x(n) \to 0$ exponentially fast, uniformly in x. □

Proposition 2.26 *Let* X *be a Markov chain with the following properties:*

(1) $\exists\kappa > 0$ *such that* $\mathbb{P}(X_n = x) > \kappa$ *for every* $n \geq 1$ *and* $x \in \mathfrak{S}_n$ *(so* $|\mathfrak{S}_n| < 1/\kappa$*); (2)* $\phi(n) \xrightarrow[n\to\infty]{} 0$.

Then X *satisfies the* γ*-step ellipticity condition for all* γ *large enough.*

Proof By (1), all state spaces are finite. Define a measure on $\mathfrak{S}_n$ by $\mu_n(E) = \mathbb{P}(X_n \in E)$, and let $p_n(x, y) := \dfrac{\mathbb{P}(X_{n+1} = y|X_n = x)}{\mathbb{P}(X_{n+1} = y)}$. This is well-defined by (1), and

(a) By construction, $\pi_{n,n+1}(x, \mathrm{d}y) = p_n(x, y)\mu_{n+1}(\mathrm{d}y)$.
(b) By (1), $p_n(x, y) \leq 1/\mathbb{P}(X_{n+1} = y) \leq \kappa^{-1}$.
(c) By (2), for all γ large enough, $\phi(\gamma) < \frac{1}{2}\kappa$. For such γ,

$$\int_{\mathfrak{S}_{n+1}} \cdots \int_{\mathfrak{S}_{n+\gamma-1}} p_n(x, y_1) \prod_{i=1}^{\gamma-2} p_{n+i}(y_i, y_{i+1})p_{n+\gamma-1}(y_{\gamma-1}, z)\, \mu_{n+1}(\mathrm{d}y_1) \cdots \mu_{n+\gamma}(\mathrm{d}y_{n+\gamma-1})$$
$$= \mathbb{P}(X_{n+\gamma} = z|X_n = x) \geq \mathbb{P}(X_{n+\gamma} = z) - \phi(\mathscr{F}_n, \mathscr{F}_{n+\gamma}) \geq \kappa - \phi(\gamma) > \frac{1}{2}\kappa.$$

We obtain the γ-ellipticity condition with ellipticity constant $\frac{1}{2}\kappa$. □

2.6 Reduction to Point Mass Initial Distributions

In this section we explain how to reduce limit theorems for general Markov arrays to the special case when the initial distributions are Dirac measures.

Lemma 2.27 *Suppose* f *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X*. Then there is an a.s. uniformly bounded additive functional* g *on a uniformly elliptic Markov array* Y *such that:*

1. *The initial distributions of* Y *are point mass measures;*
2. $S_N(\mathsf{f})$ *and* $S_N(\mathsf{g})$ *are equal in distribution, for all* N*;*
3. $u^{(N)}(\mathsf{g}) = u^{(N)}(\mathsf{f})$ *and* $d_n^{(N)}(\xi, \mathsf{f}) = d_n^{(N)}(\xi, \mathsf{g})$ *for all* $\xi \in \mathbb{R}$*, and* $5 \le n \le k_N$*.*

Proof Suppose X has row lengths $k_N + 1$, state spaces $\mathfrak{S}_n^{(N)}$, initial distributions $\pi^{(N)}(dx)$, and transition probabilities $\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(dy)$, where $p_n^{(N)}$ and $\mu_n^{(N)}$ satisfy the uniform ellipticity conditions, with constant ϵ_0.

Construct a Markov array Y, with the following data:

- *Row Lengths*: $k_N + 1$.
- *State Spaces*: $\widetilde{\mathfrak{S}}_1^{(N)} = \{x_0\}$ (a single point), $\widetilde{\mathfrak{S}}_2^{(N)} := \mathfrak{S}_1^{(N)} \times \mathfrak{S}_2^{(N)}$, and $\widetilde{\mathfrak{S}}_n^{(N)} := \mathfrak{S}_n^{(N)}$ for $3 \le n \le k_N + 1$.
- *Initial Distributions*: $\widetilde{\pi}^{(N)}$:= point mass measure at x_0.
- *Transition Probabilities*: $\widetilde{\pi}_{n,n+1}^{(N)}(x, dy) := \widetilde{p}_n^{(N)}(x, y)\widetilde{\mu}_{n+1}^{(N)}(dy)$, where
  - $\widetilde{p}_1^{(N)} \equiv 1$ and $\widetilde{\mu}_2^{(N)}(dx_1, dx_2) := p_1^{(N)}(x_1, x_2)\pi^{(N)}(dx_1)\mu_2^{(N)}(dx_2)$;
  - $\widetilde{p}_2^{(N)}((x_1, x_2), x_3) := p_2^{(N)}(x_2, x_3)$ and $\widetilde{\mu}_3^{(N)} := \mu_3^{(N)}$;
  - $\widetilde{p}_n^{(N)} := p_n^{(N)}$ and $\widetilde{\mu}_{n+1}^{(N)} := \mu_{n+1}^{(N)}$ for $3 \le n \le k_N$.

Note that $(Y_1^{(N)}, \dots, Y_{k_N+1}^{(N)})$ and $(x_0, (X_1^{(N)}, X_2^{(N)}), X_3^{(N)}, \dots, X_{k_N+1}^{(N)})$ have the same joint distribution, and the initial distributions of Y are point mass measures.

Next construct an additive functional g on Y: $g_1^{(N)}(x_0, (x_1, x_2)) := f_1^{(N)}(x_1, x_2)$, $g_2^{(N)}((x_1, x_2), x_3) := f_2^{(N)}(x_2, x_3)$, and $g_n^{(N)} \equiv f_n^{(N)}$ for $3 \le n \le k_N$.

Clearly $\operatorname{ess\,sup} |\mathsf{g}| < \infty$, and $S_N(\mathsf{f})$ and $S_N(\mathsf{g})$ are equal in distribution, for all N.

We check that Y is uniformly elliptic. Clearly, $0 \le \widetilde{p}_i^{(N)} \le \epsilon_0^{-1}$ for all i. Next,

$$\int \widetilde{p}_1^{(N)}(x_0, (x_1, x_2))\widetilde{p}_2^{(N)}((x_1, x_2), x_3)\widetilde{\mu}_2^{(N)}(dx_1, dx_2) = \iint p_2^{(N)}(x_2, x_3)p_1^{(N)}(x_1, x_2)\mu_2^{(N)}(dx_2)\pi^{(N)}(dx_1) \ge \epsilon_0,$$

$$\text{and} \quad \int \widetilde{p}_2^{(N)}((x_1, x_2), x_3)\widetilde{p}_3^{(N)}(x_3, x_4)\widetilde{\mu}_3^{(N)}(dx_3) = \int p_2^{(N)}(x_2, x_3)p_3^{(N)}(x_3, x_4)\mu_3^{(N)}(dx_3) \ge \epsilon_0.$$

For $i \ge 3$, $\int \widetilde{p}_i^{(N)}(x_i, x_{i+1})\widetilde{p}_{i+1}^{(N)}(x_{i+1}, x_{i+2})\widetilde{\mu}_{i+1}^{(N)}(dx_{i+1}) \ge \epsilon_0$ is obvious. □

2.7 Notes and References

For a comprehensive treatment of Markov chains on general state spaces, see Doob's book [62]. For a comprehensive account of mixing conditions, see [16].

Uniform ellipticity is one of a plethora of contraction conditions for Markov operators, that were developed over the years as sufficient conditions for mixing results such as Proposition 2.13. We mention in particular the works of Markov [137], Doeblin [51, 52], Hajnal [94], Doob [62], and Dobrushin [50] (see also Seneta [179] and Sethuraman & Varadhan [181]).

We note that in analysis literature the notion of *ellipticity* has a different meaning. A second order differential operator D on a d dimensional manifold M is called elliptic if it can be written in local coordinates as

$$D\phi = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(z)\frac{\partial^2\phi}{\partial z_i \partial z_j} + \sum_{j=1}^{d} b_j(z)\frac{\partial\phi}{\partial z_j}$$

where matrix (a_{ij}) is positive definite.

If $\{X_t\}$ is a diffusion process on a *compact* manifold then considering our process at integer times we obtain the process which satisfies one step ellipticity condition in our sense. In fact, a weaker condition called *hypoellipticity* introduced in [101] is sufficient for this purpose.

However, if the phase space is not compact, then the analytic and probabilistic notions of ellipticity are different. For example, the Brownian Motion considered in §3.3 is elliptic in the analytic sense but not in the probabilistic sense. In fact, the Brownian Motion is null recurrent and the theory of this book does not apply to it. We refer the reader to [195] and [102] for more information about elliptic and hypoelliptic operators.

Proposition 2.8 is similar in spirit to Doeblin's estimates for the stationary probability vector of a Markov chain satisfying Doeblin's condition [51, 52].

The contraction coefficients in section 2.2.2 are also called "ergodicity coefficients." They play a major role in Dobrushin's proof of the CLT for inhomogeneous Markov chains [50]. Our treatment of contraction coefficients follows [181] closely. Lemma 2.11 and its proof are taken from there.

The construction we call "change of measure" is crucial for the analysis of large deviations, see §7.3.1.

The quantities $\mathfrak{D}(X,\xi)$ were first used by Mukhin in [148]. They play a central role in the local limit theorem for sums of independent random variables. For details and additional references, see §§8.2, 8.7.

The balance of hexagon is related to classical constructions in dynamical systems, which we would like to explain. Consider invertible maps $\mathcal{T}_n : \Omega_n \to \Omega_{n+1}$ between metric spaces (Ω_n, d_n). Given $a_0 \in \Omega_0$, let

$$a_n := \begin{cases} (\mathcal{T}_{n-1} \circ \cdots \circ \mathcal{T}_0)(a_0) & n > 0, \\ (\mathcal{T}_{-1} \circ \cdots \circ \mathcal{T}_{-n})^{-1}(a_0) & n < 0. \end{cases}$$

This has the merit that for all $n > m$ in $\mathbb{Z}$, $(\mathcal{T}_n \circ \mathcal{T}_{n-1} \circ \cdots \mathcal{T}_m)(a_m) = a_n$.

- We say that $a_0, b_0 \in \Omega_0$ are *in the same stable manifold*, and write $a_0 \sim_s b_0$, if $d_n(a_n, b_n) \to 0$ exponentially as $n \to +\infty$.
- We say that $a_0, b_0 \in \Omega_0$ are *in the same unstable manifold*, and write $a_0 \sim_u b_0$, if $d_n(a_n, b_n) \to 0$ exponentially as $n \to -\infty$.

(In the classical dynamical setup the equivalence classes are indeed submanifolds, but this is not the case in the general setup we consider.)

Given a sequence of uniformly Hölder functions $f_n : \Omega_n \to \mathbb{R}$, and points $a_0, b_0, c_0, d_0 \in \Omega_0$ such that $a_0 \sim_s b_0, \quad c_0 \sim_s d_0, \quad a_0 \sim_u d_0, \quad b_0 \sim_u c_0$ we define the *periodic cycle functional*

$$\Delta(a,b,c,d) = \sum_{n\in\mathbb{Z}} [f_n(a_n) - f_n(b_n) + f_n(c_n) - f_n(d_n)].$$

To see that the series converges, use the decomposition $\Delta = \sum_{n\geq 0}[f_n(a_n) - f_n(b_n)] + \sum_{n\geq 0}[f_n(c_n) - f_n(d_n)] + \sum_{n<0}[f_n(a_n) - f_n(d_n)] + \sum_{n<0}[f_n(c_n) - f_n(b_n)]$.

To relate this expression to our setting we assume that our Markov chain is defined for all $n \in \mathbb{Z}$ (if it is not the case we can extend it to negative n in an arbitrary way so that the ellipticity conditions are satisfied). Let Ω_n be the space of sequences $\{X_k\}$ with $X_k \in \mathfrak{S}_{n+k}$ and put $d_n(\{X_k\},\{Z_k\}) = 2^{-\max(\ell : X_j = Z_j \text{ for } |j|<\ell)}$. We regard $f_n(X_n, X_{n+1})$ as a functions on Ω_n which depend only on coordinates 0 and 1 of a sequence $\{X_k\}$ from Ω_n.

Let $\mathcal{T}_n$ be the shift. Given $\{X_k\}, \{Z_k\} \in \Omega_0, Y_{n-1}, Y_n$ let

$$\begin{aligned} a_0 = c_0 &= \{\dots Z_{n-3}, Z_{n-2}, Z_{n-1}, Y_n, X_{n+1}, X_{n+2}, \dots\}, \\ b_0 = d_0 &= \{\dots Z_{n-3}, Z_{n-2}, Y_{n-1}, X_n, X_{n+1}, X_{n+2}, \dots\}. \end{aligned}$$

A direct computation shows that $\Delta(a_0, b_0, c_0, d_0) = 2\Gamma\left(Z_{n-2}, \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix}, X_{n+1}\right)$.

In the case where $(\Omega_n, \mathcal{T}_n) = (\Omega, \mathcal{T})$ do not depend on n, $\Delta(\cdot)$ appears in several problems associated to dynamics of $\mathcal{T}$ (see [19, 106]). The relevance to mixing properties is noted in [25], (cf. (6.5) in the present text.)

The application to cancellation properties of twisted transfer operators (called in Chapters 5 and 6 "perturbation operators") appear in [54]. Some of ideas of [54] are employed in Chapter 5, see in particular, Lemma 5.6. One can also define cycles of length greater than four. Such cycles could often be studied by breaking them into shorter cycles, cf. the discussion after (9.16).

We end with a warning. It is tempting to speak loosely of the hexagon measure m_{Hex} as "the distribution of pairs of independent paths $(x_{n-1}, *, *, x_{n+1})$ with the same beginning and end," but this is misleading.

Specifically, if the state spaces of X are discrete (or more generally if the measures $\mu_n^{(N)}$ are all atomic), then there is a well-defined measure m''_{Hex} on $\text{Hex}\,(N, n)$ obtained by taking two independent copies $\{X_n^{(N)}\}, \{Y_n^{(N)}\}$ of X, and looking at the distribution of $\begin{pmatrix} & X_{n-1}^{(N)} & X_n^{(N)} & \\ X_{n-2}^{(N)} & & & Y_{n+1}^{(N)} \\ & Y_{n-1}^{(N)} & Y_n^{(N)} & \end{pmatrix}$ conditioned on the event $\{X_{n-2}^{(N)} = Y_{n-2}^{(N)}, X_{n+1}^{(N)} = Y_{n+1}^{(N)}\}$ (this event has positive measure by discreteness and uniform ellipticity). The measures m_{Hex} and m''_{Hex} are quite different, and the Radon-Nikodym derivative $\frac{\mathrm{d}m''_{\text{Hex}}}{\mathrm{d}m_{\text{Hex}}}$ does not even have to be uniformly bounded away from zero and infinity in N. The reader is invited to compare the two measures in the special case when X_n are independent. (We thank E. Solan for this observation.)

Chapter 3
Variance Growth, Center-Tightness, and the Central Limit Theorem

Abstract We analyze the variance of $S_N = f_1(X_1, X_2) + \cdots + f_N(X_N, X_{N+1})$, and characterize the additive functionals for which $\text{Var}(S_N) \not\to \infty$. Then we prove Dobrushin's theorem: If $\text{Var}(S_N) \to \infty$, then S_N satisfies the central limit theorem.

3.1 Main Results

Let f be an additive functional on a Markov array X with row lengths $k_N + 1$. We let $S_N = \sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$.

For Markov chains, $k_N = N$, and $S_N = \sum_{i=1}^{N} f_i(X_i, X_{i+1})$.

3.1.1 Center-Tightness and Variance Growth

We say that f is **center-tight** if there are constants m_N such that for every $\epsilon > 0$, there exists M for which $\mathbb{P}[|S_N - m_N| > M] < \epsilon$ for all N.

We shall see in Theorem 3.8 below that f is center-tight iff $\text{Var}(S_N) \not\to \infty$. Obviously, in such a situation the right hand side in $\mathbb{P}[S_N - z_N \in (a,b)] \overset{?}{\sim} \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$ can be made larger than one by choosing $|a - b|$ sufficiently big, and the asymptotic relation in the "standard" LLT fails. One could hope for a different universal asymptotic behavior, but this is hopeless:

Example 3.1 (Non-Universality in the LLT for Center-Tight Functionals) Let X_n be identically distributed independent random variables with uniform distribution on $[0, 1]$. Choose an *arbitrary* sequence of random variables $\{Z_n\}_{n\geq 1}$ taking values in $[0, 1]$. By the isomorphism theorem for Lebesgue spaces, there are measurable functions $g_n : [0, 1] \to [0, 1]$ such that $g_0 \equiv 0$, and $g_n(X_n) = Z_n$ in distribution. Let $f_n(X_n, X_{n+1}) := g_{n+1}(X_{n+1}) - g_n(X_n)$. Then $S_N = Z_{N+1}$ in distribution, f is center-tight, and $\mathbb{P}(S_N \in (a,b)) = \mathbb{P}(Z_{N+1} \in (a,b))$ is completely arbitrary.

Every Markov array admits center-tight additive functionals. Here are three constructions which lead to such examples (in the uniformly bounded, uniformly elliptic case, *all* center-tight additive functionals arise this way, see Theorem 3.8 below):

Example 3.2 (Gradients) A gradient f on a Markov chain X is an additive functional of the form $f_n(x,y) := a_{n+1}(y) - a_n(x)$, where $a_n : \mathfrak{S}_n \to \mathbb{R}$ are measurable, and $\text{ess sup}\,|\mathsf{a}| < \infty$. Similarly, gradients for Markov arrays are defined by $f_n^{(N)}(x,y) = a_{n+1}^{(N)}(y) - a_n^{(N)}(x)$, where $a_n^{(N)} : \mathfrak{S}_n^{(N)} \to \mathbb{R}$ are measurable and uniformly essentially bounded. We write

$$\mathsf{f} = \nabla \mathsf{a},$$

and say that f is the **gradient** of a and a is the **potential** of f.[1]

If $\mathsf{f} = \nabla \mathsf{a}$, then $S_N(\mathsf{f})$ is telescopic, and $|S_N(\mathsf{f})| \leq 2\text{ess sup}\,|\mathsf{a}|$. So f is center-tight (take $m_N := 0$ and $M := 3\text{ess sup}\,|\mathsf{a}|$).

[1] In the ergodic theoretic literature, f is called a **coboundary** and a is called a **transfer function**.

Example 3.3 (Summable Variance) An additive functional f has **summable variance**, if it is a.s. uniformly bounded, and $V_\infty < \infty$, where

$$V_\infty := \begin{cases} \sum\limits_{n=1}^{\infty} \mathrm{Var}[f_n(X_n, X_{n+1})] & \mathsf{X} \text{ is a Markov chain} \\ \sup\limits_N \sum\limits_{n=1}^{k_N} \mathrm{Var}[f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})] & \mathsf{X} \text{ is a Markov array with row lengths } k_N + 1. \end{cases}$$

If X is uniformly elliptic and $|\mathsf{f}| \leq K$ a.s., then summable variance implies center-tightness. This follows from Chebyshev's inequality and the following lemma:

Lemma 3.4 *Let* f *be a uniformly bounded functional on a uniformly elliptic Markov array. Then* $\mathrm{Var}(S_N) \leq \overline{V}_N \left(1 + \frac{2C_{mix}}{1-\theta}\right)$ *where* $\overline{V}_N := \sum\limits_{n=1}^{k_N} \mathrm{Var}(f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}))$*, and* C_{mix} *and* $0 < \theta < 1$ *are as in Prop. 2.13.*

Proof We give the proof for Markov chains (the proof for arrays is identical). Recall (2.13)

$$\mathrm{Var}\,(S_N) = \sum_{n=1}^{N} \mathrm{Var}(f_n) + 2\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} \mathrm{Cov}(f_n, f_m) \leq \overline{V}_N + 2C_{mix}\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} \theta^{m-n}\sqrt{\mathrm{Var}(f_n)\mathrm{Var}(f_m)},$$

$$\leq \overline{V}_N + 2C_{mix}\sum_{j=1}^{N-1}\theta^j \sum_{n=1}^{N-j}\sqrt{\mathrm{Var}(f_n)\mathrm{Var}(f_{n+j})} \overset{!}{<} \overline{V}_N + \frac{2C_{mix}\overline{V}_N}{1-\theta} \quad (\overset{!}{<} \text{ uses that } ab \leq (a^2+b^2)/2).$$ □

Example 3.5 Suppose X is uniformly elliptic. Then every additive functional of the form $\mathsf{f} = \mathsf{g} + \mathsf{h}$, where g is a gradient and h has summable variance, is center-tight. The proof is a simple union bound, and we omit it.

Henceforth, we assume the following conditions:

(E) $\mathsf{X} = \{X_n^{(N)}\}$ is a **uniformly elliptic** inhomogeneous Markov array with row lengths $k_N + 1$, and ellipticity constant ϵ_0. We denote the state spaces by $\mathfrak{S}_n^{(N)}$, the initial distributions by $\pi^{(N)}$, and the transition probabilities by $\pi_{n,n+1}^{(N)}(x, dy) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(dy)$, where $p_n^{(N)}$ are as in the definition of uniform ellipticity.

(B) $\mathsf{f} = \{f_n^{(N)}\}$ is an **a.s. uniformly bounded** additive functional on X, satisfying the bound $|\mathsf{f}| \leq K$ almost surely.

Let $V_N := \mathrm{Var}(S_N)$ and $U_N := \sum\limits_{n=3}^{k_N}(u_n^{(N)})^2$, where $u_n^{(N)}$ are as in (2.26).

Theorem 3.6 (Variance Growth) *There are constants* $C_1, C_2 > 0$ *which only depend on* ϵ_0, K *such that for all* N*,*

$$C_1^{-1}U_N - C_2 \leq \mathrm{Var}(S_N) \leq C_1 U_N + C_2.$$

Corollary 3.7 *Suppose* X *is a* <u>*Markov chain*</u>*. Either* $\mathrm{Var}(S_N) \to \infty$ *or* $\mathrm{Var}(S_N)$ *is bounded. Moreover,* $\mathrm{Var}(S_N) \asymp \sum\limits_{n=3}^{N} u_n^2$*, with the* u_n *from* (2.26).

(The the first part of the corollary is clearly false for arrays.) We return to arrays:

Theorem 3.8 $\mathrm{Var}(S_N)$ *is bounded iff* f *is center-tight iff* $\mathsf{f} = \nabla \mathsf{a} + \mathsf{h}$ *where* a *is a uniformly bounded potential, and* h *has summable variance.*

Corollary 3.9 f *is center-tight iff* $\sup_{N \geq 1} U_N < \infty$.

Theorem 3.6 is a statement on the localization of cancellations. In general, if the variance of an additive functional of a stochastic process does not tend to infinity, then there must be some strong cancellations in S_N. A priori, these cancellations may involve summands which are far from each other. Theorem 3.6 says that strong cancellations must already occur among three consecutive terms $f_{n-2}^{(N)} + f_{n-1}^{(N)} + f_n^{(N)}$: this is what $u_N^{(N)}$ measures.

If f depends only on one variable (i.e. $f_n^{(N)}(x, y) = f_n^{(N)}(x)$), and if we have the one-step ellipticity condition $\epsilon_0 \le p_n^{(N)}(x, y) \le \epsilon_0^{-1}$, then one can show that there are constants $\widehat{C}_1, \widehat{C}_2$ such that

$$\widehat{C}_1^{-1} \sum_n \mathrm{Var}(f_n(X_n)) - \widehat{C}_2 \le V_N \le \widehat{C}_1 \left(\sum_n \mathrm{Var}(f_n(X_n)) \right) + \widehat{C}_2 \tag{3.1}$$

(see [50, 181] for an even more general statement). See the end of §3.2.2.

The estimate (3.1) does *not* hold when $f_n^{(N)}$ depends on two variables. For example, if $f_n^{(N)}$ is a gradient with bounded potential, then V_N is bounded, but $\sum_{n=1}^{N} \mathrm{Var}(f_n^{(N)}(X_n, X_{n+1}))$ could be arbitrarily large.

3.1.2 The Central Limit Theorem and the Two-Series Theorem

Theorem 3.10 (Dobrushin) *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X. *If* $\mathrm{Var}(S_N) \to \infty$, *then for every interval,*

$$\mathbb{P}\left[\frac{S_N - \mathbb{E}(S_N)}{\sqrt{\mathrm{Var}(S_N)}} \in (a, b) \right] \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-t^2/2} \mathrm{d}t.$$

The proof we give (due to Sethuraman & Varadhan) is based on McLeish's martingale central limit theorem. This is recalled in §3.2.3.

Suppose X is a Markov chain, and $\mathbb{E}(S_N) = 0$ for all N. Dobrushin's Theorem compares S_N to the Gaussian distribution with variance V_N. In §3.3 we will state and prove the **almost sure invariance principle**, which compares $(S_N, S_{N+1}, \ldots)$ to a path of Brownian motion, at times $V_N, V_{N+1}, \ldots$. One consequence is the **law of the iterated logarithm**

$$\limsup_{N\to\infty} \frac{S_N}{\sqrt{2V_N \ln\ln V_N}} = 1, \quad \liminf_{N\to\infty} \frac{S_N}{\sqrt{2V_N \ln\ln V_N}} = -1.$$

See §3.3 for precise statements and proofs.

Dobrushin's CLT implies that if $V_N \to \infty$, then for any *bounded* continuous function $\phi : \mathbb{R} \to \mathbb{R}$ we have

$$\lim_{N\to\infty} \mathbb{E}\left[\phi\left(\frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \right) \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(z) \mathrm{e}^{-z^2/2} dz.$$

We will now discuss the (unbounded) case $\phi(x) = x^r$ ($r \in \mathbb{N}$):

Theorem 3.11 (Lifshits) *Let* f *be a bounded additive functional of a uniformly elliptic Markov chain such that* $\mathbb{E}(S_N) = 0$ *for all* N. *If* $\mathrm{Var}(S_N) \to \infty$, *then*

$$\lim_{N\to\infty} \frac{\mathbb{E}[S_N^r]}{V_N^{r/2}} = \begin{cases} 0 & r \text{ is odd,} \\ (r-1)!! := (r-1)(r-3)\cdots 3\cdot 1 & r \text{ is even.} \end{cases} \tag{3.2}$$

Recall that the $\boldsymbol{r^{th}}$**-moment** of an L^r random variable X is the number $\mathbb{E}[X^r]$. The right-hand-side of (3.2) is well-known to be the r-th moment of the standard normal distribution. Therefore (3.2) is simply the statement that the moments of $S_N/\sqrt{V_N}$ converge to the moments of the standard Gaussian distribution.

The next result, which describes the case when $\mathrm{Var}(S_N) \not\to \infty$, is a version of the "two-series theorem" of Khintchin and Kolmogorov (originally proved for iid's):

Theorem 3.12 *Let* f *be an a.e. uniformly bounded additive functional of a uniformly elliptic inhomogeneous Markov chain* X.

(1) If $\sum_{n=1}^{\infty} \mathrm{Var}[f_n(X_n, X_{n+1})] < \infty$*, then* $\lim_{n\to\infty}(S_N - \mathbb{E}(S_N))$ *exists a.s., and is finite.*
(2) $\mathrm{Var}(S_N) \not\to \infty$ *iff there exist measurable functions* $a_n : \mathfrak{S}_n \to \mathbb{R}$ *such that* $\operatorname{ess\,sup} |\mathsf{a}| < \infty$*, and* $\lim_{n\to\infty}(S_N - a_{N+1}(X_{N+1}) - \mathbb{E}(S_N))$ *exists a.s., and is finite.*

The theorem makes no sense for Markov arrays. For arrays, S_N live on different uncoupled probability spaces, and they cannot be evaluated at the same point.

Example 3.13 (Optimality of Theorem 3.12) Let X_n be a sequence of iid random variables taking values ± 1 with probability $1/2$.

Let $a_n(x) = \sigma_n x$, and $\mathsf{f} := \nabla \mathsf{a}$, then $S_N = a_{N+1}(X_{N+1}) - a_1(X_1)$, and the a.s. convergence of $S_N - \mathbb{E}(S_N)$ reduces to the a.s. convergence of $a_N(X_N)$.

- If $\sigma_n^2 := 1/n$, then $|a_N| \le 1/\sqrt{N}$ and so $a_N \to 0$ a.e. Thus $\lim(S_N - \mathbb{E}(S_N))$ exists a.s., even though $\sum \mathrm{Var}[f_n] = \infty$. This shows that part 1 of Theorem 3.12 cannot be strengthened to an iff statement.
- If $\sigma_n^2 := 1$, then $a_N(X_N) = X_N$, which oscillates a.s. without converging. So $\lim(S_N - \mathbb{E}(S_N))$ does not exist, even though $\mathrm{Var}(S_N)$ is bounded. However, $S_N - a_{N+1}(X_{N+1}) - \mathbb{E}(S_N)$ converges a.s. (to $-a_1(X_1)$). This shows that sometimes, the term $a_{N+1}(X_{N+1})$ in part 2 of Theorem 3.12 is really necessary.

3.2 Proofs

3.2.1 The Gradient Lemma

Lemma 3.14 (Gradient Lemma) *Suppose* f *is an additive functional on a uniformly elliptic Markov array* X *with state spaces* $\mathfrak{S}_n^{(N)}$*, and assume* $\operatorname{ess\,sup} |\mathsf{f}| \le K$*. Then*

$$\mathsf{f} = \widetilde{\mathsf{f}} + \nabla \mathsf{a} + \mathsf{c},$$

where $\widetilde{\mathsf{f}}, \mathsf{a}, \mathsf{c}$ *are additive functionals on* X *with the following properties:*

(a) $|\mathsf{a}| \le 2K$ *and* $a_n^{(N)}(x)$ *are measurable functions on* $\mathfrak{S}_n^{(N)}$.
(b) $|\mathsf{c}| \le K$ *and* $c_n^{(N)}$ *are constant functions.*
(c) $|\widetilde{\mathsf{f}}| \le 6K$ *and* $\widetilde{f}_n^{(N)}(x, y)$ *satisfy* $\|\widetilde{f}_n^{(N)}\|_2 \le u_n^{(N)}$ *for all* $3 \le n \le k_N + 1$.

If X *is a Markov chain, we can choose* $f_n^{(N)} = f_n$, $a_n^{(N)} = a_n$, $c_n^{(N)} = c_n$.

Proof for Doeblin Chains: Before proving the lemma in full generality, we consider the simple but important special case of Doeblin chains (Example 2.7).

Recall that a Doeblin chain is a Markov chain X with finite state spaces $\mathfrak{S}_n$ of uniformly bounded cardinality, whose transition matrices $\pi_{xy}^n := \pi_{n,n+1}(x, \{y\})$ satisfy the following properties:

(E1) $\exists \epsilon_0' > 0$ s.t. for all $n \ge 1$ and $(x, y) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$, either $\pi_{xy}^n = 0$ or $\pi_{xy}^n > \epsilon_0'$;
(E2) for all n, for all $(x, z) \in \mathfrak{S}_n \times \mathfrak{S}_{n+2}$, $\exists y \in \mathfrak{S}_{n+1}$ such that $\pi_{xy}^n \pi_{yz}^{n+1} > 0$.

We re-label the states in $\mathfrak{S}_n$ so that $\mathfrak{S}_n = \{1, \dots, d_n\}$ where $d_n \le d$, and in such a way that $\pi^n_{11} > 0$ for all n. Assumption (E2) guarantees that for every $n \ge 3$ and every $x \in \mathfrak{S}_n$ there exists a state $\xi_{n-1}(x) \in \mathfrak{S}_{n-1}$ s.t. $\pi^{n-2}_{1,\xi_{n-1}(x)}\pi^{n-1}_{\xi_{n-1}(x),x} > 0$. Let

$$
\begin{aligned}
&a_1 \equiv 0, \quad a_2 \equiv 0, \quad \text{and } a_n(x) := f_{n-2}(1, \xi_{n-1}(x)) + f_{n-1}(\xi_{n-1}(x), x) \text{ for } n \ge 3\\
&c_1 := 0, \quad c_2 := 0, \quad \text{and } c_n := f_{n-2}(1,1) \text{ for } n \ge 3\\
&\widetilde{\mathsf{f}} := \mathsf{f} - \nabla \mathsf{a} - \mathsf{c}.
\end{aligned}
$$

We claim that $\widetilde{\mathsf{f}}, \mathsf{a}, \mathsf{c}$ satisfy our requirements.

To explain why and to motivate the construction, consider the special case $u_n = 0$. In this $\|\widetilde{\mathsf{f}}\|_2 = 0$ and the lemma reduces to constructing functions $b_n : \mathfrak{S}_n \to \mathbb{R}$ s.t. $\mathsf{f} = \nabla \mathsf{b} + \mathsf{c}$. We first try to solve $\mathsf{f} = \nabla \mathsf{b}$ with $\mathsf{c} = 0$. Any solution must satisfy

$$f_n(x, y) = b_{n+1}(y) - b_n(x). \tag{3.3}$$

Keeping x fixed and solving for $b_{n+1}(y)$ we find that

$$b_n(y) = b_2(x_2) + f_2(x_2, x_3) + \cdots + f_{n-2}(x_{n-2}, x_{n-1}) + f_{n-1}(x_{n-1}, y)$$

for all paths $(x_2, \dots, x_{n-1}, y)$ with positive probability. The path $x_2 = \cdots = x_{n-2} = 1$, $x_{n-1} = \xi_{n-1}(y)$ suggests defining

$$b_2 \equiv 0\,,\ b_n(y) := \sum_{k=2}^{n-3} f_k(1,1) + f_{n-2}(1, \xi_{n-1}(y)) + f_{n-1}(\xi_{n-1}(y), y).$$

This works: for every $n \ge 3$, if $\pi^n_{xy} > 0$ then

$$b_{n+1}(y) - b_n(x) = [f_{n-2}(1,1) + f_{n-1}(1, \xi_n(y)) + f_n(\xi_n(y), y) - f_{n-2}(1, \xi_{n-1}(x)) - f_{n-1}(\xi_{n-1}(x), x) - f_n(x,y)] + f_n(x,y)$$

$$\therefore\ b_{n+1}(y) - b_n(x) = \Gamma_n \begin{pmatrix} 1 & \begin{matrix} 1 \\ \xi_{n-1}(x) \end{matrix} & \begin{matrix} \xi_n(y) \\ x \end{matrix} & y \end{pmatrix} + f_n(x,y) \overset{!}{=} f_n(x,y). \tag{3.4}$$

Here is the justification of $\overset{!}{=}$. In the setup we consider, the natural measure on the level n hexagons is atomic, and every admissible hexagon has positive mass. So $u_n = 0$ implies that $\Gamma_n(P) = 0$ for every admissible hexagon, and $\overset{!}{=}$ follows.

We proved (3.3), but we are not yet done because it is not clear that $\operatorname{ess\,sup} |\mathsf{b}| < \infty$.

To fix this decompose $b_n(y) = a_n(y) + \sum_{k=2}^{n-3} f_k(1,1)$. Then $|\mathsf{a}| \le 2K$, and a direct calculation shows that $f_n(x,y) = a_{n+1}(y) - a_n(x) + f_{n-2}(1,1)$, whence $\mathsf{f} = \nabla \mathsf{a} + \mathsf{c}$ with a essentially bounded. This proves the lemma in case $u_n = 0$.

The general case $u_n \ge 0$ is done the same way: (3.4) implies that $\widetilde{\mathsf{f}} := \mathsf{f} - \nabla a - \mathsf{c} \equiv \mathsf{f} - \nabla \mathsf{b}$ is given by

$$\widetilde{f}_n(x,y) = f_n(x,y) - (a_{n+1}(y) - a_n(x)) - c_n = -\Gamma_n \begin{pmatrix} 1 & \begin{matrix} 1 \\ \xi_{n-1}(x) \end{matrix} & \begin{matrix} \xi_n(y) \\ x \end{matrix} & y \end{pmatrix}.$$

If $|\mathsf{f}| \le K$, then $|\Gamma_n| \le 6K$, whence $|\widetilde{\mathsf{f}}| \le 6K$. Next,

$$\|\widetilde{f}_n\|_2^2 \le \mathbb{E}\left[\Gamma_n \begin{pmatrix} 1 & \begin{matrix} 1 \\ \xi_{n-1}(X_n) \end{matrix} & \begin{matrix} \xi_n(X_{n+1}) \\ X_n \end{matrix} & X_{n+1} \end{pmatrix}^2\right].$$

In the scenario we consider the space of admissible hexagons has a finite number of elements, and each has probability uniformly bounded below. So there is a global constant C which only depends on $\sup |\mathfrak{S}_n|$ and on ϵ_0' in (E2) such that

$$\mathbb{E}\left[\Gamma_n\left(1 \;\; \begin{matrix} 1 \\ \xi_{n-1}(X_n) \end{matrix} \;\; \begin{matrix} \xi_n(X_{n+1}) \\ X_n \end{matrix} \;\; X_{n+1}\right)^2\right] \le C\mathbb{E}[\Gamma(P)^2],$$

where P is a random hexagon in (Hex $(n), m_{\mathrm{Hex}}$). So $\|\widetilde{\mathsf{f}}\|_2 \le \sqrt{C} \cdot u_n^2$.

(The gradient lemma says that we can choose a and c so that $C = 1$. The argument we gave does not quite give this, but the value of the constant is not important for the applications we have in mind.)

The Proof of the Gradient Lemma in the General Case: Recall the ladder process $\mathsf{L} = \{\underline{L}_n^{(N)}\}$, $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$ from §2.3.3. In what follows we omit the superscripts $^{(N)}$ on the right hand side of identities.

Define $F_n^{(N)}(\underline{L}_n^{(N)}) := F_n(\underline{L}_n) := f_{n-2}(Z_{n-2}, Y_{n-1}) + f_{n-1}(Y_{n-1}, X_n)$ and

$$\Gamma_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) := \Gamma_n(\underline{L}_n, \underline{L}_{n+1}) := \Gamma\left(Z_{n-2} \;\; \begin{matrix} Z_{n-1} \\ Y_{n-1} \end{matrix} \;\; \begin{matrix} Y_n \\ X_n \end{matrix} \;\; X_{n+1}\right), \quad \text{see (2.25).}$$

Then we have the following identity:

$$f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) = F_{n+1}(\underline{L}_{n+1}) - F_n(\underline{L}_n) + f_{n-2}(Z_{n-2}, Z_{n-1}) - \Gamma_n(\underline{L}_n, \underline{L}_{n+1}). \tag{3.5}$$

Next define $a_n^{(N)} : \mathfrak{S}_n^{(N)} \to \mathbb{R}$ and $c_n^{(N)} \in \mathbb{R}$ by

$$a_n^{(N)}(\xi) := \mathbb{E}(F_n(\underline{L}_n)|X_n = \xi) \quad (3 \le n \le k_N), \tag{3.6}$$
$$c_n^{(N)} := \mathbb{E}[f_{n-2}(Z_{n-2}, Z_{n-1})]. \tag{3.7}$$

By assumption, $|\mathsf{f}| \le K$, so $|\mathsf{a}| \le 2K$ and $|\mathsf{c}| \le K$.

Let $\widetilde{\mathsf{f}} := \mathsf{f} - \nabla \mathsf{a} - \mathsf{c}$. To prove the lemma, we need to bound $\widetilde{\mathsf{f}}$ in L^∞, and in L^2.

Claim: *For every* $(\xi, \eta) \in \mathfrak{S}_n \times \mathfrak{S}_{n+1}$,

$$c_n^{(N)} = \mathbb{E}\left[\mathbb{E}\left(f_{n-2}(Z_{n-2}, Z_{n-1})\middle|\begin{matrix} X_{n+1} = \eta \\ X_n = \xi \end{matrix}\right)\right],$$
$$a_n^{(N)}(\xi) = \mathbb{E}\left(F_n(\underline{L}_n)\middle|\begin{matrix} X_{n+1} = \eta \\ X_n = \xi \end{matrix}\right) \text{ and } a_{n+1}^{(N)}(\eta) = \mathbb{E}\left(F_{n+1}(\underline{L}_{n+1})\middle|\begin{matrix} X_{n+1} = \eta \\ X_n = \xi \end{matrix}\right).$$

Proof of the Claim. We use Lemma 2.22. The identity for $c_n^{(N)}$ is because $\{Z_n\}$ is independent from $\{X_n\}$. The identity for $a_n^{(N)}$ is because conditioned on X_n, $\underline{L}_n$ is independent of X_{n+1}. The identity for $a_{n+1}^{(N)}$ is because conditioned on X_{n+1}, $\underline{L}_{n+1}$ is independent of X_n.

With the claim proved, we can proceed to bound $\widetilde{\mathsf{f}}$. Taking the conditional expectation $\mathbb{E}(\,\cdot\,|X_{n+1}^{(N)} = \eta\,,\ X_n^{(N)} = \xi)$ on both sides of (3.5), we find that

$$f_n^{(N)}(\xi, \eta) = a_{n+1}(\eta) - a_n(\xi) + c_n - \mathbb{E}\left(\Gamma_n(\underline{L}_n, \underline{L}_{n+1})\middle|\begin{matrix} X_{n+1} = \eta \\ X_n = \xi \end{matrix}\right),$$

whence $\widetilde{f}_n(\xi, \eta) := -\mathbb{E}\left(\Gamma_n(\underline{L}_n, \underline{L}_{n+1})\middle|\begin{matrix} X_{n+1} = \eta \\ X_n = \xi \end{matrix}\right)$.

Clearly $|\widetilde{\mathsf{f}}| \le 6K$. To bound the L^2 norm we recall that the marginal distribution of $\{X_n\}$ with respect to the distribution of the ladder process is precisely the distribution of our original array. Therefore,

$$\|\widetilde{f}_n^{(N)}\|_2^2 = \mathbb{E}[\mathbb{E}(\Gamma_n(\underline{L}_n, \underline{L}_{n+1})|X_{n+1}, X_n)^2] \overset{!}{\le} \mathbb{E}[\Gamma_n(\underline{L}_n, \underline{L}_{n+1})^2],$$

where $\overset{!}{\le}$ is because conditional expectations contract L^2-norms. By Lemma 2.22(d), $\Gamma_n^{(N)}(\underline{L}_n, \underline{L}_{n+1})$ is equal in distribution to the balance of a random level N hexagon at position n, whence $\|\widetilde{f}_n^{(N)}\|_2^2 \le \mathbb{E}(\Gamma_n^2) = (u_n^{(N)})^2$. $\square$

3.2.2 The Estimate of $\mathrm{Var}(S_N)$

We prove Theorem 3.6. Let X and f be as in assumptions (E) and (B), in particular $|\mathsf{f}| \leq K$ a.s., and the row lengths are $k_N + 1$. Our aim is to bound $\mathrm{Var}(S_N)$ above and below by affine functions of the structure constants U_N. Henceforth, we fix N, and drop the superscripts $^{(N)}$. So $X_n^{(N)} = X_n$, $f_n^{(N)} = f_n$, $u_n^{(N)} = u_n$ etc.

Preparatory Estimate. *Let* $\Phi_t := f_t(X_t, X_{t+1}) + f_{t+1}(X_{t+1}, X_{t+2}) + f_{t+2}(X_{t+2}, X_{t+3})$. *There is a positive constant* C_0 *independent of* N *such that for every* $1 \leq t \leq k_N - 1$,

$$\mathbb{E}\left[\mathrm{Var}\left(\Phi_t | X_t, X_{t+3}\right)\right] \geq C_0 u_{t+2}^2. \tag{3.8}$$

Proof. By uniform ellipticity, $\pi_{n,n+1}(x_n, \mathrm{d}y) = p_n(x_n, x_{n+1})\mu_{n+1}(\mathrm{d}x_{n+1})$, with $p_n(\cdot,\cdot)$ as in the uniform ellipticity condition, with ellipticity constant ϵ_0.

By Corollary 2.9, we may take $\mu_n = P_n$ for $n \geq 3$, where $P_n(E) := \mathbb{P}(X_n \in E)$. Henceforth integration variables ranging over subsets of $\mathfrak{S}_k$ will be denoted by x_k or x'_k, and we will use the following short-hand for integrals and densities:

$$p(x_k, \ldots, x_{k+\ell}) := \prod_{j=k}^{k+\ell-1} p_j(x_j, x_{j+1}), \quad \int \varphi(x_k)\mathrm{d}x_k := \begin{cases} \int_{\mathfrak{S}_k} \varphi(x_k)\mu_k(\mathrm{d}x_k) & k > t \\ \int_{\mathfrak{S}_t} \varphi(x_t)P_t(\mathrm{d}x_t) & k = t. \end{cases}$$

The joint distribution of $(X_t, \ldots, X_{t+3})$ is $p(x_t, \ldots, x_{t+3})\mathrm{d}x_t \cdots \mathrm{d}x_{t+3}$. Therefore:

$$\begin{aligned}
&\mathbb{E}[\mathrm{Var}(\Phi_t|X_t, X_{t+3})] = \iiiint \mathrm{Var}\left(\Phi_t \middle| \begin{matrix} X_t = x_t \\ X_{t+3} = x_{t+3} \end{matrix}\right) p(x_t, \ldots, x_{t+3})\mathrm{d}x_t \cdots \mathrm{d}x_{t+3} \\
&= \iint \mathrm{d}x_t \mathrm{d}x_{t+3} \Big[\mathrm{Var}(\Phi_t|X_t = x_t, X_{t+3} = x_{t+3}) \int p(x_t, x_{t+1}) \left(\int p(x_{t+1}, x_{t+2})p(x_{t+2}, x_{t+3})\mathrm{d}x_{t+2}\right) \mathrm{d}x_{t+1}\Big] \\
&\geq \epsilon_0 \iint \mathrm{d}x_t \mathrm{d}x_{t+3} \Big[\mathrm{Var}(\Phi_t|X_t = x_t, X_{t+3} = x_{t+3})\Big], \quad \text{by uniform ellipticity.}
\end{aligned}$$

To continue, we need the following two facts. Firstly, the distribution of $(X_t, X_{t+1}, X_{t+2}, X_{t+3})$ conditioned on $X_t = x_t, X_{t+3} = x_{t+3}$ is

$$\nu_{x_t, x_{t+3}} := \delta_{x_t} \times \frac{p(x_t, x_{t+1}, x_{t+2}, x_{t+3})\mathrm{d}x_{t+1}\mathrm{d}x_{t+2}}{p_t(x_t \to x_{t+3})} \times \delta_{x_{t+3}}$$

(see (2.18)). Secondly, for any two identically distributed independent random variables W, W', $\mathrm{Var}(W) = \frac{1}{2}\mathbb{E}[(W - W')^2]$. It follows that

$$\begin{aligned}
&\mathbb{E}[\mathrm{Var}(\Phi_t|X_t, X_{t+3})] \\
&\geq \epsilon_0 \iint \mathrm{d}x_t \mathrm{d}x_{t+3} \iiiint \frac{p(x_t, x_{t+1}, x_{t+2}, x_{t+3})p(x_t, x'_{t+1}, x'_{t+2}, x_{t+3})}{p_t(x_t \to x_{t+3})^2} \times \\
&\times \frac{1}{2} \begin{bmatrix} f_t(x_t, x_{t+1}) + f_{t+1}(x_{t+1}, x_{t+2}) + f_{t+2}(x_{t+2}, x_{t+3}) \\ -f_t(x_t, x'_{t+1}) - f_{t+1}(x'_{t+1}, x'_{t+2}) - f_{t+2}(x'_{t+2}, x_{t+3}) \end{bmatrix}^2 \mathrm{d}x_{t+1}\mathrm{d}x'_{t+1}\mathrm{d}x_{t+2}\mathrm{d}x'_{t+2}.
\end{aligned}$$

- $p_t(x_t \to x_{t+3})$ are bounded above by ϵ_0^{-3}, see (2.18) and recall that $p(x, y) \leq \epsilon_0^{-1}$.
- The expression in the square brackets is the balance $\Gamma\left(x_t \begin{matrix} x_{t+1} & x_{t+2} \\ x'_{t+1} & x'_{t+2} \end{matrix} x_{t+3}\right)$.
- The density in front of the square brackets is the density of $\mathrm{d}m'_{\mathrm{Hex}}$ from (2.23).
- $\epsilon_0^8 \leq \frac{\mathrm{d}m'_{\mathrm{Hex}}}{\mathrm{d}m_{\mathrm{Hex}}} \leq \epsilon_0^{-8}$, by (2.24).

Thus

$$\mathbb{E}[\mathrm{Var}(\Phi_t|X_t, X_{t+3})] \ge \frac{1}{2}\epsilon_0^9 \int_{\mathrm{Hex}\,(N,t+2)} \Gamma^2 \mathrm{d}m_{\mathrm{Hex}} = \frac{1}{2}\epsilon_0^9 u_{t+2}^2.$$

Lower Bound for the Variance. Let us split $U_N = \sum_{n=3}^{k_N} u_n^2$ into three sums:

$$U_N = \sum_{\gamma=0,1,2} U_N(\gamma), \text{ where } U_N(\gamma) := \sum_{n=3}^{k_N} u_n^2 1_{[n=\gamma \bmod 3]}(n).$$

For every N there is at least one $\gamma_N \in \{0,1,2\}$ such that $U_N(\gamma_N) \ge \frac{1}{3}U_N$. Let $\alpha_N := \gamma_N + 1$, and define β_N and M_N by

$$k_N - \beta_N + 1 = \max\{n \le k_N : n = \alpha_N \bmod 3\} = 3M_N + \alpha_N.$$

With these choices, $\alpha_N, \beta_N \in \{1,2,3\}$, and $M_n \in \mathbb{N} \cup \{0\}$.

We begin by bounding the variance of $S'_N := \sum_{k=\alpha_N}^{k_N-\beta_N} f_j(X_j, X_{j+1})$ from below. Observe that $S'_N = F_0 + \cdots + F_{M_N-1}$, where

$$F_k(\xi_1,\xi_2,\xi_3,\xi_4) := f_{3k+\alpha_N}(\xi_1,\xi_2) + f_{3k+\alpha_N+1}(\xi_2,\xi_3) + f_{3k+\alpha_N+2}(\xi_3,\xi_4).$$

S'_N is a function of the following variables:

$$\boxed{X_{\alpha_N}}, X_{\alpha_N+1}, X_{\alpha_N+2}, \boxed{X_{\alpha_N+3}}, X_{\alpha_N+4}, X_{\alpha_N+5}, \cdots, \boxed{X_{k_N-\beta_N+1}},$$

where we have boxed the terms with indices congruent to α_N mod 3. Let $\mathscr{F}_N$ denote the σ-algebra generated by the boxed random variables. Conditioned on $\mathscr{F}_N$, F_k are independent. Therefore,

$$\mathrm{Var}(S'_N|\mathscr{F}_N) = \sum_{k=0}^{M_N-1} \mathrm{Var}(F_k|\mathscr{F}_N) = \sum_{k=0}^{M_N-1} \mathrm{Var}(F_k|X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N}).$$

By Jensen's inequality, $\mathrm{Var}(S'_N) \ge \mathbb{E}(\mathrm{Var}(S'_N|\mathscr{F}_N))$. It follows that

$$\mathrm{Var}(S'_N) \ge \sum_{k=0}^{M_N-1} \mathbb{E}\Big(\mathrm{Var}(F_k|X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N})\Big) \equiv \sum_{k=0}^{M_N-1} \mathbb{E}\Big(\mathrm{Var}(\Phi_{3k+\alpha_N}|X_{3k+\alpha_N}, X_{3(k+1)+\alpha_N})\Big)$$

$$\overset{(3.8)}{\ge} C_0 \sum_{k=0}^{M_N-1} u_{3k+\alpha_N+2}^2 = C_0 \sum_{k=0}^{M_N-1} u_{3(k+1)+\gamma_N}^2 \quad (\because \alpha_N = \gamma_N + 1)$$

$$\ge C_0 \sum_{n=3}^{k_N} u_n^2 1_{[n=\gamma_n \bmod 3]}(n) - 4C_0 \sup\{u_j^2\} \ge C_0 U_N(\gamma_N) - 4C_0 \cdot (6K)^2 \ge \frac{1}{3}C_0 U_N - 200 C_0 K^2, \text{ by the choice of } \gamma_N.$$

Now we claim that $\mathrm{Var}(S_N) \ge \mathrm{Var}(S'_N) - \text{const}$. Let $A_N := \{j \in \mathbb{N} : 1 \le j < \alpha_N \text{ or } k_N - \beta_N < j \le k_N\}$. $S_N = S'_N + \sum_{j\in A_N} f_j$, therefore $\mathrm{Var}(S_N) = \mathrm{Var}(S'_N) + \mathrm{Var}\Big(\sum_{j\in A_N} f_j\Big) + 2\sum_{j\in A_N} \mathrm{Cov}(S'_N, f_j)$.

Since $|A_N| \le 6$, $\mathrm{Var}(\sum_{j\in A_N} f_j)$ is uniformly bounded by a constant only depending on K. By Lemma 2.14, $\sum_{j\in A_N} \mathrm{Cov}(S'_N, f_j)$ is also uniformly bounded by a constant depending only on K and ϵ_0.

It follows that $\mathrm{Var}(S_N) \ge \mathrm{Var}(S'_N) - \text{const.} \ge \text{const.} U_N - \text{const.}$, where the constants depends only on K and the ellipticity constant ϵ_0.

Upper Bound for the Variance. Write $\mathsf{f} = \widetilde{\mathsf{f}} + \nabla \mathsf{a} + \mathsf{c}$ as in the gradient lemma. In particular, $\mathrm{Var}(\widetilde{f}_n(X_{n-1}, X_n)) \le u_n^2$. Then $\mathrm{Var}\left(\sum_{n=1}^{k_N} f_n\right) = \mathrm{Var}\left(\sum_{n=1}^{k_N} \widetilde{f}_n\right) + \mathrm{Var}\,(a_{k_N+1} - a_1) + 2\mathrm{Cov}\left(\sum_{n=1}^{k_N} \widetilde{f}_n, a_{k_N+1} - a_1\right)$. The first term is smaller than const.U_N + const. due to the gradient lemma and Lemma 3.4. The second term is bounded since $|\mathsf{a}| \leq 2K$. The third term is smaller than a constant, due to Lemma 2.14. Moreover, looking at these lemmas, we see that the constants only depend on K and ϵ_0. □

The Case of One-Step Ellipticity. To prove that the one-step ellipticity condition implies (3.1), we use quadrilaterals $Q_n^{(N)} = \left(X_{n-1}^{(N)} \begin{smallmatrix} X_n^{(N)} \\ Y_n^{(N)} \end{smallmatrix} Y_{n+1}^{(N)} \right)$ instead of hexagons. The corresponding structure constants can then be shown to satisfy

$$(\overline{u}_n^{(N)})^2 \asymp \iint |f_n^{(N)}(y_1) - f_n^{(N)}(y_2)|^2 \mu_n^{(N)}(\mathrm{d}y_1)\mu_n^{(N)}(\mathrm{d}y_2) = 2\mathrm{Var}(f_n^{(N)}). \tag{3.9}$$

Then one proceeds as in the proof of Theorem 3.6 above.

3.2.3 McLeish's Martingale Central Limit Theorem

A **martingale difference array** with row lengths k_N is a (possibly non-Markov) array Δ of random variables $\Delta = \{\Delta_j^{(N)} : N \geq 1, 1 \leq j \leq k_N\}$ together with an array of σ-algebras $\{\mathscr{F}_j^{(N)} : N \geq 1, 1 \leq j \leq k_N\}$, so that:

(1) For each N, $\Delta_1^{(N)}, \ldots, \Delta_{k_N}^{(N)}$ are random variables on the same probability space $(\mathfrak{S}_N, \mathscr{F}_N, \mu_N)$.
(2) $\mathscr{F}_1^{(N)} \subset \mathscr{F}_2^{(N)} \subset \mathscr{F}_3^{(N)} \subset \cdots \subset \mathscr{F}_{k_N}^{(N)}$ are sub σ-algebras of $\mathscr{F}_N$.
(3) $\Delta_j^{(N)}$ is $\mathscr{F}_j^{(N)}$–measurable, $\mathbb{E}(|\Delta_j^{(N)}|) < \infty$, and $\mathbb{E}(\Delta_{j+1}^{(N)}|\mathscr{F}_j^{(N)}) = 0$.

We say that Δ has **finite variance**, if every $\Delta_j^{(N)}$ has finite variance. Notice that $\mathbb{E}(\Delta_j^{(N)}) = 0$ for all $j = 2, \ldots, k_{N+1}$. If in addition $\mathbb{E}(\Delta_1^{(N)}) = 0$ for all N, then we say that Δ has **mean zero**.

A **martingale difference sequence** is a martingale difference array such that $\Delta_i^{(N)} = \Delta_i$ and $\mathscr{F}_i^{(N)} = \mathscr{F}_i$ for all N.

Example 3.15 Suppose $\{S_n\}$ is a martingale relative to $\{\mathscr{F}_n\}$, then $\Delta_1 := S_1$, $\Delta_j := S_j - S_{j-1}$ is a martingale difference sequence.

The following observation is the key to many of the properties of martingale difference arrays:

Lemma 3.16 *Suppose Δ is a martingale difference array with finite variance, then for each N, $\Delta_1^{(N)}, \ldots, \Delta_{k_N}^{(N)}$ are uncorrelated, and if Δ has mean zero, then* $\mathrm{Var}\left(\sum_{n=1}^{k_N} \Delta_n^{(N)}\right) = \sum_{n=1}^{k_N} \mathbb{E}\left[(\Delta_n^{(N)})^2\right]$.

Proof Fix N and write $\Delta_j^{(N)} = \Delta_j$, $\mathscr{F}_j^{(N)} = \mathscr{F}_j$. If $i < j$, then

$$\mathbb{E}(\Delta_j\Delta_i) = \mathbb{E}[\mathbb{E}(\Delta_j\Delta_i|\mathscr{F}_{j-1})] = \mathbb{E}[\mathbb{E}(\Delta_i\mathbb{E}(\Delta_j|\mathscr{F}_{j-1}))] = \mathbb{E}(\Delta_i \cdot 0) = 0.$$

The identity for the variance immediately follows. □

Theorem 3.17 (McLeish's Martingale Central Limit Theorem) *Let $\Delta = \{\Delta_j^{(N)}\}$ be a martingale difference array with row lengths k_N, zero mean, and finite variance, and let $V_N := \sum_{j=1}^{k_N} \mathbb{E}[(\Delta_j^{(N)})^2]$. Suppose:*

(1) $\max\limits_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}}$ *has uniformly bounded L^2 norm;*
(2) $\max\limits_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}} \xrightarrow[N\to\infty]{} 0$ *in probability; and*

(3) $\displaystyle\frac{1}{V_N}\sum_{n=1}^{k_N}(\Delta_n^{(N)})^2 \xrightarrow[N\to\infty]{} 1$ *in probability.*

Then for all intervals (a,b)*,* $\displaystyle\mathbb{P}\left[\frac{1}{\sqrt{V_N}}\sum_{j=1}^{k_N}\Delta_j^{(N)} \in (a,b)\right] \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt.$

We make some preparations for the proof.

A sequence of random variables $\{Y_n\}$ on $(\Omega,\mathscr{F},\mu)$ is called **uniformly integrable** if for every ϵ, $\exists K$ such that $\mathbb{E}(|Y_n|1_{[|Y_n|>K]}) < \epsilon$ for all n. This is strictly stronger than tightness (there are tight sequences of non-integrable random variables).

Example 3.18 (Bounded Moments and Uniform Integrability) If $\sup\|Y_n\|_p < \infty$ for some $p > 1$, then $\{Y_n\}$ is uniformly integrable. To see this, let $M_p := \sup\|Y_n\|_p$, and suppose $\frac{1}{p}+\frac{1}{q} = 1$. By Markov's inequality,

$$\mu[|Y_n| > K] \le \frac{1}{K^p}M_p^p,$$

and by Hölder's inequality, $\mathbb{E}(|Y_n|1_{[|Y_n|>K]}) \le M_p\mu[|Y_n| > K]^{1/q} = O(K^{-p/q})$.

Lemma 3.19 *Suppose* $Y_n, Y \in L^1(\Omega,\mathscr{F},\mu)$*, then* $Y_n \xrightarrow[n\to\infty]{L^1} Y$ *iff* $\{Y_n\}$ *are uniformly integrable, and* $Y_n \xrightarrow[n\to\infty]{} Y$ *in probability. In this case,* $\mathbb{E}(Y_n) \xrightarrow[n\to\infty]{} \mathbb{E}(Y)$*.*

Proof The proof is standard, but we include it for completeness.

Proof of ($\Rightarrow$): Suppose $\|Y_n - Y\|_1 \to 0$, then it is easy to see that $\mathbb{E}(Y_n) \to \mathbb{E}(Y)$, and that $Y_n \to Y$ in probability. It remains to check uniform integrability.

Since $Y \in L^1$, $\lim_{K\to\infty}\mathbb{E}(|Y|1_{[|Y|\ge K]}) = 0$. Given ε take K so that $\mathbb{E}(|Y|1_{[|Y|\ge K]}) < \varepsilon$. Next choose $\delta > 0$ so small that

$$K\delta + \mathbb{E}(|Y|1_{[|Y|\ge K]}) < \varepsilon.$$

For this δ, $\mathbb{E}(|Y|1_F) < \varepsilon$ for all measurable sets F such that $\mu(F) < \delta$.

By Markov's inequality, $\mathbb{P}[|Y_n| > L] \le L^{-1}\sup\|Y_n\|_1 = O(L^{-1})$, so there exists $L > K$ such that $\mathbb{P}[|Y_n| > L] < \delta$ for all n. By the choice of δ,

$$\int_{[|Y_n|>L]}|Y_n|\,d\mu \le \int_{[|Y_n|>L]}|Y|\,d\mu + \int_{[|Y_n|>L]}|Y_n - Y|\,d\mu < \varepsilon + \|Y_n - Y\|_1.$$

Since $\|Y_n - Y\|_1 \to 0$, there exists an N so that $\mathbb{E}(|Y_n|1_{[|Y_n|>L]}) < \varepsilon$ for all $n \ge N$.

Since $Y_n \in L^1$, for some $M > L$ big enough, $\mathbb{E}(|Y_n|1_{[|Y_n|>M]}) < \varepsilon$ for all $1 \le n \le N-1$. So $\mathbb{E}(|Y_n|1_{[|Y_n|>M]}) < \varepsilon$ for all n.

Proof of ($\Leftarrow$): Given a random variable Z, let $Z^K := Z1_{[|Z|\le K]}$. Since $\{Y_n\}$ is uniformly integrable, for every ϵ there is a $K > 1$ such that $\|Y_n^K - Y_n\|_1 < \epsilon$ for all n. By the dominated convergence theorem, $\|Y^K - Y\|_1 < \epsilon$ for all K large enough. Thus for all K large enough, for all n,

$$\begin{aligned}\|Y_n - Y\|_1 &\le \|Y_n^K - Y^K\|_1 + 2\epsilon \le \epsilon\mu[|Y_n^K - Y^K| \le \epsilon] + 2K\mu[|Y_n^K - Y^K| > \epsilon] + 2\epsilon\\ &\le 3\epsilon + 2K\Big(\mu[|Y_n - Y| > \epsilon] + \mu[|Y_n| > K] + \mu[|Y| > K]\Big)\\ &\le 3\epsilon + 2K\mu[|Y_n - Y| > \epsilon] + 2\mathbb{E}(|Y_n|1_{[|Y_n|>K]}) + 2\mathbb{E}(|Y|1_{|Y|>K}).\end{aligned}$$

Using the assumption that $Y_n \to Y$ in probability, we obtain

$$\limsup_{n\to\infty}\|Y_n - Y\|_1 \le 3\epsilon + 2\sup_n\mathbb{E}(|Y_n|1_{[|Y_n|>K]}) + 2\mathbb{E}(|Y|1_{|Y|>K}) \overset{!}{<} 7\varepsilon,$$

where the last inequality follows from the choice of K. Now take $\varepsilon \to 0$. □

Lemma 3.20 (McLeish) *Let $\{W_j^{(N)} : 1 \le j \le k_N, N \ge 1\}$ be a triangular array of random variables,*[2] *where $W_1^{(N)}, \dots, W_{k_N}^{(N)}$ are defined on the same probability space. Fix $t \in \mathbb{R}$ and let $T_N(t) := \prod_{j=1}^{k_N} \left(1 + \mathrm{i}tW_j^{(N)}\right)$. Suppose*

(1) *$\{T_N(t)\}$ is uniformly integrable and $\mathbb{E}(T_N) \xrightarrow[N\to\infty]{} 1$;*
(2) *$\sum_{j=1}^{k_N} (W_j^{(N)})^2 \xrightarrow[N\to\infty]{} 1$ in probability;*
(3) *$\max_{1\le j\le k_N} |W_j^{(N)}| \xrightarrow[N\to\infty]{} 0$ in probability.*

Then $\mathbb{E}(\mathrm{e}^{\mathrm{i}t(W_1^{(N)}+\cdots+W_{k_N}^{(N)})}) \xrightarrow[N\to\infty]{} \mathrm{e}^{-\frac{1}{2}t^2}$.

Proof Define a function $r(x)$ on $[-1,1]$ by the identity $\mathrm{e}^{\mathrm{i}x} = (1+\mathrm{i}x)\mathrm{e}^{-\frac{1}{2}x^2+r(x)}$. Equivalently, $r(x) = -\log(1+\mathrm{i}x) + \mathrm{i}x + \frac{1}{2}x^2 = O(|x|^3)$. Fix C such that $|r(x)| \le C|x|^3$ for $|x| < 1$.

Substituting $S_N := W_1^{(N)} + \cdots + W_{k_N}^{(N)}$ in $\mathrm{e}^{\mathrm{i}x} = (1+\mathrm{i}x)\mathrm{e}^{-\frac{1}{2}x^2+r(x)}$ gives (in what follows we drop the superscripts $^{(N)}$ and abbreviate $T_n := T_n(t)$):

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}tS_N}) = \mathbb{E}\Big(\prod_{j=1}^{k_N} \mathrm{e}^{\mathrm{i}tW_j}\Big) = \mathbb{E}(T_N \mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{k_N} t^2 W_j^2 + r(tW_j)}) = \mathbb{E}(T_N U_N), \text{ where}$$

$$U_N := \exp\Big[-\frac{t^2}{2}\sum_{j=1}^{k_N} \left(W_j^{(N)}\right)^2 + r(tW_j^{(N)})\Big].$$

T_N and U_N have the following properties:

(a) $\mathbb{E}(T_N) \xrightarrow[N\to\infty]{} 1$, by assumption.
(b) $\{T_N\}$ is uniformly integrable by assumption, and $|T_N U_N| = |\mathrm{e}^{\mathrm{i}tS_N}| = 1$.
(c) $U_N \xrightarrow[N\to\infty]{} \mathrm{e}^{-\frac{1}{2}t^2}$ in probability, because $\sum_{j=1}^{k_N} \left(W_j^{(N)}\right)^2 \xrightarrow[N\to\infty]{} 1$ in probability, and by the assumptions of the lemma, with asymptotic probability one, $\left|\sum_{i=1}^{k_N} r(tW_j^{(N)})\right| \le C|t|^3 \left(\max_{1\le j\le k_N} \left|W_j^{(N)}\right|\right) \sum_{j=1}^{k_N} \left(W_j^{(N)}\right)^2 \xrightarrow[N\to\infty]{\text{prob}} 0$. □

Let $L := \mathrm{e}^{-\frac{1}{2}t^2}$, then

$$\begin{aligned} |\mathbb{E}(\mathrm{e}^{\mathrm{i}tS_N}) - L| = |\mathbb{E}(T_N U_N) - L| &\le |\mathbb{E}(T_N(U_N - L))| \quad + L|\mathbb{E}(T_N) - 1| \\ &= |\mathbb{E}(T_N(U_N - L))| + o(1), \text{by (a)}. \end{aligned} \tag{3.10}$$

Next, $\mathbb{P}[|T_N(U_N - L)| > \epsilon] \le \mathbb{P}[|T_N| > K] + \mathbb{P}[|U_N - L| > \epsilon/K]$, for all K and ε. Therefore by (b) and (c), $T_N(U_N - L) \xrightarrow[N\to\infty]{} 0$ in probability. Finally, by (b), $|T_N(U_N - L)| \le 1 + L|T_N|$, and $T_N(U_N - L)$ is uniformly integrable. By Lemma 3.19, $\mathbb{E}(T_N(U_N - L)) \to 0$, and by (3.10), $\mathbb{E}(\mathrm{e}^{\mathrm{i}tS_N}) \to \mathrm{e}^{-\frac{1}{2}t^2}$. □

Proof of the Martingale CLT ([140]): Let $\Delta = \{\Delta_j^{(N)}\}$ be a martingale difference array with row lengths k_N, which satisfies the assumptions of Theorem 3.17, and let

[2] Not necessarily a martingale difference array or a Markov array.

$$S_N := \sum_{j=1}^{k_N} \Delta_j^{(N)} \text{ and } V_N := \mathrm{Var}(S_N) \equiv \sum_{j=1}^{k_N} \mathbb{E}\left[(\Delta_j^{(N)})^2\right] \quad \text{(see Lemma 3.16).}$$

It is tempting to apply McLeish's lemma to the normalized array $\Delta_j^{(N)}/\sqrt{V_N}$, but to do this we need to check the uniform integrability of $\prod_{j=1}^n (1 + \mathrm{i}t\Delta_j^{(N)}/\sqrt{V_N})$ and this is difficult. It is easier to work with the following array of truncations:

$$W_1^{(N)} := \tfrac{1}{\sqrt{V_N}}\Delta_1^{(N)}\ ,\ W_n^{(N)} := \tfrac{1}{\sqrt{V_N}}\Delta_n^{(N)} 1_{[\sum_{k=1}^{n-1}(\Delta_k^{(N)})^2 \le 2V_N]}.$$

It is easy to check that $\{W_n^{(N)}\}$ is a martingale difference array relative to $\mathscr{F}_n^{(N)}$, and that $\{W_n^{(N)}\}$ has zero mean, and finite variance. In addition, $S_N^* := \sum_{n=1}^{k_N} W_n^{(N)}$ are close to $S_N/\sqrt{V_N}$ in probability:

$$\mathbb{P}\left[S_N^* \ne \tfrac{S_N}{\sqrt{V_N}}\right] \le \mathbb{P}\left[\exists 1 \le j \le k_N \text{ s.t. } \sum_{k=1}^{j-1}(\Delta_k^{(N)})^2 > 2V_N\right]$$
$$\le \mathbb{P}\Big[\sum_{j=1}^{k_N}(\Delta_k^{(N)})^2 > 2V_N\Big] \xrightarrow[N\to\infty]{} 0,\ \because\ \frac{1}{V_N}\sum_{j=1}^{k_N}\left(\Delta_j^{(N)}\right)^2 \xrightarrow[N\to\infty]{\text{prob}} 1 \text{ by assumption.}$$

Thus to prove the theorem, it is enough to show that S_N^* converges in distribution to the standard Gaussian distribution. To do this, we check that $\{W_n^{(N)}\}$ satisfies the conditions of McLeish's lemma.

Fix $t \in \mathbb{R}$, and let $T_N = T_N(t) := \prod_{j=1}^{k_N}(1 + \mathrm{i}tW_j^{(N)})$, and $J_N := \max\{2 \le j \le k_N : \sum_{k=1}^{j-1}(\Delta_n^{(N)})^2 \le 2V_N\}$. (or $J_N = 1$ if the maximum is over the empty set). Writing $W_j = W_j^{(N)}$ and $\Delta_j = \Delta_j^{(N)}$, we obtain

$|T_N| = \prod_{j=1}^{k_N}(1 + t^2W_j^2)^{1/2} = \prod_{j=1}^{J_N}\left(1 + \dfrac{t^2\Delta_j^2}{V_N}\right)^{1/2}$. Thus

$$|T_N| = \left(\prod_{j=1}^{J_N-1}\left(1 + \frac{t^2\Delta_j^2}{V_N}\right)\right)^{1/2} \cdot \left(1 + \frac{t^2\Delta_{J_N}^2}{V_N}\right)^{1/2} \le \exp\Big(\frac{t^2}{2V_N}\sum_{j=0}^{J_N-1}\Delta_j^2\Big)\Big(1 + \frac{t^2\Delta_{J_N}^2}{V_N}\Big)^{1/2} \le \mathrm{e}^{t^2}\Big(1 + |t| \max_{1\le j\le k_N}\Big|\frac{\Delta_j^{(N)}}{\sqrt{V_N}}\Big|\Big).$$

So $\|T_N(t)\|_2^2 \le \mathrm{e}^{2t^2}\mathbb{E}\left[\left(1 + |t| \max_{1\le j\le k_N}\left|\Delta_j^{(N)}/\sqrt{V_N}\right|\right)^2\right]$.

By the first assumption of the theorem, $\sup_{N\in\mathbb{N}}\|T_N(t)\|_2 < \infty$ for each t. Thus $\{T_N(t)\}_{N\ge1}$ is uniformly integrable for each t (see Example 3.18). Next, successive conditioning shows that $\mathbb{E}(T_N) = 1 + \mathrm{i}t\mathbb{E}\left(\Delta_1^{(N)}\right) = 1$. The first condition of McLeish's lemma is verified.

Next, $\mathbb{P}\left[\sum_{n=1}^{k_N}(W_n^{(N)})^2 \ne \sum_{n=1}^{k_N}(\frac{\Delta_n^{(N)}}{\sqrt{V_N}})^2\right] \le \mathbb{P}\Big[\exists 1 \le n \le k_N \text{ s.t. } \sum_{j=1}^{n}(\Delta_j^{(N)})^2 > 2V_N\Big] \le \mathbb{P}\left[\sum_{n=1}^{k_N}(\Delta_n^{(N)})^2 > 2V_N\right] \xrightarrow[N\to\infty]{}$ 0, because $\dfrac{1}{V_n}\sum_{j=1}^{k_N}(\Delta_n^{(N)})^2 \to 1$ in probability.

The second condition of McLeish's lemma now follows from assumption 3 of the theorem. The third condition of McLeish's lemma follows for similar reasons.

So McLeish's lemma applies to $\{W_n^{(N)}\}$, and $\mathbb{E}(\mathrm{e}^{\mathrm{i}tS_N^*}) \to \mathrm{e}^{-\frac{1}{2}t^2}$ for all $t \in \mathbb{R}$. By Lévy's continuity theorem, S_N^* converges in distribution to the standard Gaussian distribution. As explained above, this implies that $\frac{S_N}{\sqrt{V_N}}$ converges in distribution to the standard Gaussian distribution. □

3.2.4 Proof of the Central Limit Theorem

We prove Theorem 3.10.

Let $\mathsf{X} = \{X_n^{(N)}\}$ be a uniformly elliptic Markov array with row lengths $k_N + 1$, and let $\mathsf{f} = \{f_n^{(N)}\}$ be an a.s. uniformly bounded additive functional on X such that $V_N := \mathrm{Var}(S_N) \to \infty$.

Without loss of generality, $\mathbb{E}[f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})] = 0$ and $|f_n^{(N)}| \leq K$ for all n, N. Let $\mathscr{F}_n^{(N)}$ denote the σ-algebra generated by $X_1^{(N)}, \ldots, X_{n+1}^{(N)}$, and let $\mathscr{F}_0^{(N)}$ denote the trivial σ-algebra.

Fix N and write $f_k = f_k^{(N)}(X_k^{(N)}, X_{k+1}^{(N)})$ and $\mathscr{F}_k = \mathscr{F}_k^{(N)}$, then $\mathbb{E}(f_k|\mathscr{F}_k) = f_k$, $\mathbb{E}(f_k|\mathscr{F}_0) = \mathbb{E}(f_k) = 0$, and therefore

$$S_N = \sum_{k=1}^{k_N} f_k = \sum_{k=1}^{k_N} (\mathbb{E}(f_k|\mathscr{F}_k) - \mathbb{E}(f_k|\mathscr{F}_0)) = \sum_{k=1}^{k_N} \sum_{n=1}^{k} (\mathbb{E}(f_k|\mathscr{F}_n) - \mathbb{E}(f_k|\mathscr{F}_{n-1})).$$

So $S_N = \sum_{n=1}^{k_N} \sum_{k=n}^{k_N} (\mathbb{E}(f_k|\mathscr{F}_n) - \mathbb{E}(f_k|\mathscr{F}_{n-1})) = \sum_{n=1}^{k_N} \Delta_n^{(N)}$, where

$$\Delta_n^{(N)} := \sum_{k=n}^{k_N} (\mathbb{E}(f_k^{(N)}|\mathscr{F}_n^{(N)}) - \mathbb{E}(f_k^{(N)}|\mathscr{F}_{n-1}^{(N)})).$$

The array $\{\Delta_n^{(N)} : 1 \leq n \leq k_N; N \geq 1\}$ is a martingale difference array relative to the filtrations $\mathscr{F}_n^{(N)}$, with zero mean and finite variance. To prove the theorem, it suffices to check that $\{\Delta_n^{(N)}\}$ satisfies the conditions of the martingale CLT.

Step 1: $\max\limits_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}}$ *has uniformly bounded* L^2 *norm, and* $\max\limits_{1 \leq j \leq k_N} \frac{|\Delta_j^{(N)}|}{\sqrt{V_N}} \xrightarrow[N\to\infty]{\text{prob}} 0$.

Proof. The proof is based on the exponential mixing of uniformly elliptic Markov arrays (Proposition 2.13): Let $K :=$ ess sup $|\mathsf{f}|$, then there are constants $C_{mix} > 1$ and $0 < \theta < 1$ such that for all $k \geq n$, $\|\mathbb{E}(f_k^{(N)}|\mathscr{F}_n^{(N)})\|_\infty \leq C_{mix} K \theta^{k-n-1}$.

It follows that $|\Delta_j^{(N)}| < 2C_{mix} K \sum\limits_{\ell=-1}^{\infty} \theta^\ell = \dfrac{2C_{mix} K \theta^{-1}}{1-\theta}$. The step follows, since $V_N \to \infty$, by the assumptions of the theorem.

Step 2: $\dfrac{1}{V_N} \sum\limits_{n=1}^{k_N} \left(\Delta_n^{(N)}\right)^2 \xrightarrow[N\to\infty]{} 1$ *in probability.*

Proof. We follow [181] closely. Let $Y_i^{(N)} := (\Delta_i^{(N)})^2 / V_N$. We will show that $\left\|\sum_{i=1}^{k_N} Y_i^{(N)} - 1\right\|_2^2 \xrightarrow[N\to\infty]{} 0$, and use the general fact that L^2-convergence implies convergence in probability (by Chebyshev's inequality).

Notice that $\mathbb{E}\Big(\sum\limits_{i=1}^{k_N} Y_i^{(N)}\Big) = 1$, because by Lemma 3.16, this expectation equals $\frac{1}{V_N} \times \mathrm{Var}\left(\sum_{n=1}^{k_N} \Delta_n^{(N)}\right) = \frac{1}{V_N} \mathrm{Var}(S_N) = 1$. So

$$\left\|\sum_{i=1}^{k_N} Y_i^{(N)} - 1\right\|_2^2 = \mathbb{E}\left[\left(\sum_{i=1}^{k_N} Y_i^{(N)}\right)^2\right] - 2\mathbb{E}\Big[\sum_{i=1}^{k_N} Y_i^{(N)}\Big] + 1 = \mathbb{E}\Big[\sum_{i=1}^{k_N} (Y_i^{(N)})^2\Big] + 2\mathbb{E}\Big[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}\Big] - 2 + 1$$

$$= O\left(\max_{1 \leq \ell \leq k_N} \|Y_\ell^{(N)}\|_\infty\right) \cdot \mathbb{E}\Big[\sum_{\ell=1}^{k_N} Y_i^{(N)}\Big] + 2\mathbb{E}\Big[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}\Big] - 1.$$

We saw in the proof of Step 1 that $\|\Delta_j^{(N)}\|_\infty$ are uniformly bounded. Thus

$$\max_{1 \leq \ell \leq k_N} \|Y_\ell^{(N)}\|_\infty = O(1/V_N). \tag{3.11}$$

So $\Big\|\sum_{i=1}^{k_N} Y_i^{(N)} - 1\Big\|_2^2 = 2\mathbb{E}[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}] - 1 + o(1)$. It remains to show that

$$2\mathbb{E}\Big[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}\Big] \xrightarrow[N\to\infty]{} 1. \tag{3.12}$$

Define the oscillation of an L^1-element φ to be the infimum of the oscillations of all its a.e. versions, see (2.8). The proof of (3.12) is based on the following claim:

$$\mathrm{Osc}(N) := \max_{1\le i\le k_N} \mathrm{Osc}\left(\mathbb{E}\Big(\sum_{j=i+1}^{k_N} Y_j^{(N)}\Big|\mathscr{F}_i^{(N)}\Big)\right) \xrightarrow[N\to\infty]{} 0. \tag{3.13}$$

Before proving this, we explain why (3.13) implies (3.12).

Henceforth we fix N and drop some of the superscripts $^{(N)}$. We start from

$$2\mathbb{E}\Big[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}\Big] = 2\mathbb{E}\Big[\sum_{i=1}^{k_N} Y_i \sum_{j=i+1}^{k_N} Y_j\Big] = 2\mathbb{E}\Big[\sum_{i=1}^{k_N} Y_i \mathbb{E}\Big(\sum_{j=i+1}^{k_N} Y_j|\mathscr{F}_i\Big)\Big].$$

Call the conditional expectation φ, then $\varphi = \mathbb{E}(\varphi) \pm \mathrm{Osc}(N)$ a.e., where $x = y \pm \epsilon$ means that $y - \epsilon \le x \le y + \epsilon$. Therefore,

$$\begin{aligned}
2\mathbb{E}\Big[\sum_{i<j} Y_i^{(N)} Y_j^{(N)}\Big] &= 2\mathbb{E}\Big[\sum_{i=1}^{k_N} Y_i \mathbb{E}\Big(\sum_{j=i+1}^{k_N} Y_j\Big)\Big] \pm 2\mathbb{E}\Big[\sum_{i=1}^{k_N} Y_i\Big]\mathrm{Osc}(N) = 2\sum_{i=1}^{k_N}\mathbb{E}(Y_i)\sum_{j=i+1}^{k_N}\mathbb{E}(Y_j) \pm 2\mathrm{Osc}(N) \quad (\because \sum_{i=1}^{k_N}\mathbb{E}(Y_i) = 1)\\
&= \Big(\sum_{i=1}^{k_N}\mathbb{E}(Y_i)\Big)^2 - \sum_{i=1}^{k_N}\mathbb{E}(Y_i)^2 \pm 2\mathrm{Osc}(N) = 1 + O\Big(\max_{1\le i\le k_N}\|Y_i\|_\infty\Big) \pm 2\mathrm{Osc}(N), \because \sum \mathbb{E}(Y_i)^2 \le \underbrace{\sum \mathbb{E}(Y_i)}_{=1} \max\|Y_i\|_\infty\\
&= 1 + O(V_N^{-1}) + O(\mathrm{Osc}(N)), \text{ see (3.11)}.
\end{aligned}$$

So (3.13) implies (3.12), and with it the step. We turn to the proof of (3.13). First we note that a routine modification of the proof of Lemma 3.16 shows that for all $j, k > i$, $\mathbb{E}(\Delta_j\Delta_k|\mathscr{F}_i) = 0$. It follows that

$$\begin{aligned}
\mathbb{E}\Big(\sum_{j=i+1}^{k_N} Y_j\Big|\mathscr{F}_i\Big) &\equiv \frac{1}{V_N}\mathbb{E}\Big(\sum_{j=i+1}^{k_N}\Delta_j^2\Big|\mathscr{F}_i\Big) = \frac{1}{V_N}\mathbb{E}\Big(\big(\sum_{n=i+1}^{k_N}\Delta_n\big)^2\Big|\mathscr{F}_i\Big) = \frac{1}{V_N}\mathbb{E}\Big[\Big(\sum_{n=i+1}^{k_N}\sum_{k=n}^{k_N}[\mathbb{E}(f_k|\mathscr{F}_n) - \mathbb{E}(f_k|\mathscr{F}_{n-1})]\Big)^2\Big|\mathscr{F}_i\Big]\\
&= \frac{1}{V_N}\mathbb{E}\Big[\Big(\sum_{k=i+1}^{k_N}\sum_{n=i+1}^{k}\mathbb{E}(f_k|\mathscr{F}_n) - \mathbb{E}(f_k|\mathscr{F}_{n-1})\Big)^2\Big|\mathscr{F}_i\Big] = \frac{1}{V_N}\mathbb{E}\Big[\Big(\sum_{k=i+1}^{k_N}[f_k - \mathbb{E}(f_k|\mathscr{F}_i)]\Big)^2\Big|\mathscr{F}_i\Big]\\
&= \frac{1}{V_N}\sum_{k,\ell=i+1}^{k_N}\mathbb{E}\Big[(f_k - \mathbb{E}(f_k|\mathscr{F}_i))(f_\ell - \mathbb{E}(f_\ell|\mathscr{F}_i))\Big|\mathscr{F}_i\Big]\\
&= \frac{1}{V_N}\sum_{k,\ell=i+1}^{k_N}\mathbb{E}\Big[f_k f_\ell + \mathbb{E}(f_k|\mathscr{F}_i)\mathbb{E}(f_\ell|\mathscr{F}_i) - f_k\mathbb{E}(f_\ell|\mathscr{F}_i) - f_\ell\mathbb{E}(f_k|\mathscr{F}_i)\Big|\mathscr{F}_i\Big].\\
&= \frac{1}{V_N}\sum_{k,\ell=i+1}^{k_N}\left[\mathbb{E}(f_k f_\ell|\mathscr{F}_i) - \mathbb{E}(f_\ell|\mathscr{F}_i)\mathbb{E}(f_k|\mathscr{F}_i)\right].
\end{aligned} \tag{3.14}$$

Proposition 2.13(1) applied to $\{X_n^{(N)}\}_{i+1\le n\le k_N+1}$, shows that there exists $C_0 > 0$ and $0 < \theta < 1$ which only depend on ϵ_0 such that for all $k > i+1$, and for every bounded measurable function $u : \mathfrak{S}_k^{(N)} \times \mathfrak{S}_{k+1}^{(N)} \to \mathbb{R}$,

$$\mathrm{Osc}\left[\mathbb{E}(u(X_k^{(N)}, X_{k+1}^{(N)})|\mathscr{F}_i^{(N)})\right] \le C_0\theta^{k-i}\mathrm{Osc}(u).$$

This, (2.11), and the inequalities $|f_j| \le K$, $\mathrm{Osc}(u) \le 2\|u\|_\infty$ and

$$\mathrm{Osc}(uv) \le \|u\|_\infty \mathrm{Osc}(v) + \|v\|_\infty \mathrm{Osc}(u)$$

imply the existence of constants $C_1 > 0$ and $0 < \theta < 1$ such that for every $N \ge 1$ and $i+2 \le k \le \ell \le k_N$,

$$\mathrm{Osc}(\mathbb{E}(f_\ell|\mathscr{F}_i)\mathbb{E}(f_k|\mathscr{F}_i)) \le \mathrm{Osc}(\mathbb{E}(f_\ell|\mathscr{F}_i))\|\mathbb{E}(f_k|\mathscr{F}_i)\|_\infty + \|\mathbb{E}(f_\ell|\mathscr{F}_i)\|_\infty \mathrm{Osc}(\mathbb{E}(f_k|\mathscr{F}_i)) \le C_1\theta^{k-i}\theta^{\ell-i}, \text{ and}$$

$$\mathrm{Osc}\Big(\mathbb{E}[f_k f_\ell|\mathscr{F}_i]\Big) = \mathrm{Osc}\Big(\mathbb{E}[f_k\mathbb{E}(f_\ell|\mathscr{F}_k)|\mathscr{F}_i]\Big) \le C_0\theta^{k-i}\mathrm{Osc}(f_k\mathbb{E}(f_\ell|\mathscr{F}_k))$$
$$\le C_0\theta^{k-i}[K\cdot \mathrm{Osc}(\mathbb{E}(f_\ell|\mathscr{F}_k)) + \mathrm{Osc}(f_k)\|\mathbb{E}(f_\ell|\mathscr{F}_k)\|_\infty] \le C_1\theta^{k-i}\theta^{\ell-k}.$$

We have stated these bounds for $k, \ell \ge i+2$, but in fact they remain valid for $k = i+1$ or $\ell = i+1$, if we increase C_1 to guarantee that $C_1\theta^2 > 2K^2$.

Substituting these bounds in (3.14), we find that

$$\mathrm{Osc}(N) \le \max_{1\le i\le k_N} \frac{1}{V_N} \sum_{k,\ell=i+1}^{\infty} C_1\theta^{k-i}\theta^{|\ell-k|} + C_1\theta^{k-i}\theta^{\ell-i} = O(V_N^{-1}) \xrightarrow[N\to\infty]{} 0.$$

This proves (3.13), and Step 2.

Steps 1 and 2 verify the conditions of the martingale CLT. So $\frac{1}{\sqrt{V_N}}S_N \equiv \frac{1}{\sqrt{V_N}}\sum_{n=1}^{k_N}\Delta_n^{(N)}$ converges in distribution to the standard normal distribution. □

3.2.5 Convergence of Moments

We prove Theorem 3.11. It is sufficient to prove the following lemma:

Lemma 3.21 *Let* f *be a centered bounded additive functional of a uniformly elliptic Markov chain such that* $V_N \to \infty$. *Then for each* $r \in \mathbb{N}$ *there is a constant* C_r *such that for all* N,

$$\left|\mathbb{E}\left[S_N^r\right]\right| \le C_r V_N^{\lfloor r/2\rfloor}.$$

Proof of Theorem 3.11 Assuming Lemma 3.21: Suppose r is even. We have already remarked that by Dobrushin's CLT, for any *bounded* continuous function $\phi : \mathbb{R} \to \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{E}\left[\phi\left(\frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\phi(z)\mathrm{e}^{-z^2/2}dz. \tag{3.15}$$

Let $\mathbf{N}$ be a Gaussian random variable with mean zero and variance one. Applying (3.15) to the bounded continuous function $\phi_M(x) = x^r \wedge M$, we obtain

$$\lim_{N\to\infty}\mathbb{E}\left[\left(\tfrac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^r \wedge M\right] = \mathbb{E}\left(\mathbf{N}^r \wedge M\right).$$

By the dominated convergence theorem and the assumption that r is even,

$$\lim_{M\to\infty}\lim_{N\to\infty}\mathbb{E}\left[\left(\tfrac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^r \wedge M\right] = \lim_{M\to\infty}\mathbb{E}\left(\mathbf{N}^r \wedge M\right) = \mathbb{E}(\mathbf{N}^r) = (r-1)!!.$$

It remains to see that

$$\lim_{M\to\infty}\lim_{N\to\infty}\mathbb{E}\left[\left(\tfrac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}\right)^r\wedge M\right]=\lim_{N\to\infty}\mathbb{E}\left[\left(\tfrac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}\right)^r\right]. \tag{3.16}$$

By Lemma 3.21, $\left\|\left(\dfrac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}\right)^r\right\|_2^2\leq C_{2r}$ for all N. Therefore $(\frac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}})^r$ are uniformly integrable. It is not difficult to see that this implies (3.16). This proves the theorem for even r. The proof when r is odd is similar, except that one has to use $\phi_M(x)=(x^r\wedge M)\vee(-M)$, and the identity $\mathbb{E}(\mathbf{N}^r)=0$. □

The rest of the section is a proof of Lemma 3.21. By the gradient lemma (and the Cauchy-Schwarz inequality), it is sufficient to prove Lemma 3.21 under the additional assumption that there is some constant $C>0$ as follows:

$$\sum_{n=3}^{N}\widetilde{u}_n^2\leq CV_N,\ \text{ where } \widetilde{u}_n:=\|f_n\|_{L^2}.$$

Let $f_n:=f_n(X_n,X_{n+1})$. The proof proceeds by expanding S_N^r into a sum of r-tuples $f_{n_1}\cdots f_{n_r}$ $(n_1\leq\cdots\leq n_r)$, and by estimating the expectation of each tuple. Consider an r-tuple $f_{n_1}\cdots f_{n_r}$ where $n_1\leq n_2\leq\cdots\leq n_r$. Segments of the form $[n_j,n_{j+1}]$ will be called *edges*. A *marking* is a non-empty collection of edges satisfying the following two conditions. Firstly, each vertex n_j belongs to at most one edge. The vertices belonging to an edge are called *bound*, the other vertices are called *free*. Secondly, we require that for every free vertex n_l, either

(i) there exists a minimal $f(l)>l$ such that $n_{f(l)}$ is bound, and for all $l\leq i<f(l)$, $n_{i+1}-n_i\leq n_{f(l)+1}-n_{f(l)}$; or
(ii) there exists a maximal $p(l)<l$ such that $n_{p(l)}$ is bound, and for all $p(l)<i\leq l$, $n_i-n_{i-1}\leq n_{p(l)}-n_{p(l)-1}$.

If (i) holds we will say that n_l is associated to the edge $[n_{f(l)},n_{f(l)+1}]$ otherwise it is associated to $[n_{p(l)-1},n_{p(l)}]$.

Lemma 3.22 *There are constants $L=L(r)>0$ and $0<\theta<1$ such that*

$$\left|\mathbb{E}\Big[\prod_{i=1}^{r}f_{n_i}\Big]\right|\leq L\sum_{\textit{markings}}\ \prod_{[n_j,n_{j+1}]\textit{ is an edge}}\left(\theta^{(n_{j+1}-n_j)}\ \ \widetilde{u}_{n_j}\widetilde{u}_{n_{j+1}}\right).$$

Proof If $r=1$ then the result holds since $\mathbb{E}[f_n]=0$ (in this case there are no markings, and we let the empty sum be equal to zero).

If $r=2$ then the lemma says that $|\mathbb{E}\,[f_{n_1}f_{n_2}]|\leq\text{const.}\theta^{n_2-n_1}\|f_{n_1}\|_{L^2}\|f_{n_2}\|_{L^2}$. This is a consequence of uniform ellipticity, see Proposition 2.25.

For $r\geq 3$ we use induction. Take j such that $n_{j+1}-n_j$ is the largest. Then

$$\mathbb{E}\left[\prod_{i=1}^{r}f_{n_i}\right]=\mathbb{E}\left[\prod_{i=1}^{j}f_{n_i}\right]\mathbb{E}\left[\prod_{i=j+1}^{r}f_{n_i}\right]+O\left(\theta^{(n_{j+1}-n_j)}\left\|\prod_{i=1}^{j}f_{n_i}\right\|_2\left\|\prod_{i=j+1}^{r}f_{n_i}\right\|_2\right).$$

Let $K:=\text{ess sup}\,|\mathsf{f}|$, then the second term is less than $C_{mix}\theta^{(n_{j+1}-n_j)}\widetilde{u}_{n_j}\widetilde{u}_{n_{j+1}}K^{r-2}$. Thus this term is controlled by the marking with only one marked edge $[n_j,n_{j+1}]$. Applying the inductive assumption to each factor in the first term gives the result. □

Lemma 3.23 $\exists\overline{C}_r>0$ *such that for every set C of r-tuples $1\leq n_1\leq\cdots\leq n_r\leq N$,*

$$\Gamma_{\mathsf{C}}:=\sum_{(n_1,\ldots,n_r)\in\mathsf{C}}\left|\mathbb{E}\Big[\prod_{i=1}^{r}f_{n_i}\Big]\right|\leq\overline{C}_rV_N^{\lfloor r/2\rfloor}.$$

Lemma 3.23 implies Lemma 3.21 since $\mathbb{E}(S_N^r)$ is a linear combination of expectations of products along r-tuples, with combinatorial coefficients which only depend on r. Therefore it suffices to prove Lemma 3.23.

Proof By Lemma 3.22, $\Gamma_C \le L \sum_{(n_1,\dots,n_r)\in C} \sum_{\substack{\text{markings } (e_1,\dots,e_s)\\ \text{of } (n_1,\dots,n_r)}} \prod_{j=1}^{s} \left(\widetilde{u}_{e_j^-}\widetilde{u}_{e_j^+}\theta^{(e_j^+-e_j^-)}\right)$ where the marked edges are $e_j = [e_j^-, e_j^+]$, $j = 1, \dots, s$.

Collecting all terms with a fixed set of marked edges $(e_1, \dots, e_s)$, we obtain

$$\Gamma_C \le C(r) \sum_s \sum_{(e_1,\dots,e_s)} \prod_{j=1}^{s} \left(\widetilde{u}_{e_j^-}\widetilde{u}_{e_j^+}\theta^{(e_j^+-e_j^-)}(e_j^+ - e_j^-)^{r-2}\right) \tag{3.17}$$

where $C(r)\prod_j (e_j^+ - e_j^-)^{r-2}$ accounts for all tuples which admit a marking $(e_1, \dots e_s)$. Indeed, for every edge $e = [e^-, e^+]$ there are at most $0 \le j \le r-2$ vertices which may be associated to e, and these vertices are inside

$$[e^- - (r-2)(e^+ - e^-), e^-) \cup (e^+, e^+ + (r-2)(e^+ - e^-)]\,.$$

Thus there are at most $2(r-2)(e^+ - e^-)$ choices to place each vertex associated to a given edge. This gives the following bound for the number of possibilities for tuples with marking $(e_1, \dots, e_s)$:

$$\prod_e \left(\sum_{j=0}^{r-2} [2(r-2)(e^+ - e^-)]^j\right) \le C(r) \prod_e (e^+ - e^-)^{r-2}.$$

The sum over $(e_1, \dots e_s)$ in (3.17) is estimated by $\left(\sum_{n=1}^{N-1}\sum_{m=1}^{N-n} \widetilde{u}_n \widetilde{u}_{n+m} \theta^m m^{r-2}\right)^s$. For each m, $\sum_n \widetilde{u}_n \widetilde{u}_{n+m} = O(V_N)$ due to the Cauchy-Schwartz inequality and because $\sum_{n=1}^{N} \widetilde{u}_n^2 \le CV_N$ by assumption. Summing over m gives $\Gamma_C \le \text{const.} \sum_{2s\le r} V_N^s$ where the condition $2s \le r$ appears because each edge involves two distinct vertices, and no vertex belongs to more than one edge. The result follows. □

3.2.6 Characterization of Center-Tight Additive Functionals

We prove Theorem 3.8. Suppose f is an a.s. uniformly bounded functional on a uniformly elliptic array X. We will show that the following conditions are equivalent:

(a) $\mathrm{Var}(S_N) = O(1)$;
(b) f is the sum of a gradient and an additive functional with summable variance;
(c) f is center-tight.

(a)⇒(b): By the gradient lemma $\mathsf{f} = \nabla \mathsf{a} + (\widetilde{\mathsf{f}} + \mathsf{c})$, where a is a.s. uniformly bounded, $c_n^{(N)}$ are uniformly bounded constants, and $\|\widetilde{f}_n\|_2 \le u_n^{(N)}$. By Theorem 3.6, $\sup_N \sum_{n=3}^{k_N} (u_n^{(N)})^2 < \infty$, so $\widetilde{\mathsf{f}} + \mathsf{c}$ has summable variance, proving (b).

(b)⇒(c): We already saw that gradients and functionals with summable variance are center-tight. Since the sum of center-tight functionals is center-tight, (c) is proved.

(c)⇒(a): Assume by way of contradiction that $\exists N_i \uparrow \infty$ such that $V_{N_i} := \mathrm{Var}(S_{N_i}) \to \infty$. By Dobrushin's central limit theorem, $\frac{S_{N_i} - \mathbb{E}(S_{N_i})}{\sqrt{V_{N_i}}}$ converges in distribution to a standard Gaussian distribution, so $|\exp(itS_{N_i}/\sqrt{V_{N_i}})| \to$

$e^{-t^2/2}$. But center-tightness implies that there are constants m_N such that $\frac{S_{N_i}-m_{N_i}}{\sqrt{V_{N_i}}}$ converges in distribution to zero, so $|\exp(\mathrm{i}tS_{N_i}/\sqrt{V_{N_i}})| \to 1$, a contradiction. □

3.2.7 Proof of the Two-Series Theorem

We prove Theorem 3.12.

Part (1). We suppose that f has summable variance, and prove that $S_N - \mathbb{E}(S_N)$ converges a.e. to a finite limit.

Let $f_0^* := 0$, $f_n^* := f_n(X_n, X_{n+1}) - \mathbb{E}[f_n(X_n, X_{n+1})]$, let $\mathscr{F}_0$ denote the trivial σ-algebra, and let $\mathscr{F}_n$ denote the σ-algebra generated by $X_1, \ldots, X_n$. Then f_k^* is $\mathscr{F}_{k+1}$-measurable, so $f_k^* = \mathbb{E}(f_k^*|\mathscr{F}_{k+1}) - \mathbb{E}(f_k^*|\mathscr{F}_0) = \sum_{n=0}^{k} \mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)$. Therefore,

$$
\begin{aligned}
\sum_{k=1}^{N} f_k^* &= \sum_{k=1}^{N}\sum_{n=0}^{k}\Big[\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big] = \sum_{n=0}^{N}\sum_{k=n}^{N}\Big[\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big] \\
&\overset{!}{=} \sum_{n=0}^{N}\sum_{k=n}^{\infty}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) - \sum_{n=0}^{N}\sum_{k=N+1}^{\infty}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) \\
&\overset{!}{=} \sum_{n=0}^{N}\sum_{k=n}^{\infty}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) - \sum_{k=N+1}^{\infty}\sum_{n=0}^{N}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) \\
&= \sum_{n=0}^{N}\sum_{k=n}^{\infty}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) - \sum_{k=N+1}^{\infty}\mathbb{E}(f_k^*|\mathscr{F}_{N+1}). \qquad (3.18)
\end{aligned}
$$

To justify the marked equalities, we need to show that the infinite sums converge absolutely in L^2. By (2.12), $\|\mathbb{E}(f_k^*|\mathscr{F}_{n+1})\|_2 + \|\mathbb{E}(f_k^*|\mathscr{F}_n)\|_2 \le 2C_{mix}\sqrt{\mathrm{Var}(f_k)}\theta^{k-n-1}$. Since $\sum_{n=1}^{\infty}\mathrm{Var}(f_n) < \infty$, $\sum_{n=0}^{N}\sum_{k=n}^{\infty}\Big\|\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big\|_2 < \infty$.

$$
\text{Let } \Delta_n := \sum_{k=n}^{\infty}\Big(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\Big) \text{ and } Z_N := \sum_{k=N+1}^{\infty}\mathbb{E}(f_k^*|\mathscr{F}_{N+1}). \qquad (3.19)
$$

Equation (3.18) leads to the following **martingale-coboundary decomposition**:[3]

$$
S_N - \mathbb{E}(S_N) = \sum_{n=0}^{N}\Delta_n - Z_N. \qquad (3.20)
$$

To finish the proof, we show that $\sum_{n=0}^{\infty}\Delta_n$ and $\lim_{N\to\infty} Z_N$ exist a.s.

Claim 1. $M_N := \sum_{n=0}^{N-1}\Delta_n$ *is a martingale relative to* $\{\mathscr{F}_N\}$, *and* $\sup\|M_N\|_2 < \infty$. *Consequently,* $\lim M_N$ *exists and is finite almost surely.*

Proof of the Claim. The martingale property is because $M_{N+1} - M_N = \Delta_N$, and

[3] Indeed, we will soon see that $\sum_{n=0}^{N}\Delta_n$ is a martingale (Claim 1); and $Z_N \equiv Z_N - Z_{-1} = \sum_{n=0}^{N}(Z_n - Z_{n-1})$, a sum of "coboundaries" $Z_n - Z_{n-1}$ which tend to zero. (See Footnote 1 on page 31.)

$$\mathbb{E}(\Delta_N|\mathscr{F}_N) \stackrel{!}{=} \sum_{k=N}^{\infty} \mathbb{E}(\mathbb{E}(f_k^*|\mathscr{F}_{N+1})|\mathscr{F}_N) - \mathbb{E}(\mathbb{E}(f_k^*|\mathscr{F}_N)|\mathscr{F}_N) = 0.$$

To justify $\stackrel{!}{=}$, we note that the series $\Delta_N = \sum_{k=N}^{\infty} \left[\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n)\right]$ converges in L^2, because $\|\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k|\mathscr{F}_n)\|_\infty = O(\theta^{k-n})$. Therefore, the conditional expectation can be calculated term-by-term.

Next we show that $\|M_N\|_2$ is uniformly bounded.

$$\begin{aligned}
\|M_{N+1}\|_2 &\le \left\|\sum_{n=0}^{N}\sum_{k=n}^{\infty}(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n))\right\|_2 \le \left\|\sum_{k=0}^{\infty}\sum_{n=0}^{k\wedge N}(\mathbb{E}(f_k^*|\mathscr{F}_{n+1}) - \mathbb{E}(f_k^*|\mathscr{F}_n))\right\|_2 \\
&= \left\|\sum_{k=0}^{\infty}\mathbb{E}(f_k^*|\mathscr{F}_{(k\wedge N)+1})\right\|_2 \le \left\|\sum_{k=0}^{N} f_k^*\right\|_2 + \left\|\sum_{k=N+1}^{\infty}\mathbb{E}(f_k^*|\mathscr{F}_{N+1})\right\|_2 \\
&\le \sqrt{\sum_{k=0}^{N}\|f_k^*\|_2^2 + 2\sum_{0\le k<\ell\le N}\mathrm{Cov}(f_k^*, f_\ell^*)} + \sum_{k=N+1}^{\infty}\|\mathbb{E}(f_k^*|\mathscr{F}_{N+1})\|_\infty \\
&\le \sqrt{\sum_{k=0}^{\infty}\|f_k^*\|_2^2 + 2C_{mix}\sum_{0\le k<\ell\le\infty}\theta^{\ell-k}\|f_k^*\|_2\|f_\ell^*\|_2} + C_{mix}\sum_{k=N+1}^{\infty}\|f_k^*\|_\infty\theta^{k-(N+1)}.
\end{aligned}$$

The last expression is uniformly bounded, because $\sum\|f_k^*\|_2^2 = \sum \mathrm{Var}(f_k) < \infty$, and $\sum_{0\le k<\ell<\infty}\theta^{\ell-k}\|f_k^*\|_2\|f_\ell^*\|_2 \le \sum_{r=1}^{\infty}\theta^r\sum_{k=0}^{\infty}\|f_k^*\|_2\|f_{k+r}^*\|_2 \stackrel{!}{\le} \frac{1}{1-\theta}\sum_{k=0}^{\infty}\|f_k^*\|_2^2$, by the inequality $|ab| \le \frac{1}{2}(a^2+b^2)$.

Claim 2. $Z_N \xrightarrow[N\to\infty]{} 0$ *a.s.*

Proof of the Claim. In fact we will show that $\sum Z_N^2 < \infty$ almost surely.

$$\frac{1}{2}\mathbb{E}\left(\sum_{N=1}^{\infty} Z_N^2\right) \le \sum_{N=1}^{\infty}\sum_{k_2\ge k_1>N}\mathbb{E}[\mathbb{E}(f_{k_1}^*|\mathscr{F}_{N+1})\mathbb{E}(f_{k_2}^*|\mathscr{F}_{N+1})] = \sum_{N=1}^{\infty}\sum_{k_2\ge k_1>N}\mathbb{E}[f_{k_2}^*\mathbb{E}(f_{k_1}^*|\mathscr{F}_{N+1})]$$

$$\stackrel{(2.13)}{\le} C_{mix}\sum_{N=1}^{\infty}\sum_{k_2\ge k_1>N}\theta^{k_2-(N+1)}\|f_{k_2}^*\|_2\|\mathbb{E}(f_{k_1}^*|\mathscr{F}_{N+1})\|_2 \stackrel{(2.12)}{\le} C_{mix}^2\sum_{N=1}^{\infty}\sum_{k_2\ge k_1>N}\theta^{k_2-(N+1)}\|f_{k_2}^*\|_2\cdot\theta^{k_1-(N+1)}\|f_{k_1}^*\|_2$$

$$= C_{mix}^2\theta^{-2}\sum_{j\ge 0}\theta^j\sum_{k>0}\theta^{2k}\sum_{N=1}^{\infty}\|f_{k+N+j}^*\|_2\|f_{k+N}^*\|_2, \text{ where } j = k_2 - k_1,\ k = k_1 - N.$$

Since $ab \le \frac{a^2+b^2}{2}$, the innermost sum is less than $\sum\|f_n^*\|_2^2$, which is finite by the assumption that f has summable variance. So $\mathbb{E}(\sum Z_n^2) < \infty$, proving the claim.

By Claims 1 and 2 and equation (3.20), $\lim(S_N - \mathbb{E}(S_N))$ exists almost surely, and the first part of Theorem 3.12 is proved. □

Part (2). Suppose first that $\mathrm{Var}(S_N) \not\to \infty$. By Corollary 3.7 (a direct consequence of Theorem 3.6), $\mathrm{Var}(S_N)$ is bounded, and therefore f is center-tight. By Theorem 3.8, $\mathsf{f} = \nabla\mathsf{a} + \mathsf{h}$ where h has summable variance.

Trading constants between h and a, we may arrange that $\mathbb{E}[a_n(X_n)] = 0$ for all n. Then $\mathbb{E}(S_N) = \mathbb{E}[S_N(\mathsf{h})]$, and since $S_N(\mathsf{f}) = S_N(\nabla\mathsf{a}) + S_N(\mathsf{h})$, we get $S_N - a_{N+1}(X_{N+1}) - \mathbb{E}(S_N) = [S_N(\mathsf{h}) - \mathbb{E}(S_N(\mathsf{h}))] - a_1(X_1)$.

The last expression has a finite a.s. limit, by part (1). This proves the direction ($\Rightarrow$).

To see ($\Leftarrow$), assume that there are uniformly bounded measurable functions $a_n : \mathfrak{S}_n \to \mathbb{R}$ so that $S_N - a_{N+1}(X_{N+1}) - \mathbb{E}(S_N)$ has a finite limit almost everywhere, and assume by contradiction that $V_N \to \infty$.

Then $\lim_{N\to\infty} \dfrac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} = \lim_{N\to\infty} \dfrac{S_N - a_{N+1}(X_{N+1}) - \mathbb{E}(S_N)}{\sqrt{V_N}} = 0$ a.s. , whence $\frac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N\to\infty]{\text{dist}} 0$. But this contradicts Dobrushin's CLT, which says that if $V_N \to \infty$, then $\frac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}$ converges in distribution to the Gaussian law. □

*3.3 The Almost Sure Invariance Principle

The CLT approximates the distribution of S_N by the Gaussian distribution with variance V_N. The next result approximates $(S_N, S_{N+1}, S_{N+2}, \dots)$ by a path of standard Brownian motion at times $(V_N, V_{N+1}, V_{N+2}, \dots)$.

We remind the reader that a **standard Brownian Motion** on a probability space $\widetilde{\Omega}$ is a one-parameter family of real-valued functions $W(t) : \widetilde{\Omega} \to \mathbb{R}$ $(t \geq 0)$ s.t.:

(1) $(t, \omega) \mapsto W(t)(\omega)$ is measurable;
(2) $W(0) \equiv 0$;
(3) $W(t) - W(s)$ is normally distributed with mean zero and variance $|t - s|$;
(4) for all $0 \leq t_1 \leq \cdots \leq t_n$, the random variables $W(t_i) - W(t_{i-1})$ $(i = 2, \dots, n)$ are independent;
(5) for a.e. $\omega \in \widetilde{\Omega}$, the function $t \mapsto W(t)$ is continuous on $[0, \infty)$.

Theorem 3.24 *Let* f *be a non center-tight uniformly bounded additive functional of a uniformly elliptic Markov chain* X*. Let* $f_n := f_n(X_n, X_{n+1})$ *and suppose that* $\mathbb{E}(f_n) = 0$ *for all* n*. Denote* $S_N = \sum_{n=1}^{N} f_n$ *and let* V_N *be the variance of* S_N*. Then there exist a number* $\delta > 0$*, a probability space* $(\widetilde{\Omega}, \widetilde{\mathbb{P}})$*, and measurable functions* $\widetilde{S}_N, W(t), \widetilde{N} : \widetilde{\Omega} \to \mathbb{R}$ $(N \in \mathbb{N}, t \geq 0)$ *such that*

- $\{S_N\}_{N\geq 1}$ *and* $\{\widetilde{S}_N\}_{N\geq 1}$ *have the same distribution;*
- $W(t)$ *is a standard Brownian motion on* $\widetilde{\Omega}$*;*
- *for a.e.* $\omega \in \widetilde{\Omega}$*,* $|\widetilde{S}_N(\omega) - W(V_N)(\omega)| \leq V_N^{1/2-\delta}$ *for* $N \geq \widetilde{N}(\omega)$*.*

Corollary 3.25 (Law of the Iterated Logarithm) *With probability one,*

$$\limsup_{N\to\infty} \frac{S_N}{\sqrt{V_N \ln\ln V_N}} = \sqrt{2}, \quad \liminf_{N\to\infty} \frac{S_N}{\sqrt{V_N \ln\ln V_N}} = -\sqrt{2}.$$

The proof of these results relies on properties of martingales which we now recall, and which can be found in [95, Theorems A.1 and 2.2]:

Proposition 3.26 (Skorokhod's Embedding Theorem for Martingales) *Suppose that* $(\Delta_n, \mathcal{F}_n)$ *is a martingale difference sequence with mean zero and finite variance, defined on a probability space* Ω*. Then there is a probability space* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$*, a filtration* $\widetilde{\mathcal{F}}_n \subset \widetilde{\mathcal{F}}$ *and measurable functions* $\widetilde{\boldsymbol{\Delta}}_n, W(t), \tau_n : \widetilde{\Omega} \to \mathbb{R}$ *such that* $(\widetilde{\boldsymbol{\Delta}}_n, \widetilde{\mathcal{F}}_n)$ *is a martingale difference sequence,* $W(t)$ *is a standard Brownian Motion,* τ_n *is* $\widetilde{\mathcal{F}}_n$ *measurable, and the following holds with* $T_0 = 0$*,* $T_N = \tau_1 + \cdots + \tau_N$*:*

(a) $\{\widetilde{\boldsymbol{\Delta}}_n\}_{n\geq 1}$ *and* $\{\Delta_n\}_{n\geq 1}$ *have the same law;*
(b) $\widetilde{\boldsymbol{\Delta}}_n = W(T_n) - W(T_{n-1}) = W(T_{n-1} + \tau_n) - W(T_{n-1})$*;*
(c) $\mathbb{E}(\tau_n | \widetilde{\mathcal{F}}_{n-1}) = \mathbb{E}(\widetilde{\boldsymbol{\Delta}}_n^2 | \widetilde{\mathcal{F}}_{n-1})$*;*
(d) For each $p \geq 1$ *there is* C_p *such that* $\mathbb{E}(\tau_n^p | \widetilde{\mathcal{F}}_{n-1}) \leq C_p \mathbb{E}(\Delta_n^{2p} | \mathcal{F}_{n-1})$*.*

Proposition 3.27 (Doob's Maximal Inequality for Martingales) *Let* Δ_n *be a martingale difference, and let* $S_N = \sum_{n=1}^{N} \Delta_n$*. Then for each* $p > 1$*,* $\mathbb{E}\left(\max_{n\in[1,N]} |S_n|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left(|S_N|^p\right)$ *.*

Proof of Theorem 3.24. It suffices to prove the result in the case when f_n is a martingale difference sequence, with respect to the σ-algebras $\mathcal{F}_n$ generated by $X_1, \dots, X_n, X_{n+1}$. We can always reduce to this case, by replacing f_n with the Δ_n defined in (3.19). By (3.20) and Lemma 2.14, $S_N = \sum_{n=1}^{N} \Delta_n + O(1)$ and $\mathrm{Var}(S_N) = \mathrm{Var}\Big(\sum_{n=1}^{N} \Delta_n\Big) + O(1)$. So if we can prove the theorem for $\sum_{n=1}^{N} \Delta_n$, then we can prove it for $\sum_{n=1}^{N} f_n$.

Note that Δ_n is a martingale difference with zero mean and finite variance. The reader can easily verify using the Markov property that Δ_n has measurable versions of the form $\Delta_n = \Delta_n(X_n, X_{n+1})$. By (2.11), $\sup_n \operatorname{ess\,sup} |\Delta_n| < \infty$.

Henceforth we assume that f_n is a martingale difference sequence with zero mean. By Lemma 3.16, V_N is monotonically increasing.

Let $\{\widetilde{\mathbf{\Delta}}_n\}$, τ_n, T_N, $W(t)$ be the objects provided by Skorokhod's embedding theorem. We claim that W and $\widetilde{S}_N = \sum_{n=1}^{N} \widetilde{\mathbf{\Delta}}_N$ satisfy the conditions of Theorem 3.24. The proof consists of several steps.

Lemma 3.28 $\mathbb{E}(T_N) = V_N$*, and* $\mathrm{Var}(T_N) \le CV_N$ *for some positive constant* C*.*

Proof $\mathbb{E}(T_N) = V_N$, by Proposition 3.26 and Lemma 3.16. Split $T_N - \mathbb{E}(T_N) = \sum_{n=1}^{N} \left[\tau_n - \widetilde{\mathbf{\Delta}}_n^2\right] + \sum_{n=1}^{N} \left[\widetilde{\mathbf{\Delta}}_n^2 - \mathbb{E}(\widetilde{\mathbf{\Delta}}_n^2)\right]$.

By the Cauchy-Schwarz inequality, it is enough to show that the variance of each sum is $O(V_N)$. By Proposition 3.26(c), the first sum is a martingale, hence by Lemma 3.16,

$$\mathrm{Var}\left(\sum_{n=1}^{N} \left[\tau_n - \widetilde{\mathbf{\Delta}}_n^2\right]\right) = \sum_{n=1}^{N} \mathbb{E}\left(\left[\tau_n - \widetilde{\mathbf{\Delta}}_n^2\right]^2\right) \overset{!}{\le} \text{const.} \sum_{n=1}^{N} \mathbb{E}(\widetilde{\mathbf{\Delta}}_n^4),$$

where the last step uses the Cauchy-Schwarz inequality, and Proposition 3.26(d) with $p = 2$. Since Δ_n are uniformly bounded,

$$\sum_{n=1}^{N} \mathbb{E}(\widetilde{\mathbf{\Delta}}_n^4) = \sum_{n=1}^{N} \mathbb{E}(\Delta_n^4) \le \text{const.} \sum_{n} \mathbb{E}(\Delta_n^2) \le \text{const.} V_N. \tag{3.21}$$

Next, $\mathrm{Var}\left(\sum_{n=1}^{N} \left[\widetilde{\mathbf{\Delta}}_n^2 - \mathbb{E}(\widetilde{\mathbf{\Delta}}_n^2)\right]\right) = \mathrm{Var}\left(\sum_{n=1}^{N} \left[\Delta_n^2 - \mathbb{E}(\Delta_n^2)\right]\right) \le \text{const.} \sum_{n=1}^{N} \mathbb{E}\left(\left[\Delta_n^2 - \mathbb{E}(\Delta_n^2)\right]^2\right) \le \text{const.} V_N$
where the first step follows by Prop. 3.26, second follows by Lemma 3.4 and the third follows by (3.21) This completes the proof. □

Let $\alpha_1 > 1$ be a numerical parameter to be chosen later. Let N_k be the smallest number such that $V_{N_k} \ge k^{\alpha_1}$. We denote $\mathcal{V}_k := V_{N_k}$. First, we will prove Theorem 3.24 along the sequence N_k. Then we will estimate the oscillation of $\widetilde{S}_N$ and $W(V_N)$ between consecutive N_k, and deduce the theorem for $N_k \le N < N_{k+1}$.

Lemma 3.29 *Suppose that* $\alpha_2 > \frac{1+1/\alpha_1}{2}$*. Then with probability one, for all large N we have* $|T_{N_k} - \mathcal{V}_k| \le \mathcal{V}_k^{\alpha_2}$*.*

Proof We saw above that $\mathcal{V}_k = \mathbb{E}(T_{N_k})$. By Lemma 3.28 and Chebyshev's inequality,

$$\mathbb{P}\left(|T_{N_k} - \mathcal{V}_k| > \mathcal{V}_k^{\alpha_2}\right) \le C\mathcal{V}_k^{1-2\alpha_2} \le Ck^{\alpha_1(1-2\alpha_2)}.$$

Since $\alpha_1(2\alpha_2 - 1) > 1$, the result follows from the Borel–Cantelli lemma. □

Lemma 3.30

(a) If $\alpha_3 > \alpha_2/2$*, then with probability one, for all large k,* $|\widetilde{S}_{N_k} - W(\mathcal{V}_k)| \le \mathcal{V}_k^{\alpha_3}$*.*
(b) If $\alpha_3 > \frac{1}{2}(1 - \alpha_1^{-1})$*, then with probability one, for all large k,* $\max_{t \in [\mathcal{V}_k, \mathcal{V}_{k+1}]} |W(t) - W(\mathcal{V}_k)| \le \mathcal{V}_k^{\alpha_3}$*.*

Proof By Lemma 3.29 and the identity $\widetilde{S}_{N_k} = W(T_{N_k})$,

$$\mathbb{P}\{|\widetilde{S}_{N_k} - W(\mathcal{V}_k)| \geq \mathcal{V}_k^{\alpha_3} \quad \text{for infinitely many } k\} \leq \mathbb{P}\Big\{\max_{t\in[\mathcal{V}_k-\mathcal{V}_k^{\alpha_2},\mathcal{V}_k+\mathcal{V}_k^{\alpha_2}]} |W(t) - W(\mathcal{V}_k)| \geq \mathcal{V}_k^{\alpha_3} \text{ for infinitely many } k\Big\}.$$

Thus to prove part (a), it suffices to show that $\sum_k \mathbb{P}\left(\max_{t\in[\mathcal{V}_k-\mathcal{V}_k^{\alpha_2},\mathcal{V}_k+\mathcal{V}_k^{\alpha_2}]} |W(t) - W(\mathcal{V}_k)| \geq \mathcal{V}_k^{\alpha_3}\right) < \infty.$

It suffices to check that $\sum_k \mathbb{P}\left(\max_{t\in[\mathcal{V}_k,\mathcal{V}_k+\mathcal{V}_k^{\alpha_2}]} |W(t) - W(\mathcal{V}_k)| \geq \mathcal{V}_k^{\alpha_3}\right) < \infty$, the interval $[\mathcal{V}_k - \mathcal{V}_k^{\alpha_2}, \mathcal{V}_k]$ completely analogous.

By the reflection principle, for any $a<b$ and $h>0$, $\mathbb{P}\left(\max_{t\in[a,b]}[W(t) - W(a)] \geq h\right)$ and $\mathbb{P}\left(\min_{t\in[a,b]}[W(t) - W(a)] \leq -h\right)$ are both equal to $2\mathbb{P}\left(W(b)-W(a) \geq h\right)$ (see [104, property (1.7.4)].) Hence

$$\mathbb{P}\left(\max_{t\in[\mathcal{V}_k,\mathcal{V}_k+\mathcal{V}_k^{\alpha_2}]} |W(t) - W(\mathcal{V}_k)| \geq \mathcal{V}_k^{\alpha_3}\right) \leq 4\mathbb{P}\left(W(\mathcal{V}_k + \mathcal{V}_k^{\alpha_2}) - W(\mathcal{V}_k) \geq \mathcal{V}_k^{\alpha_3}\right)$$

$$\overset{!}{=} 4\int_{\mathcal{V}_k^{\alpha_3-(\alpha_2/2)}}^{\infty} \frac{\mathrm{e}^{-u^2/2}}{\sqrt{2\pi}}\mathrm{d}u \overset{!!}{\leq} 4\mathrm{e}^{-(1/2)\mathcal{V}_k^{2\alpha_3-\alpha_2}} \leq 4\mathrm{e}^{-ck^{\beta_1}}, \tag{3.22}$$

where $c>0$ and $\beta_1=(2\alpha_3 - \alpha_2)\alpha_1 > 0$. (!) holds since $[W(\mathcal{V}_k + \mathcal{V}_k^{\alpha_2}) - W(\mathcal{V}_k)]/\mathcal{V}_k^{\alpha_2/2}$ is a standard normal, (!!) is because for $h > 0$, $\int_h^\infty \mathrm{e}^{-u^2/2}\mathrm{d}u \leq \int_h^\infty \frac{u}{h}\mathrm{e}^{-u^2/2}\mathrm{d}u = \frac{\mathrm{e}^{-h^2/2}}{h}$. Since (3.22) is summable in k, part (a) follows.

To prove part (b) we first claim that

$$\mathrm{Var}(S_{N_{k+1}} - S_{N_k}) \asymp k^{\alpha_1-1}. \tag{3.23}$$

Indeed $\mathcal{V}_{k+1} = \mathrm{Var}(S_{N_{k+1}}) = \mathcal{V}_k + \mathrm{Var}(S_{N_{k+1}} - S_{N_k}) + 2\mathrm{Cov}(S_{N_k}, S_{N_{k+1}} - S_{N_k})$. Since V_N is increasing, N_k is increasing. By (2.13) and the uniform boundness of f_n, there is $0 < \theta < 1$ such that

$$\mathrm{Cov}(S_{N_k}, S_{N_{k+1}} - S_{N_k}) = \sum_{n_1=1}^{N_k}\sum_{n_2=N_k+1}^{N_{k+1}} \mathrm{Cov}(f_{n_1}, f_{n_2}) \leq \text{const.} \sum_{n_1=1}^{N_k}\sum_{n_2=N_k+1}^{N_{k+1}} \theta^{n_2-n_1} \leq \text{const.} \sum_{m=1}^{\infty} m\theta^m \leq \text{const.}$$

So $\mathrm{Var}(S_{N_{k+1}}–S_{N_k})=\mathcal{V}_{k+1}–\mathcal{V}_k + O(1)$ and (3.23) follows from the definition of N_k.

Now similarly to part (a), we obtain $\mathbb{P}\left(\max_{t\in[\mathcal{V}_k,\mathcal{V}_{k+1}]} |W(t) - W(T_{N_k})| \geq \mathcal{V}_k^{\alpha_3}\right) \leq \mathrm{e}^{-ck^{\beta_2}}$, where $c > 0$ and

$$\beta_2 = 2\alpha_1\alpha_3 - (\alpha_1 - 1) = \alpha_1\left[2\alpha_3 + \frac{1}{\alpha_1} - 1\right] > 0. \tag{3.24}$$

Part (b) can now be proved the same way we proved part (a). □

Lemma 3.31 *If* $\alpha_3 > \dfrac{1 - 1/\alpha_1}{2}$*, then with probability one, for all large k*

$$\max_{n\in[N_k,N_{k+1}]} |\widetilde{S}_n - \widetilde{S}_{N_k}| \leq \mathcal{V}_k^{\alpha_3}.$$

Proof It suffices to prove the proposition with S_N in place of $\widetilde{S}_N$. For each $p \in 2\mathbb{N}$,

$$\mathbb{P}\left(\max_{n\in[N_k,N_{k+1}]}|S_n - S_{N_k}| \geq \mathcal{V}_k^{\alpha_3}\right) \overset{(1)}{\leq} \mathbb{E}\left(\max_{n\in[N_k,N_{k+1}]}|S_n - S_{n_k}|^p\right)\Big/\mathcal{V}_k^{\alpha_3 p}$$
$$\overset{(2)}{\leq} \text{const.}\frac{\mathbb{E}[|S_{N_{k+1}} - S_{N_k}|^p]}{\mathcal{V}_k^{\alpha_3 p}} \overset{(3)}{\leq} \text{const.}\frac{[\mathrm{Var}(S_{N_{k+1}} - S_{N_k})]^{p/2}}{\mathcal{V}_k^{\alpha_3 p}} \overset{(4)}{\leq} \text{const.}k^{-p\beta_2/2},$$

where β_2 is given by (3.24). Inequality (1) is by Markov's inequality. Inequality (2) is by Doob's maximal inequality. Inequality (3) uses the assumption that p is even, and the moment bound in Lemma 3.21. Inequality (4) uses (3.23).

Thus the lemma follows from Borel–Cantelli lemma after taking $p > \frac{1}{\beta_2}$. □

We are now ready to prove Theorem 3.24. Take

$$\alpha_3 > \tfrac{1}{2}\max(\alpha_2, 1 - \alpha_1^{-1}). \tag{3.25}$$

Given N take k such that $N_k \leq N < N_{k+1}$. Then $|\widetilde{S}_N - W(V_N)| \leq |\widetilde{S}_{N_k} - W(\mathcal{V}_k)| + |W(V_N) - W(\mathcal{V}_k)| + |\widetilde{S}_N - \widetilde{S}_{N_k}|$.

We claim that with probability one, each term is less than or equal to $\mathcal{V}_k^{\alpha_3}$, for all k large enough. For the first term, this follows from Lemma 3.30(a). For the second term, this follows from Lemma 3.30(b) and the monotonicity of V_N. For the last term, this follows from Lemma 3.31.

Thus, for every $\alpha_4 > \alpha_3$, with probability 1, for all large N, $|\widetilde{S}_N - W(V_N)| \leq 3\mathcal{V}_k^{\alpha_3} \leq \mathcal{V}_k^{\alpha_4} \leq V_N^{\alpha_4}$. Since α_2 can be taken arbitrary close to $\frac{1+1/\alpha_1}{2}$, we conclude from (3.25) that α_4 could be taken arbitrary close to $\min_{\alpha_1>1}\max\left(\frac{1+(1/\alpha_1)}{4}, \frac{1-(1/\alpha_1)}{2}\right) = \frac{1}{3}$. This shows that the theorem holds for any $\delta < \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$.

We note for future reference that in this proof, the final choice of constants is $\alpha_1 \downarrow 3$, $\alpha_2 \downarrow \frac{2}{3}$, $\alpha_3 \downarrow \frac{1}{3}$,and $\alpha_4 \downarrow \frac{1}{3}$, where $\alpha \downarrow c$ means that $\alpha > c$, and α can be taken arbitrary close to c. □

Proof of the Law of the Iterated Logarithm (Corollary 3.25). The law of the iterated logarithm for the Brownian Motion [68, Theorem 8.5.1] says that with probability one, $\limsup_{t\to\infty}\frac{W(t)}{\sqrt{t\ln\ln t}} = \sqrt{2}$. It follows that with probability one:

$$\begin{aligned}
\sqrt{2} &= \limsup_{t\to\infty}\frac{W(t)}{\sqrt{t\ln\ln t}} = \limsup_{k\to\infty}\frac{W(\mathcal{V}_k)}{\sqrt{\mathcal{V}_k\ln\ln\mathcal{V}_k}} && \text{by Lemma 3.30 with } \alpha_3 \approx \frac{1}{3}\\
&= \limsup_{k\to\infty}\frac{\widetilde{S}_{N_k}}{\sqrt{\mathcal{V}_k\ln\ln\mathcal{V}_k}} && \text{by Theorem 3.24}\\
&= \limsup_{N\to\infty}\frac{\widetilde{S}_N}{\sqrt{V_N\ln\ln V_N}} && \text{by Lemma 3.31 with } \alpha_3 \approx \frac{1}{3}.
\end{aligned}$$

By Theorem 3.24, $\{S_N\}$ and $\{\widetilde{S}_N\}$ are equal in law. Therefore, with probability one, $\limsup\frac{S_N}{\sqrt{V_N\ln\ln V_N}} = \sqrt{2}$.

By the symmetry f ↔ −f, the liminf is a.s. $-\sqrt{2}$ as well. □

3.4 Notes and References

The connection between the non-growth of the variance and a representation in terms of gradients is well-known for stationary stochastic processes. The first result in this direction we are aware of is Leonov's theorem [129]. He showed that the asymptotic variance of a homogeneous additive functional of a stationary homogeneous Markov chain is zero iff the additive functional is the sum of a gradient and a constant. Rousseau-Egele [168] and Guivarc'h & Hardy [88] extended this to the context of dynamical systems preserving an invariant Gibbs measure. Kifer [112], Conze & Raugi [31], Dragičević, Froyland & González-Tokman [64] have proved versions of Leonov's theorem for random and/or sequential dynamical systems (see §9.4).

The connection between center-tightness and gradients is a central feature of the theory of cocycles over ergodic transformations. Suppose $T : X \to X$ is an ergodic probability preserving transformation on a non-atomic probability space. For every measurable $f : X \to \mathbb{R}$, $\{f \circ T^n\}$ is a stationary stochastic process, and $S_N = f + f \circ T + \cdots + f \circ T^{N-1}$ are called the **ergodic sums of the cocycle f.** A **coboundary** is a function of the form $f = g - g \circ T$ with g measurable. Schmidt characterized cocycles with center-tight S_N as those arising from coboundaries [178, page 181]. These results extend to cocycles taking values in locally compact groups, see Moore & Schmidt [145] and Aaronson & Weiss [8]. For more on this, see Aaronson [1, chapter 8], and Bradley [16, chapters 8,19]. We also refer to [83] for an analogous result in the continuous setting.

The characterization of center-tightness for inhomogeneous Markov chains in Theorem 3.8 seems to be new. The inhomogeneous theory is different from the stationary theory in that there is another cause for center-tightness: Having summable variance. This cannot happen in the stationary homogeneous world, unless f_i =const.

We have already commented that if X satisfies the one-step ellipticity condition and $f_k = f_k(X_k)$, then the variance estimate in Theorem 3.6 can be replaced by the simpler estimate (3.1), see [50],[181],[56]. Theorem 3.6 for $f_k = f_k(X_k, X_{k+1})$ seems to be new.

Theorem 3.10 is a special case of a result of Dobrushin [50], which also applies to some unbounded additive functionals. Our proof follows the paper of Sethuraman & Varadhan [181] closely, except for minor changes needed to deal with additive functionals of the form $f_k(X_k, X_{k+1})$, instead of $f_k(X_k)$.

McLeish's lemma, the martingale CLT, and their proofs are due to McLeish [140]. We refer the reader to Hall & Heyde [95] for the history of this result, further extensions, and references.

Theorem 3.12 extends the Kolmogorov-Khintchin "two-series theorem" [117]. There are other extensions to sums of dependent random variables, for example for martingales (Hall & Heyde [95, chapter 2]), for sums of negatively dependent random variables (Matula [138]) and for Birkhoff sums of expanding maps (Conze & Raugi [31]). The proofs of theorems 3.10 and 3.12 use Gordin's "martingale-coboundary decomposition" [81], see also [95],[120].

Theorem 3.11 is due to B.A. Lifshits [131]. Actually, he proved a more general result which also applies to the unbounded additive functionals considered in Dobrushin's paper. For an extension to ϕ-mixing processes, see [143, Theorem 6.17].

The first results on the law of the iterated logarithm (LIL) were obtained by Khinchin [108] and Kolmogorov [115], in the setting of bounded i.i.d. random variables. Kolmogorov's result also applies to certain sequences of non-identically distributed but independent and bounded random variables, and this was further extended to the unbounded case by Hartman and Wintner [96].

The almost sure invariance principle (ASIP) for independent identically distributed random variables and its application to the proof of the LIL are due to Strassen [193]. He used Skorokhod's embedding theorem [186]. Komlós, Major and Tusnády found an alternative proof, which gives better error rates [118].

In the stationary case, a classical application of ASIP is the *functional central limit theorem*, which says that a random function $W_N : [0, 1] \to \mathbb{R}$ obtained by linear interpolation of the points $W_N(\frac{n}{N}) = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{N}}$ converges in law as $N \to \infty$ to a Brownian Motion. In the inhomogeneous case, such results are only available after a random time change, so they are more complicated to state. See [95] for the precise statements.

Skorokhod's embedding theorem was extended to martingales in [36, 66, 194], and Stout extended the ASIP to martingales differences in [192]. Philipp & Stout [157] gave a further extension to weakly dependent random variables, including Markov chains. Cuny and Merlevède [35] proved the ASIP for reverse martingale differences. The martingale methods can be used to prove the ASIP for time-series of dynamical systems, homogeneous and inhomogeneous, see e.g. [43], [63], [97] (this is a very partial list).

For an alternative approach to the ASIP, based on perturbation operators, see [86]. This was applied in the inhomogeneous setup in [65].

Chapter 4
The Essential Range and Irreducibility

Abstract The local limit theorem may fail for additive functionals whose range can be reduced by subtracting a center-tight functional. In this chapter we study the structure of such functionals, and calculate the smallest possible algebraic range which can be obtained this way.

4.1 Definitions and Motivation

Let $\mathsf{f} = \{f_n\}$ be an additive functional of a Markov chain $\mathsf{X} := \{X_n\}$. The **algebraic range** of (X, f) is the intersection $G_{alg}(\mathsf{X}, \mathsf{f})$ of all closed groups G such that

$$\forall n\ \exists c_n \in \mathbb{R} \text{ s.t. } \mathbb{P}[f_n(X_n, X_{n+1}) - c_n \in G] = 1. \tag{4.1}$$

Later (Lemma 4.15), we will see that $G_{alg}(\mathsf{X}, \mathsf{f})$ itself satisfies (4.1), therefore $G_{alg}(\mathsf{X}, \mathsf{f})$ is the smallest closed group satisfying (4.1).

Example 4.1 (The Simple Random Walk) Suppose $\{X_n\}$ are independent random variables such that $\mathbb{P}(X_n = \pm 1) = \frac{1}{2}$, and let $f_n(x, y) = x$. Then $S_n = X_1 + \cdots + X_n$ is the simple random walk on $\mathbb{Z}$. The algebraic range in this case is $2\mathbb{Z}$.

Proof: $G_{alg} \subset 2\mathbb{Z}$, because we can take $c_n := -1$. Assume by contradiction that $G_{alg} \subsetneq 2\mathbb{Z}$, then $G_{alg} = t\mathbb{Z}$ for $t \geq 4$, and the supports of S_n are cosets of $t\mathbb{Z}$. But this is false, because $\exists a_1, a_2$ s.t. $|a_1 - a_2| < t$ and $\mathbb{P}(S_n = a_i) \neq 0$: For n odd take $a_i = (-1)^i$, and for n even take $a_i = 1 + (-1)^i$. □

The **lattice case** is the case when $G_{alg}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ for some $t \geq 0$. The **non-lattice case** is the case when $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$. This distinction appears naturally in the study of the LLT for the following reason. If $G_{alg}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ and $\gamma_N := c_1 + \cdots + c_N$, then

$$\mathbb{P}(S_N \in \gamma_N + t\mathbb{Z}) = 1 \text{ for all } N.$$

In this case, $\mathbb{P}[S_N - z_N \in (a, b)] = 0$ whenever $|a - b| < t$ and $z_N + (a, b)$ falls inside the gaps of $\gamma_N + t\mathbb{Z}$, and $\mathbb{P}(S_N - z_N \in (a, b)) \overset{?}{\sim} \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$ fails. This is the **lattice obstruction** to the local limit theorem.

There is a related, but more subtle, obstruction. An additive functional f is called **reducible** on X, if there is another additive functional g on X such that $\mathsf{f} - \mathsf{g}$ is center-tight, and $G_{alg}(\mathsf{X}, \mathsf{g}) \subsetneq G_{alg}(\mathsf{X}, \mathsf{f})$. In this case we say that g is a **reduction** of f, and that $G_{alg}(\mathsf{X}, \mathsf{g})$ is a **reduced range** of f. An additive functional without reductions is called **irreducible**.

Example 4.2 (Simple Random Walk with Continuous First Step) Suppose $\{X_n\}_{n\geq 1}$ are independent real valued random variables such that X_1 has a continuous distribution $\mathfrak{F}$ with compact support, and $X_2, X_3, \ldots$ are equal to ± 1 with equal probabilities. Let $f_n(x, y) = x$, then $S_n = X_1 + X_2 + \cdots + X_n$. Because of the continuously distributed first step, $G_{alg}(\mathsf{f}) = \mathbb{R}$. But if we subtract from f the center-tight functional c with components

$$c_n(x, y) = x \text{ when } n = 1 \text{ and } c_n(x, y) \equiv 0 \text{ when } n > 1,$$

then the result $\mathsf{g} := \mathsf{f} - \mathsf{c}$ has algebraic range $2\mathbb{Z}$. So f is reducible.

The reduction g satisfies the lattice local limit theorem (see Chapter 1), because it generates the (delayed) simple random walk. But for a general distribution $\mathfrak{F}$, $\mathsf{f} = \mathsf{g} + \mathsf{c}$ does not satisfy the LLT, lattice or non-lattice.

This can be seen by direct calculations, using the fact that the distribution of S_n is the convolution of $\mathfrak{F}$ and the centered binomial distribution. See chapter 6 for details.

Here we see an instance of the **reducibility obstruction** to the LLT: A situation when the LLT fails because the additive functional is a sum of a lattice term which satisfies the lattice LLT and a non-lattice center-tight term which spoils it.

The reducibility obstruction to the LLT raises the following questions:

(1) Given a reducible additive functional f, is there an "optimal" center-tight functional c such that $\mathsf{f} - \mathsf{c}$ is irreducible?

(2) What is the algebraic range of the optimal reduction?

Motivated by these questions, we introduce the **essential range** of f:

$$G_{ess}(\mathsf{X}, \mathsf{f}) := \bigcap \left\{ G_{alg}(\mathsf{X}, \mathsf{g}) : \mathsf{f} - \mathsf{g} \text{ is center-tight} \right\}.$$

This is a closed sub-group of $G_{alg}(\mathsf{X}, \mathsf{f})$. In this language, f is irreducible iff $G_{ess}(\mathsf{X}, \mathsf{f}) = G_{alg}(\mathsf{X}, \mathsf{f})$, and an optimal reduction is g such that $\mathsf{f} - \mathsf{g}$ is center-tight, and $G_{alg}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{f})$.

4.2 Main Results

4.2.1 Markov Chains

The questions raised at the end of the last section can be answered using the structure constants $d_n(\xi)$ introduced in (2.26).

Assume henceforth that f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X. Define the **co-range** of f to be the set

$$H(\mathsf{X}, \mathsf{f}) := \left\{ \xi \in \mathbb{R} : \sum_{n=3}^{\infty} d_n(\xi)^2 < \infty \right\}.$$

Theorem 4.3 *If* f *is center-tight then* $H(\mathsf{X}, \mathsf{f}) = \mathbb{R}$*, and if not then either* $H(\mathsf{X}, \mathsf{f}) = \{0\}$*, or* $H(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ *for some* $t \geq \pi/(3 \operatorname{ess\,sup} |\mathsf{f}|)$.

Theorem 4.4 *(a) If* $H(\mathsf{X}, \mathsf{f}) = 0$*, then* $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$.
(b) If $H(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ *with* $t \neq 0$*, then* $G_{ess}(\mathsf{X}, \mathsf{f}) = \frac{2\pi}{t}\mathbb{Z}$.
(c) If $H(\mathsf{X}, \mathsf{f}) = \mathbb{R}$*, then* $G_{ess}(\mathsf{X}, \mathsf{f}) = \{0\}$.

Theorem 4.5 *There exists an irreducible uniformly bounded additive functional* g *such that* $\mathsf{f} - \mathsf{g}$ *is center-tight, and* $G_{alg}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{f})$.

Corollary 4.6 *If* $G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ *with* $t \neq 0$*, then* $|t| \leq 6 \operatorname{ess\,sup} |\mathsf{f}|$.

The corollary follows directly from Theorems 4.3 and 4.4(b).

4.2.2 Markov Arrays

Let f be an additive functional on a Markov array X with row lengths $k_N + 1$, then we can make the following definitions:

- The **algebraic range** $G_{alg}(\mathsf{X},\mathsf{f})$ is the intersection of all closed subgroups G of $\mathbb{R}$ such that for all N and $1 \le k \le k_N$, $\exists c_k^{(N)} \in \mathbb{R}$ s.t. $\mathbb{P}\left[f_k^{(N)}(X_k^{(N)}, X_{k+1}^{(N)}) - c_k^{(N)} \in G\right] = 1$.
- The **essential range** $G_{ess}(\mathsf{X},\mathsf{f})$ is the intersection of the algebraic ranges of all additive functionals of the form $\mathsf{f} - \mathsf{h}$ where h is center-tight.
- The **co-range** is $H(\mathsf{X},\mathsf{f}) := \{\xi \in \mathbb{R} : \sup_N \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty\}$.
- An additive functional f is called **irreducible** if $G_{ess}(\mathsf{X},\mathsf{f}) = G_{alg}(\mathsf{X},\mathsf{f})$.

This is consistent with the definitions for Markov chains, see Corollary 4.8 below.

Theorem 4.7 *The results of Theorems 4.3, 4.4, 4.5 and of Corollary 4.6 hold for all a.s. uniformly bounded additive functionals on uniformly elliptic Markov arrays.*

Corollary 4.8 *Suppose* $\mathsf{f} = \{f_n\}$ *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* $\mathsf{X} = \{X_n\}$. *Let* $\widetilde{\mathsf{f}} = \{f_n^{(N)}\}$ *be an additive functional on a Markov array* $\widetilde{\mathsf{X}} = \{X_n^{(N)}\}$ *s.t.* $f_n^{(N)} = f_n$ *and* $X_n^{(N)} = X_n$. *Then* $G_{alg}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = G_{alg}(\mathsf{X},\mathsf{f})$, $G_{ess}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = G_{ess}(\mathsf{X},\mathsf{f})$, $H(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = H(\mathsf{X},\mathsf{f})$.

Proof The equality of the algebraic ranges and of the co-ranges is trivial, but the equality of the essential ranges requires justification, because some center-tight functionals on Markov arrays are not of the form $h_n^{(N)} = h_n$. However, since the co-ranges agree, the essential ranges also agree, by the version of Theorem 4.4 for arrays. □

4.2.3 Hereditary Arrays

Some results for Markov chains do not extend to general Markov arrays. Of particular importance is the following fact, which we will need for the proof of the LLT in Chapter 5. Recall the definition of $D_N(\xi)$ from (2.26).

Theorem 4.9 *Suppose* f *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov* <u>*chain*</u> X, *then*

$$D_N(\xi) \xrightarrow[N\to\infty]{} \infty \text{ uniformly on compact subsets of } H(\mathsf{X},\mathsf{f})^c. \tag{4.2}$$

Proof Suppose $\xi \in \mathbb{R}\setminus H(\mathsf{X},\mathsf{f})$, then $\sup_N D_N(\xi) = \infty$, so $D_N(\xi) = \sum_{k=3}^{N} d_k(\xi)^2 \xrightarrow[N\to\infty]{} \sum_{k=3}^{\infty} d_k(\xi)^2 \equiv \sup_N D_N(\xi) = \infty$. The convergence is uniform on compacts, because in the case of Markov chains, $D_N(\xi)$ is non-decreasing, and $\xi \mapsto D_N(\xi)$ are continuous. □

Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array X. We call (X,f)

- **hereditary**, if $D_N(\xi) \xrightarrow[N\to\infty]{} \infty$ for all $\xi \in \mathsf{H}(\mathsf{X},\mathsf{f})^c$; and
- **stably hereditary**, if $D_N(\xi) \xrightarrow[N\to\infty]{} \infty$ uniformly on compacts in $\mathsf{H}(\mathsf{X},\mathsf{f})^c$.

By Theorem 4.9, every a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain is stably hereditary. But as the following two examples show, this is not the case for arrays.

Example 4.10 (Irreducible but Not Hereditary) Let X_n be a sequence of independent uniform random variables with mean zero and variance one. Let

$$X_k^{(N)} = \begin{cases} X_k & 1 \le k \le N+1, \quad N \text{ odd}, \\ 0 & 1 \le k \le N+1, \quad N \text{ even}. \end{cases}$$

Let $f_k^{(N)}(x, y) := x$. Then for every $\xi \in H(\mathsf{X},\mathsf{f})^c$, $D_N(\xi) \not\to \infty$.

Proof. Suppose $P = \left(X_{n-2}\,\begin{matrix}X_{n-1} & X_n\\ Y_{n-1} & Y_n\end{matrix}\,, Y_{n+1}\right)$ is a random level $2N+1$ hexagon at position n, then $\Gamma(P) = X_{n-1} + X_n - Y_{n-1} - Y_n$ where X_i, Y_j are independent random variables, each having uniform distribution with mean zero and variance one. The distribution of $\Gamma(P)$ is independent of n and N, and as a result,

$$d_n^{(2N+1)}(\xi)^2 = \mathbb{E}(|e^{i\xi\Gamma(P)} - 1|^2) = c(\xi),$$

where $c(\xi)$ is independent of n. In addition, $c(\xi) > 0$ for $\xi \neq 0$, because the distribution of $\Gamma(P)$ is continuous. So $D_{2N+1}(\xi) = (2N-1)c(\xi) \xrightarrow[N\to\infty]{} \infty$ on $\mathbb{R}\setminus\{0\}$. Thus $H(\mathsf{X},\mathsf{f}) = \{0\}$ and $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$. But $D_N(\xi) \not\to \infty$ for $\xi \neq 0$, because $D_{2N}(\xi) = 0$. □

Example 4.11 (Hereditary but Not Stably Hereditary) Suppose X_n are a sequence of independent identically distributed random variables, equal to ± 1 with probability $\frac{1}{2}$. Form an array with row lengths $N+1$ by setting $X_n^{(N)} = X_n$, and let $f_n^{(N)}(X_n, X_{n+1}) := \frac{1}{2}\left(1 + \frac{1}{\sqrt[3]{N}}\right)X_n \ \ (1 \le n \le N+1)$. Then $D_N(\xi) \to \infty$ on $H(\mathsf{X},\mathsf{f})^c$, but the convergence is not uniform on compact subsets of $H(\mathsf{X},\mathsf{f})^c$.

Proof. $\Gamma\left(+1\,\begin{matrix}+1 & +1\\ +1 & -1\end{matrix}\,+1\right) = 1 + N^{-1/3}$. Since $\mathrm{Hex}(N,n)$ consists of 2^6 hexagons with equal probabilities, the hexagon $\left(+1\,\begin{matrix}+1 & +1\\ +1 & -1\end{matrix}\,+1\right)$ has probability 2^{-6}. Hence $d_n^{(N)}(\xi) \ge 2^{-6}|e^{i\xi(1+N^{-1/3})} - 1|^2 = \frac{1}{16}\sin^2\frac{\xi(1+N^{-1/3})}{2}$.

Therefore, $D_N(\xi) \ge \frac{N-2}{16}\sin^2\frac{\xi(1+N^{-1/3})}{2} \sim \begin{cases}16^{-1}N\sin^2\frac{\xi}{2} & \xi \notin 2\pi\mathbb{Z}\\ 64^{-1}\xi^2\sqrt[3]{N} & \xi \in 2\pi\mathbb{Z}.\end{cases}$ We see that $D_N(\xi) \to \infty$ for all $\xi \neq 0$, whence $H(\mathsf{X},\mathsf{f}) = \{0\}$, and $D_N(\xi) \to \infty$ on $H(\mathsf{X},\mathsf{f})^c$. But the convergence is not uniform on any compact neighborhood of $2\pi k$, $k \neq 0$, because $D_N(\xi_N) \equiv 0$ for $\xi_N = 2\pi k(1 + N^{-1/3})^{-1}$, and $\xi_N \to 2\pi k$. □

These examples raise the problem of deciding whether a given (X,f) is (stably) hereditary or not. We will discuss this now.

We begin with a simple class of examples, which will be important for us when we analyze the local limit theorem in the regime of large deviations:

Example 4.12 ("Change of Measure") Let Y be an array obtained from a uniformly elliptic Markov chain X using the change of measure construction (Example 2.6). Let $\varphi_n^{(N)}$ denote the weights of the change of measure. If for some constant $C > 0$,

$$C^{-1} < \varphi_n^{(N)} < C \text{ for all } n, N,$$

then for every a.s. uniformly bounded additive functional f on X, $f_n^{(N)} := f_n$ satisfies $H(\mathsf{Y},\mathsf{f}) = H(\mathsf{X},\mathsf{f})$, $G_{ess}(\mathsf{Y},\mathsf{f}) = G_{ess}(\mathsf{X},\mathsf{f})$, and (Y,f) is stably hereditary.

Proof. Let $m_{\mathrm{Hex}}^{N,n}$ be the hexagon measures of Y, and let m_{Hex}^{n} be the hexagon measures of X. It is not difficult to see that there is a constant C_* such that $C_*^{-1} \le \frac{dm_{\mathrm{Hex}}^{N,n}}{dm_{\mathrm{Hex}}^{n}} \le C_*$ for all $N \ge 0$ and $5 \le n \le N$, see Corollary 2.10 and its proof.

Thus $d_n(\xi,\mathsf{Y}) \asymp d_n(\xi,\mathsf{X})$ and $D_n(\xi,\mathsf{Y}) \asymp D_n(\xi,\mathsf{X}) + O(1)$. So $H(\mathsf{Y},\mathsf{f}) = H(\mathsf{X},\mathsf{f})$, and $G_{ess}(\mathsf{Y},\mathsf{f}) = G_{ess}(\mathsf{X},\mathsf{f})$. By Theorem 4.9, $D_N(\xi,\mathsf{X}) \to \infty$ uniformly on compacts in $H(\mathsf{X},\mathsf{f})^c$. So $D_N(\xi,\mathsf{X}) \to \infty$ uniformly on compacts in $H(\mathsf{Y},\mathsf{f})^c$. □

The hereditary property can be understood in terms of the behavior of sub-arrays. Let X be a Markov array with row lengths k_N. A **sub-array** of X is an array X' of the form $\{X_k^{(N_\ell)} : 1 \le k \le k_{N_\ell} + 1, \ell \ge 1\}$ where $N_\ell \uparrow \infty$. The **restriction** of an additive functional f on X to X' is $\mathsf{f}|_{\mathsf{X}'} = \{f_k^{(N_\ell)} : 1 \le k \le k_{N_\ell}, \ell \ge 1\}$.

Theorem 4.13 *Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array X. The following conditions are equivalent:*

(1) (X,f) is hereditary; (2) for all ξ, $\liminf_{N\to\infty} D_N(\xi) < \infty \Rightarrow \limsup_{N\to\infty} D_N(\xi) < \infty$;

(3) $H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) = H(\mathsf{X},\mathsf{f})$ for all sub-arrays X'; (4) $G_{ess}(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) = G_{ess}(\mathsf{X},\mathsf{f})$ for all sub-arrays X'.

The equivalence of (1) and (4) is the reason we call hereditary arrays "hereditary." Next we characterize the stably hereditary arrays:

Theorem 4.14 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X.

(1) If $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$*, then* (X,f) *is stably hereditary iff* $G_{ess}(\mathsf{X}',\mathsf{g}|_{\mathsf{X}'}) = \mathbb{R}$ *for all sub-arrays* X' *and all additive functionals* $\mathsf{g} = \{(1+\varepsilon_N)f_n^{(N)}\}$*, where* $\varepsilon_N \to 0$.
(2) If $G_{ess}(\mathsf{X},\mathsf{f}) \neq \mathbb{R}$*, then* (X,f) *is stably hereditary iff* (X,f) *is hereditary.*

For a hereditary array which is not stably hereditary, see Example 4.11.

4.3 Proofs

4.3.1 Reduction Lemma

Lemma 4.15 *Let* f *be an additive functional on a Markov array* X *with row lengths* $k_N + 1$*. For every* $N \geq 1$ *and* $1 \leq n \leq k_N$*, there exists a constant* $c_n^{(N)}$ *such that* $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in G_{alg}(\mathsf{X},\mathsf{f})$ *almost surely.*

Proof $G_{alg}(\mathsf{X},\mathsf{f})$ is the intersection of all closed subgroups G such that

$$\forall N\, \forall 1 \leq n \leq k_N, \exists c_n^{(N)} \text{ s.t. } f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in G \text{ almost surely.} \tag{4.3}$$

This is a closed subgroup of $\mathbb{R}$. The lemma is trivial when $G_{alg}(\mathsf{X},\mathsf{f}) = \mathbb{R}$ (take $c_n^{(N)} \equiv 0$), so we focus on the case $G_{alg}(\mathsf{X},\mathsf{f}) \neq \mathbb{R}$.

In this case (4.3) holds with some $G = t\mathbb{Z}$ with $t \geq 0$, and $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$ is a discrete random variable. Let $A_n^{(N)}$ denote the set of values attained by $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$ with positive probability. Since $G = t\mathbb{Z}$ satisfies (4.3), $A_n^{(N)}$ are subsets of cosets of $t\mathbb{Z}$, and $D_n^{(N)} := A_n^{(N)} - A_n^{(N)} \subset t\mathbb{Z}$. Let G_0 denote the group generated by $\bigcup_{N\geq 1}\bigcup_{1\leq n\leq k_N} D_n^{(N)}$. Then G_0 is a subgroup of $t\mathbb{Z}$. In particular, G_0 is closed.

By the previous paragraph, $G_0 \subset t\mathbb{Z}$ for any group $t\mathbb{Z}$ which satisfies (4.3). So $G_0 \subset G_{alg}(\mathsf{X},\mathsf{f})$. Next, we fix n, N and observe that all the values of $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})$ belong to the same translate of $A_n^{(N)} - A_n^{(N)}$, and therefore to the same coset of G_0. So G_0 satisfies (4.3), and $G_0 \supset G_{alg}(\mathsf{X},\mathsf{f})$. So $G_{alg}(\mathsf{X},\mathsf{f}) = G_0$. Since G_0 satisfies (4.3), $G_{alg}(\mathsf{X},\mathsf{f})$ satisfies (4.3). □

Lemma 4.16 (Reduction Lemma) *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X*, with row lengths* $k_N + 1$*. If* $\xi \neq 0$ *and* $\sup_N \sum_{k=3}^{k_N} d_k^{(N)}(\xi)^2 < \infty$*, then there is an additive functional* g *on* X *such that*

$$\mathsf{g} - \mathsf{f} \text{ is center-tight, } \operatorname{ess\,sup}|\mathsf{g}| < \infty, \text{ and } G_{alg}(\mathsf{X},\mathsf{g}) \subset \frac{2\pi}{\xi}\mathbb{Z}. \tag{4.4}$$

If $X_n^{(N)} = X_n$ *and* $f_n^{(N)} = f_n$ *(as in the case of additive functionals of Markov chains), then we can take* g *such that* $g_n^{(N)} = g_n$.

Proof for Doeblin Chains: As in the case of the gradient lemma, there is a simple proof in the important special case of Doeblin Markov chains (Example 2.7). Recall that Doeblin chains have finite state spaces $\mathfrak{S}_n$. Let $\pi_{xy}^n := \pi_{n,n+1}(x,\{y\})$, and relabel the states $\mathfrak{S}_n = \{1,\ldots,d_n\}$ in such a way that $\pi_{11}^n = \pi_{n,n+1}(1,\{1\}) \neq 0$ for all n. The Doeblin condition guarantees that for every $x \in \mathfrak{S}_n$, there exists a state $\xi_n(x) \in \mathfrak{S}_n$ such that $\pi_{1,\xi_n(x)}^{n-1}\pi_{\xi_n(x),1}^n > 0$.

Define, as in the proof of the gradient lemma,

$$a_1 \equiv 0, \quad a_2 \equiv 0, \quad \text{and } a_n(x) := f_{n-2}(1, \xi_{n-1}(x)) + f_{n-1}(\xi_{n-1}(x), x) \text{ for } n \geq 3$$
$$c_1 := 0, \quad c_2 := 0, \quad \text{and } c_n := f_{n-2}(1,1) \text{ for } n \geq 3 \quad \widetilde{\mathsf{f}} := \mathsf{f} - \nabla \mathsf{a} - \mathsf{c}.$$

Then $\widetilde{f_n}(x,y) = f_n(x,y) - (a_{n+1}(y) - a_n(x)) - c_n = -\Gamma_n \begin{pmatrix} 1 & 1 & \xi_n(y) & \\ & \xi_{n-1}(x) & x & y \end{pmatrix}$, where Γ_n denotes the balance of a position n hexagon, see (2.25).

For Doeblin chains, there are finitely many admissible hexagons at position n, and the hexagon measure assigns each of them a mass which is uniformly bounded from below. Let C^{-1} be a uniform lower bound for this mass, then $|e^{i\xi \widetilde{f_n}(x,y)} - 1|^2 \leq C\mathbb{E}_{m_{\text{Hex}}}(|e^{i\xi\Gamma_n} - 1|^2) = Cd_n^2(\xi)$.

Decompose $\widetilde{f_n}(x,y) = g_n(x,y) + h_n(x,y)$ where $g_n(x,y) \in \frac{2\pi}{\xi}\mathbb{Z}$ and $h_n(x,y) \in [-\frac{\pi}{\xi}, \frac{\pi}{\xi})$. Clearly $|\mathsf{g}| \leq |\mathsf{f}| + |\nabla a| + |\mathsf{c}| + |\mathsf{h}| \leq 6|\mathsf{f}| + \pi/\xi$, and $G_{alg}(\mathsf{X}, \mathsf{g}) \subset \frac{2\pi}{\xi}\mathbb{Z}$.

We show that $\mathsf{f} - \mathsf{g}$ is center-tight. We need the following inequality: [1]

$$\frac{4x^2}{\pi^2} \leq |e^{ix} - 1|^2 \leq x^2 \text{ for all } |x| \leq \pi. \tag{4.5}$$

By (4.5), $|h_n(x,y)|^2 \leq \frac{\pi^2}{4\xi^2}|e^{i\xi h_n(x,y)} - 1|^2 = \frac{\pi^2}{4\xi^2}|e^{i\xi\widetilde{f_n}(x,y)} - 1|^2 \leq C\frac{\pi^2}{4\xi^2}d_n^2(\xi)$, whence

$$\sum_{n=3}^{\infty} \mathrm{Var}(h_n(X_n, X_{n+1}) + c_n) = \sum_{n=3}^{\infty} \mathrm{Var}(h_n(X_n, X_{n+1})) \leq \frac{C\pi^2}{4\xi^2}\sum_{n=3}^{\infty} d_n^2(\xi) < \infty.$$

So $\mathsf{h} + \mathsf{c}$ has summable variance. Therefore $\mathsf{f} - \mathsf{g} = \nabla a + (\mathsf{h} + \mathsf{c})$ is center-tight. □

Preparations for the Proof in the General Case.

Lemma 4.17 *Suppose $E_1, \ldots, E_N$ are measurable events, and let W denote the random variable which counts how many of E_i occur simultaneously, then* $\mathbb{P}(W \geq t) \leq \dfrac{1}{t}\displaystyle\sum_{k=1}^{N} \mathbb{P}(E_k)$.

Proof Apply Markov's inequality to $W = \sum 1_{E_k}$. □

The expectation of an L^2 random variable W can be characterized as the constant $\mu \in \mathbb{R}$ which minimizes $\mathbb{E}(|W - \mu|^2)$. The variance is $\mathrm{Var}(W) = \min\limits_{\mu\in\mathbb{R}} \mathbb{E}(|W - \mu|^2)$.

Similarly, we define a **circular mean** of a random variable W to be any real number $\theta \in [-\pi, \pi)$ which minimizes the quantity $\mathbb{E}(|e^{i(W-\theta)} - 1|^2)$, and we define the **circular variance** to be the minimum

$$\mathrm{CVar}\,(W) := \min_{\theta\in[-\pi,\pi)} \mathbb{E}(|e^{i(W-\theta)} - 1|^2) \equiv \min_{\theta\in[-\pi,\pi)} 4\mathbb{E}(\sin^2 \tfrac{W-\theta}{2}).$$

Circular means always exist, but they are not always unique. Existence is because the function $\theta \mapsto \mathbb{E}(|e^{i(W-\theta)} - 1|^2)$ is continuous and 2π-periodic. Non-uniqueness can be seen, for example, when W is uniformly distributed on $[-\pi, \pi]$. In this case, every $\theta \in [-\pi, \pi)$ is a circular mean.

For every $x \in \mathbb{R}$, let

$$\langle x \rangle := \text{ the unique element of } [-\pi, \pi) \text{ s.t. } x - \langle x \rangle \in 2\pi\mathbb{Z}. \tag{4.6}$$

It is not difficult to see using (4.5), that for every circular mean θ,

$$\frac{4}{\pi^2}\mathrm{Var}\langle W - \theta \rangle \leq \mathrm{CVar}\,(W) \leq \mathrm{Var}(W). \tag{4.7}$$

[1] Proof of (4.5): $y = \sin\frac{x}{2}$ is concave on $[0, \pi]$, so its graph lies above the chord $y = \frac{x}{\pi}$ and below the tangent $y = \frac{x}{2}$. So $\frac{x}{\pi} \leq \sin\frac{x}{2} \leq \frac{x}{2}$ on $[0, \pi]$. Since $|e^{ix} - 1|^2 = 4\sin^2\frac{x}{2}$, we have (4.5).

Lemma 4.18 *For every real-valued random variable W, we can write $W = W_1 + W_2$ where $W_1 \in 2\pi\mathbb{Z}$ almost surely, and* $\mathrm{Var}(W_2) \le \frac{\pi^2}{4}\mathrm{CVar}\,(W)$.

Proof Take a circular mean θ, and let $W_1 := (W-\theta) - \langle W-\theta\rangle$, $W_2 := \langle W-\theta\rangle + \theta$. □

Proof of the Reduction Lemma in the General Case: Suppose f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array X, with row lengths k_N+1, and fix $\xi \neq 0$ such that $\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty$.

Let L denote the ladder process associated with X (see §2.3.3). Recall that this is a Markov array with entries $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$ $(3 \le n \le k_N)$ such that

(a) $\{X_n^{(N)}\}$, $\{Z_n^{(N)}\}$ are two independent copies of $\mathsf{X}^{(N)}$;
(b) $Y_n^{(N)}$ are conditionally independent given $\{X_i^{(N)}\}$ and $\{Z_i^{(N)}\}$; and
(c) the conditional distribution of $Y_n^{(N)}$ given $\{Z_i^{(N)}\}$ and $\{X_i^{(N)}\}$ is given by

$$\mathbb{P}\left(Y_{n-1}^{(N)} \in E \,\middle|\, \begin{matrix}\{Z_i^{(N)}\} = \{\zeta_i^{(N)}\}\\ \{X_i^{(N)}\} = \{\xi_i^{(N)}\}\end{matrix}\right) = \begin{matrix}\text{the bridge probability for } \mathsf{X} \text{ that } X_{n-1}^{(N)} \in E\\ \text{given that } X_{n-2}^{(N)} = \zeta_{n-2}^{(N)} \text{ and } X_n^{(N)} = \xi_n^{(N)}.\end{matrix}$$

Recall see (2.25), (4.6). We need the following (uniformly bounded) additive functionals on L:

$$F_n^{(N)}(\underline{L}_n^{(N)}) := f_{n-2}^{(N)}(Z_{n-2}^{(N)}, Y_{n-1}^{(N)}) + f_{n-1}^{(N)}(Y_{n-1}^{(N)}, X_n^{(N)});$$
$$\widehat{H}_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) := \left\langle \xi\Gamma\left(Z_{n-2}^{(N)} \;\begin{matrix}Z_{n-1}^{(N)}\\ Y_{n-1}^{(N)}\end{matrix}\; \begin{matrix}Y_n^{(N)}\\ X_n^{(N)}\end{matrix}\; X_{n+1}^{(N)}\right)\right\rangle, \quad H_n^{(N)} := \widehat{H}_n^{(N)} - \mathbb{E}[\widehat{H}_n^{(N)}].$$

Sometimes we will abuse notation, and write $\underline{L}_n = (Z_{n-2}, Y_{n-1}, X_n)$ and

$$F_n^{(N)}(\underline{L}_n^{(N)}) = F(\underline{L}_n)\ ,\ H_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)}) = H(\underline{L}_n, \underline{L}_{n+1})\ ,\ f_n^{(N)} = f_n. \tag{4.8}$$

By construction, $\widehat{H}_n \equiv \xi\Gamma\left(Z_{n-2} \;\begin{smallmatrix}Z_{n-1}\\ Y_{n-1}\end{smallmatrix}\; \begin{smallmatrix}Y_n\\ X_n\end{smallmatrix}\; X_{n+1}\right) \bmod 2\pi\mathbb{Z}$, and therefore,

$$\widehat{H}_n \equiv \xi f_{n-2}(Z_{n-2}, Z_{n-1}) + \xi\nabla\mathsf{F} - \xi f_n(X_n, X_{n+1}) \bmod 2\pi\mathbb{Z}.$$

Dividing by ξ and rearranging terms, we obtain the decomposition

$$f_n(X_n, X_{n+1}) = \frac{1}{\xi}\begin{bmatrix}\text{something taking}\\ \text{values in } 2\pi\mathbb{Z}\end{bmatrix} + \left[\nabla\mathsf{F} - \xi^{-1}\widehat{\mathsf{H}}\right] + f_{n-2}(Z_{n-2}, Z_{n-1}). \tag{4.9}$$

Step 1 below says that $\widehat{\mathsf{H}}$ has summable variance. So $\nabla\mathsf{F} - \xi^{-1}\widehat{\mathsf{H}}$ is center-tight.

By the structure of the ladder process, $f_{n-2}(Z_{n-2}, Z_{n-1})$ is independent from $f_n(X_n, X_{n+1})$. Fix some possible array of values ζ of Z. If we condition both sides of (4.9) on X_n, X_{n+1} and $\mathsf{Z} = \zeta$, then the left hand side remains $f_n(X_n, X_{n+1})$, but $f_{n-2}(Z_{n-2}, Z_{n-1})$ is replaced by the constant $f_{n-2}(\zeta_{n-2}, \zeta_{n-1})$.

The idea of the proof is to construct ζ so that the conditional expectation of the RHS on $\{X_n\}$ and $\mathsf{Z} = \zeta$ can still be put in the form $\frac{1}{\xi}\begin{bmatrix}\text{something taking}\\ \text{values in } 2\pi\mathbb{Z}\end{bmatrix} + \begin{bmatrix}\text{center-tight}\\ \text{(w.r.t. } \mathsf{X})\end{bmatrix} + f_{n-2}(\zeta_{n-2}, \zeta_{n-1})$.

The main difficulty is that the conditional expectation is an average, and the average of a $2\pi\mathbb{Z}$-valued quantity is not necessarily $2\pi\mathbb{Z}$-valued. We will address this difficulty by using "approximate conditional circular means," see Step 3.

Step 1: $\widehat{\mathsf{H}}$ *has summable variances. In addition,* $\mathbb{E}[(\widehat{H}_n^{(N)})^2] \le \frac{\pi^2}{4} d_n^{(N)}(\xi)^2$, $\mathbb{E}(H_n^{(N)}) = 0$, $\mathbb{E}[(H_n^{(N)})^2] \le \frac{\pi^2}{4} d_n^{(N)}(\xi)^2$, *and* $\sup_N \mathbb{E}[(H_3^{(N)} + \cdots + H_{k_N}^{(N)})^2] < \infty$.

Proof of the Step. We fix N and drop the superscripts $^{(N)}$.

- $\mathbb{E}(H_n) = \mathbb{E}(\widehat{H}_n) - \mathbb{E}(\widehat{H}_n) = 0$.
- By (4.5), $\mathbb{E}(\widehat{H}_n^2) \le \frac{\pi^2}{4}\mathbb{E}(|e^{i\widehat{H}_n} - 1|^2)$. By Lemma 2.22(d), $e^{i\widehat{H}_n}$ is equal in distribution to $e^{i\xi\Gamma(P)}$, where $\Gamma(P)$ is the balance of a random hexagon in Hex (N, n). So $\mathbb{E}(\widehat{H}_n^2) \le \dfrac{\pi^2}{4}\mathbb{E}_{m_{\text{Hex}}}(|e^{i\xi\Gamma(P)} - 1|^2) \equiv \dfrac{\pi^2}{4}d_n(\xi)^2$. By our assumptions, $\sup\limits_N \sum\limits_{n=3}^{k_N} d_n(\xi)^2 < \infty$. So $\widehat{\mathsf{H}}$ has summable variations.
- $\mathbb{E}(H_n^2) = \mathbb{E}(\widehat{H}_n^2) - \mathbb{E}(\widehat{H}_n)^2 \le \mathbb{E}(\widehat{H}_n^2) \le \frac{\pi^2}{4}d_n(\xi)^2$.
- L is uniformly elliptic, by Lemma 2.22. So Lemma 3.4 applies, and

$$\mathbb{E}\left[\left(\sum_{k=3}^{k_N} H_n\right)^2\right] = \text{Var}\left(\sum_{k=3}^{k_N} H_n\right) \le \text{const.}\sum_{k=3}^{k_N}\text{Var}(H_n) \le \text{const.}\sum_{k=3}^{k_N} d_n(\xi)^2 < \text{const.}$$

This completes the proof of Step 1.

Fix a constant D such that

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 + \sup_N \mathbb{E}\left[\left(\sum_{n=3}^{k_N} H_n^{(N)}\right)^2\right] + 4 < D. \tag{4.10}$$

Step 2 (Choice of ζ): $\exists\underline{\zeta}^{(N)} = (\zeta_1^{(N)}, \dots, \zeta_{k_N+1}^{(N)}) \in \mathfrak{S}_1^{(N)} \times \cdots \times \mathfrak{S}_{k_N+1}^{(N)}$ *s.t.*

$$\sum_{n=3}^{k_N}\mathbb{E}\Big(H_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)})^2\Big|\{Z_i^{(N)}\} = \underline{\zeta}^{(N)}\Big) < \pi^2 D, \quad \mathbb{E}\left[\left(\sum_{n=3}^{k_N} H_n^{(N)}(\underline{L}_n^{(N)}, \underline{L}_{n+1}^{(N)})\right)^2 \Big|\{Z_i^{(N)}\} = \underline{\zeta}^{(N)}\right] < \pi^2 D,$$

$$\mathbb{E}_X\left[\sum_{n=3}^{k_N}\text{CVar}\left(\xi F_n^{(N)}(\underline{L}_n^{(N)})\Big|\{Z_i^{(N)}\} = \underline{\zeta}^{(N)}, X_n^{(N)}\right)\right] < \pi^2 D, \quad |f_n^{(N)}(\zeta_n^{(N)}, \zeta_{n+1}^{(N)})| \le \text{ess sup}\,|f| \;\; \forall 3 \le n \le k_N.$$

(Here and throughout $\underline{L}_n^{(N)} = (Z_{n-2}^{(N)}, Y_{n-1}^{(N)}, X_n^{(N)})$*, and* $\mathbb{E}_X =$ *averaging on* $\{X_i^{(N)}\}$*).*

Proof of the Step. We fix N, drop the superscripts $^{(N)}$, and use convention (4.8). Let

$$\Omega_1 := \left\{\underline{\zeta} : \sum_{n=3}^{k_N}\mathbb{E}(H_n^2|\{Z_i\} = \underline{\zeta}) < \pi^2 D\right\}.$$

By Step 1,

$$\mathbb{E}_Z\left[\mathbb{E}\left(\sum_{n=3}^{k_N} H_n^2\Big|\{Z_i\} = \underline{\zeta}\right)\right] = \sum_{n=3}^{k_N}\mathbb{E}(H_n^2) \le \frac{\pi^2}{4}\sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 < \frac{\pi^2}{4}D,$$

where $\mathbb{E}_Z$ = integration over $\underline{\zeta}$ with respect to the distribution of $\{Z_i^{(N)}\}$ (recall that $\{Z_i^{(N)}\} \overset{\text{dist}}{=} \{X_i^{(N)}\}$). By Markov's inequality, $\mathbb{P}[\{Z_i^{(N)}\} \in \Omega_1] > \frac{3}{4}$.

Let $\Omega_2 := \left\{\underline{\zeta} : \mathbb{E}\left[\left(\sum_{n=3}^{k_N} H_n(\underline{L}_n, \underline{L}_{n+1})\right)^2 |\{Z_i\} = \underline{\zeta}\right] < \pi^2 D\right\}$. As before, by Markov's inequality, $\mathbb{P}[\{Z_i^{(N)}\} \in \Omega_2] \ge 1 - \frac{1}{\pi^2}$.

Let $\Omega_3 := \left\{\underline{\zeta} : \mathbb{E}_X\left[\sum\limits_{n=3}^{k_N}\text{CVar}\,(\xi F(\underline{L}_n)|\{Z_i\} = \underline{\zeta}, X_n)\right] < \pi^2 D\right\}$, and let

$$\theta^*(\underline{L}_n, X_{n+1}, Z_{n-1}) := -\xi f_{n-2}(Z_{n-2}, Z_{n-1}) + \xi F(\underline{L}_n) + \xi f_n(X_n, X_{n+1}).$$

Then $\exp[\mathrm{i}\widehat{H}_n(\underline{L}_n, \underline{L}_{n+1})] = \exp[\mathrm{i}\xi F(\underline{L}_{n+1}) - \mathrm{i}\theta^*(\underline{L}_n, X_{n+1}, Z_{n-1})]$.

Given X_{n+1} and $\{Z_i\}$, $\underline{L}_{n+1}$ is conditionally independent from $\underline{L}_n$, $\{X_i\}_{i\neq n+1}$. So

$$\begin{aligned}\mathbb{E}_{Z,X}\Big(\mathrm{CVar}\left(\xi F(\underline{L}_{n+1})|\{Z_i\}, X_{n+1}\right)\Big) &= \mathbb{E}\Big(\mathrm{CVar}\left(\xi F(\underline{L}_{n+1})|\underline{L}_n, \{Z_i\}, \{X_i\}\right)\Big)\\ &\overset{!}{\leq} \mathbb{E}\Big(\mathbb{E}(|\mathrm{e}^{\mathrm{i}\xi F(\underline{L}_{n+1})-\mathrm{i}\theta^*(\underline{L}_n, X_{n+1}, Z_{n-1})} - 1|^2|\underline{L}_n, \{X_i\}, \{Z_i\})\Big),\end{aligned}$$

where $\overset{!}{\leq}$ is because θ^* is conditionally constant, and by the definition CVar. Thus,

$$\mathbb{E}_{Z,X}(\mathrm{CVar}\,(\xi F(\underline{L}_{n+1})|\{Z_i\}, X_{n+1}))=\mathbb{E}(|\mathrm{e}^{\mathrm{i}(\xi F(\underline{L}_{n+1})-\theta^*)} - 1|^2) = \mathbb{E}(|\mathrm{e}^{\mathrm{i}\widehat{H}_n} - 1|^2)=\mathbb{E}_{m^{N,n}_{\mathrm{Hex}}}(|\mathrm{e}^{\mathrm{i}\xi\Gamma(P)} - 1|^2)=d_n(\xi)^2.$$

Summing over $3 \leq n \leq k_N - 1$ and adding the trivial bound 4 for $n = 2$, we obtain after shifting the index that $\mathbb{E}_Z\left[\mathbb{E}_X\left(\sum_{n=3}^{k_N}\mathrm{CVar}\,(\xi F(\underline{L}_n)|\{Z_i\}, X_n)\right)\right] < D$. By Markov's inequality, $\mathbb{P}(\{Z_i^{(N)}\} \in \Omega_3) \geq 1 - \frac{1}{\pi^2}$.

Finally, let $\Omega_4 := \{\underline{\zeta} : |f_n(\zeta_n, \zeta_{n+1})| \leq \text{ess sup}\,|\mathsf{f}|\}$, then $\mathbb{P}(\{Z_i^{(N)}\} \in \Omega_4) = 1$.

In summary $\mathbb{P}\left[\bigcup_{1\leq i\leq 4}\Omega_i^c\right] \leq \dfrac{2}{\pi^2} + \dfrac{1}{4} < 1$. Necessarily $\Omega_1 \cap \Omega_2 \cap \Omega_3 \cap \Omega_4 \neq \emptyset$. Any $\underline{\zeta} = \underline{\zeta}^{(N)}$ in the intersection satisfies the requirements of Step 2.

Step 3 (Choice of θ): *$\exists$ measurable functions $\theta_n^{(N)} : \mathfrak{S}_n^{(N)} \to [-\pi, \pi)$ such that*

$$\sum_{n=3}^{k_N}\mathbb{E}\Big(|\mathrm{e}^{\mathrm{i}\xi F_n^{(N)}(\underline{L}_n^{(N)})-\mathrm{i}\theta_n^{(N)}(X_n^{(N)})} - 1|^2\Big|\{Z_i^{(N)}\} = \underline{\zeta}^{(N)}\Big) < 2\pi^2 D.$$

Remark: As most of the summands must be small, many of the $\theta_n^{(N)}(X_n^{(N)})$ could be considered as "approximate circular means" of $\xi F(\underline{L}_n)$, given $\mathsf{Z} = \zeta, X_n$.

Proof of the Step. We fix N and drop the superscripts $^{(N)}$.

For every random variable W, $\theta \mapsto \mathbb{E}(|\mathrm{e}^{\mathrm{i}(W-\theta)} - 1|^2)$ is continuous, and therefore $\mathrm{CVar}\,(W) = \inf_{q\in\mathbb{Q}}\mathbb{E}(|\mathrm{e}^{\mathrm{i}(W-q)} - 1|^2)$.

In particular, $\mathrm{CVar}\,(\xi F_n|\{Z_i\} = \underline{\zeta}, X_n = \eta) = \inf_{q\in\mathbb{Q}}\mathbb{E}(|\mathrm{e}^{\mathrm{i}\xi F(\underline{L}_n)-\mathrm{i}q} - 1|^2|\{Z_i\} = \underline{\zeta}, X_n = \eta)$.

For each q, the expectation can be expressed explicitly using integrals with respect to bridge distributions, and its dependence on η is measurable. Passing to the infimum over $q \in \mathbb{Q}$, we find that $\eta \mapsto \mathrm{CVar}\,(\xi F_n|\{Z_i\} = \underline{\zeta}, X_n = \eta)$ is measurable.

Fix N and $\underline{\zeta} = \underline{\zeta}^{(N)}$. We say that $(\eta, q) \in \mathfrak{S}_n^{(N)} \times \mathbb{R}$ has "property $P_n(\eta, q)$," if

$$\mathbb{E}(|\mathrm{e}^{\mathrm{i}\xi F_n(\underline{L}_n)-\mathrm{i}q} - 1|^2|\{Z_n\} = \underline{\zeta}, X_n = \eta) \leq \mathrm{CVar}\,(\xi F_n(\underline{L}_n)|\{Z_n\} = \underline{\zeta}, X_n = \eta) + \frac{D}{n^2}.$$

By the previous paragraph, $\{\eta : P_n(\eta, q) \text{ holds}\}$ is measurable, and for every η there exists $q \in \mathbb{Q} \cap (-\pi, \pi)$ such that $P_n(\eta, q)$ holds. Let $\theta_n(\eta) = \theta_n^{(N)}(\eta) := \inf\{q : q \in \mathbb{Q} \cap (-\pi, \pi) \text{ s.t. } P_n(\eta, q) \text{ holds}\}$. Again, this is a measurable function, and since for fixed η, $P_n(\eta, q)$ is a closed property of q, $\theta_n^{(N)}(\eta)$ itself satisfies property $P_n(\eta, \theta_n^{(N)}(\xi))$. So, by choice of $\underline{\zeta}$

$$\mathbb{E}_X\left[\sum_{n=3}^{k_N}\mathbb{E}(|\mathrm{e}^{\mathrm{i}\xi F_n(\underline{L}_n)-\mathrm{i}\theta_n(X_n)} - 1|^2|\{Z_n\} = \underline{\zeta}, X_n)\right] \leq \mathbb{E}_X\left[\sum_{n=3}^{k_N}\mathrm{CVar}\left(\xi F_n(\underline{L}_n)\Big|\{Z_n\} = \underline{\zeta}, X_n\right)\right] + \frac{\pi^2}{6}D < 2\pi^2 D.$$

Step 4 (The Reduction). *Let $\underline{\zeta} = \underline{\zeta}^{(N)}$, $\theta_n = \theta_n^{(N)}$, $f_n = f_n^{(N)}$, $F_n = F_n^{(N)}$, $X_n = X_n^{(N)}$, $Z_n = Z_n^{(N)}$. Define*

$$c_n^{(N)} := f_{n-2}(\zeta_{n-2},\zeta_{n-1}) - \xi^{-1}\mathbb{E}(\widehat{H}_n^{(N)}),$$
$$a_n^{(N)}(x) := \frac{1}{\xi}\Big(\theta_n(x) + \mathbb{E}(\langle \xi F_n(\underline{L}_n) - \theta_n(X_n)\rangle | \{Z_i\} = \underline{\zeta}, X_n = x)\Big) \quad (x \in \mathfrak{S}_n^{(N)}),$$
$$\widetilde{\mathsf{f}} := \frac{1}{\xi}\Big\langle \xi(\mathsf{f} - \nabla\mathsf{a} - \mathsf{c})\Big\rangle, \quad \mathsf{g} := \mathsf{f} - \nabla\mathsf{a} - \mathsf{c} - \widetilde{\mathsf{f}}.$$

Then $\mathsf{a}, \mathsf{c}, \widetilde{\mathsf{f}}, \mathsf{g}$ *are uniformly bounded, and* $G_{alg}(\mathsf{g}) \subset \frac{2\pi}{\xi}\mathbb{Z}$*.*

Proof of the Step. By the choice of $\underline{\zeta}^{(N)}$, $|\mathsf{c}| \leq \text{ess sup}\,|\mathsf{f}| + \pi/|\xi|$, and by the definition of $\theta_n^{(N)}$ and $\langle\cdot\rangle$, $|\mathsf{a}| \leq 2\pi/|\xi|$ and $|\widetilde{\mathsf{f}}| \leq \pi/|\xi|$. So $|\mathsf{g}|$ is a.s. uniformly bounded. Next, $\mathsf{g} \equiv \dfrac{1}{\xi}(\xi(\mathsf{f} - \nabla\mathsf{a} - \mathsf{c}) - \langle \xi(\mathsf{f} - \nabla\mathsf{a} - \mathsf{c})\rangle)$.

By the definition of $\langle\cdot\rangle$, $G_{alg}(\mathsf{g}) \subset \frac{2\pi}{\xi}\mathbb{Z}$. Notice that $\mathsf{g} - \mathsf{f} = -\nabla\mathsf{a} - \mathsf{c} - \widetilde{\mathsf{f}}$. Gradients and constant functionals are center-tight, so to complete the proof of the reduction lemma, it remains to show:

Step 5: $\widetilde{\mathsf{f}}$ *is center-tight.*

Proof of the Step. We fix N, drop the superscripts $^{(N)}$, and use convention (4.8).

Let $\{Z_i\} = \{\zeta_i\}$, and $P_n := \begin{pmatrix} Z_{n-2} & \begin{smallmatrix} Z_{n-1} \\ Y_{n-1} \end{smallmatrix} & \begin{smallmatrix} Y_n \\ X_n \end{smallmatrix} & X_{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_{n-2} & \begin{smallmatrix} \zeta_{n-1} \\ Y_{n-1} \end{smallmatrix} & \begin{smallmatrix} Y_n \\ X_n \end{smallmatrix} & X_{n+1} \end{pmatrix}$. So

$$\begin{aligned} -\Gamma(P_n) &= -f_{n-2}(\zeta_{n-2},\zeta_{n-1}) - F_{n+1}(\underline{L}_{n+1}) + F_n(\underline{L}_n) + f_n(X_n, X_{n+1}) \\ &= \mathsf{f} - \mathsf{c} - \xi^{-1}\mathbb{E}(\widehat{H}_n) - \nabla\mathsf{F} = (\mathsf{f} - \nabla\mathsf{a} - \mathsf{c}) - \xi^{-1}\mathbb{E}(\widehat{H}_n) + \nabla(\mathsf{a} - \mathsf{F}). \end{aligned}$$

It follows that, conditioned on $\mathsf{Z} = \zeta$, we have the following equalities mod $2\pi\mathbb{Z}$:

$$\begin{aligned} \xi\widetilde{\mathsf{f}} &\equiv \xi(\mathsf{f} - \nabla\mathsf{a} - \mathsf{c}) \bmod 2\pi\mathbb{Z} \equiv -\xi\Gamma(P_n) + \mathbb{E}(\widehat{H}_n) + \xi\nabla(\mathsf{F} - \mathsf{a}) \bmod 2\pi\mathbb{Z} \\ &\equiv -(\widehat{H}_n - \mathbb{E}(\widehat{H}_n)) + \xi\nabla(\mathsf{F} - \mathsf{a}) \bmod 2\pi\mathbb{Z} \equiv -\mathsf{H} + \xi\nabla(\mathsf{F} - \mathsf{a}) \bmod 2\pi\mathbb{Z}. \end{aligned} \tag{4.11}$$

Define a new additive functional of L by $W_n^{(N)}(\underline{L}_n^{(N)}) := \langle \xi F(\underline{L}_n) - \theta_n(X_n)\rangle - \mathbb{E}\left(\langle \xi F(\underline{L}_n) - \theta_n(X_n)\rangle | \{Z_i\} = \underline{\zeta}, X_n\right)$.

By (4.11) and the definition of a, $\xi\widetilde{\mathsf{f}} \equiv -\mathsf{H} + \nabla\mathsf{W} \bmod 2\pi\mathbb{Z}$. Since $\xi\widetilde{\mathsf{f}}$ takes values in $[-\pi,\pi)$, $\langle \xi\widetilde{\mathsf{f}}\rangle = \xi\widetilde{\mathsf{f}}$, and so

$$\xi\widetilde{f}_n(X_n, X_{n+1}) = \langle W(\underline{L}_{n+1}) - W(\underline{L}_n) - H(\underline{L}_n, \underline{L}_{n+1})\rangle. \tag{4.12}$$

Claim. *Fix N, and given $\delta > 0$, let $T_\delta := 11\pi^2 D/\delta$. Then there exists a measurable set Ω_X of $\{X_i\}$ such that $\mathbb{P}(\Omega_X) > 1 - \delta$, and for all $\underline{\xi} \in \Omega_X$,*

(1) $\displaystyle\sum_{n=3}^{k_N} \mathbb{P}\Big(|W(\underline{L}_n)| > \frac{\pi}{3}\Big| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) < T_\delta,$

(2) $\displaystyle\sum_{n=3}^{k_N} \mathbb{P}\Big(|H(\underline{L}_n, \underline{L}_{n+1})| > \frac{\pi}{3}\Big| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) < T_\delta,$

(3) $\displaystyle\mathbb{E}\Big(\Big|\sum_{n=3}^{k_N} H(\underline{L}_n, \underline{L}_{n+1})\Big|\Big| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) < T_\delta.$

Proof of the Claim. $\underline{L}_n$ is conditionally independent of $\{X_i\}_{i\neq n}$ given $\{Z_i\}, X_n$. So

$$\begin{aligned} &\sum_{n=3}^{k_N} \mathbb{P}\Big(|W(\underline{L}_n)| \geq \frac{\pi}{4}\Big| \{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) = \sum_{n=3}^{k_N} \mathbb{P}\Big(|W(\underline{L}_n)| \geq \frac{\pi}{4}\Big| \{Z_i\} = \underline{\zeta}, X_n = \xi_n\Big) \\ &\leq \frac{16}{\pi^2}\sum_{n=3}^{k_N} \mathrm{Var}(\langle \xi F(\underline{L}_n) - \theta_n(X_n)\rangle | \{Z_i\} = \underline{\zeta}, X_n) \leq \frac{16}{\pi^2}\sum_{n=3}^{k_N} \mathbb{E}(\langle \xi F(\underline{L}_n) - \theta_n(X_n)\rangle^2 | \{Z_i\} = \underline{\zeta}, X_n) \end{aligned}$$

$$\le 4 \sum_{n=3}^{k_N} \mathbb{E}\big(|e^{i\xi F(\underline{L}_n) - i\theta_n(X_n)} - 1|^2 |\{Z_i\} = \underline{\zeta}, X_n\big), \text{ see (4.5).}$$

Integrating over $\{X_i\}$, and recalling the choice of $\theta_n^{(N)}(X_n)$ (Step 3), we find that

$$\mathbb{E}_X\left[\sum_{n=3}^{k_N} \mathbb{P}\Big(|W(\underline{L}_n)| \ge \frac{\pi}{4}\Big|\{Z_i\} = \underline{\zeta}, \{X_i\}\Big)\right] < 8\pi^2 D.$$

Hence by Markov's inequality, the set $\Omega_X^1(T) := \left\{\underline{\xi} : \sum_{n=3}^{k_N} \mathbb{P}\Big(|W(\underline{L}_n)| > \frac{\pi}{3}\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) \le T\right\}$ has probability $\mathbb{P}[\Omega_X^1(T)] > 1 - 8\pi^2 D/T$. Similarly, by Markov's inequality

$$\mathbb{P}\Big(|H_n| \ge \frac{\pi}{4}\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) \le \frac{16}{\pi^2}\mathbb{E}\Big(H_n^2\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big).$$

By the choice of $\underline{\zeta}$, $\mathbb{E}_X\left[\sum_{n=3}^{k_N} \mathbb{P}\left(|H_n| \ge \frac{\pi}{4}|\{Z_i\} = \underline{\zeta}, \{X_i\}\right)\right] < 16D.$

So the set $\Omega_X^2(T) := \left\{\underline{\xi} : \sum_{n=3}^{k_N} \mathbb{P}\Big(|H(\underline{L}_n, \underline{L}_{n+1})| > \frac{\pi}{3}\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) \le T\right\}$ has probability $\mathbb{P}[\Omega_X^2(T)] \ge 1 - 16D/T > 1 - 2\pi^2 D/T$.

Finally, conditional expectations contract L^2-norms, therefore

$$\mathbb{E}\Big[\mathbb{E}\Big(|\sum_{n=3}^{k_N} H(\underline{L}_n, \underline{L}_{n+1})|\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big)^2\Big|\{Z_i\} = \underline{\zeta}\Big]$$
$$\le \mathbb{E}\Big[\Big(\sum_{n=3}^{k_N} H(\underline{L}_n, \underline{L}_{n+1})\Big)^2\Big|\{Z_i\} = \underline{\zeta}\Big] \le \pi^2 D, \text{ see step 3.}$$

So

$$\Omega_X^3(T) := \left\{\underline{\xi} : \mathbb{E}\Big(|\sum_{n=3}^{k_N} H(\underline{L}_n, \underline{L}_{n+1})|\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) \le T\right\}$$

has probability $\mathbb{P}[\Omega_X^3(T)] > 1 - \pi^2 D/T^2$.

Then $\mathbb{P}[\Omega_X^1(T) \cap \Omega_X^2(T) \cap \Omega_X^3(T)] > 1 - \frac{11\pi^2 D}{T}$, and the claim follows.

We can now complete the proof of the Step 5 and show that $\widetilde{\mathsf{f}}$ is center-tight. Fix $\delta > 0$, Ω_X and T_δ as in the claim. Fix N and define the random set

$$A_N(\{\underline{L}_n^{(N)}\}) := \left\{3 \le n \le k_N : |W(\underline{L}_n)| \ge \frac{\pi}{3} \text{ or } |H_n(\underline{L}_n, \underline{L}_{n+1})| \ge \frac{\pi}{3}\right\}.$$

For all $\underline{\xi} \in \Omega_X$, we have the following bound by Lemma 4.17:

$$\mathbb{P}\Big(|A_N| > 4T_\delta\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) < \frac{1}{2}.$$

In addition, for all $\underline{\xi} \in \Omega_X$, $\mathbb{P}\Big(|\sum_{n=3}^{k_N} H_n| > 4T_\delta\Big|\{Z_i\} = \underline{\zeta}, \{X_i\} = \underline{\xi}\Big) \le \frac{1}{4}$.

The probabilities of these events add up to less than one, so the intersection of their complements is non-empty. Thus, for every $\underline{\xi} \in \Omega_X$, we can find $\{Y_i^{(N)}(\underline{\xi})\}_{i=2}^{k_N - 1}$ such that

$$\underline{L}_n^* := \underline{L}_n^*(\underline{\xi}) = (\zeta_{n-2}^{(N)}, Y_{n-1}^{(N)}(\underline{\xi}), \xi_n)$$

has the following two properties:

- $\left|\sum_{n=3}^{k_N} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)\right| \leq 4T_\delta$, and
- $M := \#\left\{3 \leq n \leq k_N : |W(\underline{L}_n^*)| \geq \dfrac{\pi}{3} \text{ or } |H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)| \geq \dfrac{\pi}{3}\right\} \leq 4T_\delta.$

Let $n_1 < \cdots < n_M$ be an enumeration of the indices n where $|W(\underline{L}_n^*)| \geq \frac{\pi}{3}$, or $|H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)| \geq \frac{\pi}{3}$. By (4.12), if $n_i < n < n_{i+1} - 1$,

$$\xi \widetilde{f}_n(\xi_n, \xi_{n+1}) = W(\underline{L}_{n+1}^*) - W(\underline{L}_n^*) - H_n(\underline{L}_n^*, \underline{L}_{n+1}^*),$$

because $\langle x + y + z \rangle = x + y + z$ whenever $|x|, |y|, |z| < \frac{\pi}{3}$.

So $\sum_{n=n_i}^{n_{i+1}-1} \xi \widetilde{f}_n(\xi_n, \xi_{n+1}) = -\sum_{n=n_i+1}^{n_{i+1}-1} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*) \pm 6\pi$, where we have used the bounds $|\mathsf{W}| \leq 2\pi$ and $|\xi \widetilde{f}_{n_i}| \leq \pi$. Summing over i we find that for every $\underline{\xi} \in \Omega_X$,

$$\left|\xi \sum_{n=3}^{k_N} \widetilde{f}_n(\xi_n, \xi_{n+1})\right| \leq \left|\sum_{n=3}^{k_N} H_n(\underline{L}_n^*, \underline{L}_{n+1}^*)\right| + 10M\pi \leq 4T_\delta + 40T_\delta\pi < 42\pi T_\delta.$$

Setting $C_\delta := 42\pi T_\delta / |\xi|$, we find that

$$\mathbb{P}\left(|\sum_{n=3}^{k_N} \widetilde{f}_n^{(N)}| \geq C_\delta\right) < 1 - \mathbb{P}(\Omega_X) < \delta$$

for all N, whence the (center-)tightness of $\widetilde{\mathsf{f}}$.

This proves Step 5. The Lemma follows from Steps 4 and 5. □

4.3.2 Joint Reduction

The gradient lemma in §3.2.1 modifies an additive functional by a gradient, to make the sums of the variances of its first k_N terms comparable to U_N.

The reduction lemma modifies an additive functional by a gradient, so that it can be split into the sum of a center-tight functional, and a $\dfrac{2\pi}{\xi}$-valued functional. But the sum of the variances of the $\dfrac{2\pi}{\xi}$-valued functional may be much larger than U_N.

The goal of this section is to achieve a joint reduction, so that the sum of the variances of the $\dfrac{2\pi}{\xi}$-valued functional is of order U_N. This result will be used in Chapter 6.

Lemma 4.19 (Integer Reduction Lemma) *Let* X *be a uniformly elliptic Markov chain, and* f *an integer-valued additive functional on* X *such that* $|f| \leq K$ *a.s. For every* N,

$$f_n(x, y) = g_n^{(N)}(x, y) + a_{n+1}^{(N)}(x) - a_n^{(N)}(y) + c_n^{(N)} \quad (n = 3, \ldots, N)$$

where

(1) $c_n^{(N)}$ *are integers such that* $|c_n^{(N)}| \leq K$,
(2) $a_n^{(N)}$ *are measurable integer-valued functions on* $\mathfrak{S}_n$ *such that* $|a_n^{(N)}| \leq 2K$,

(3) $g_n^{(N)}$ *are measurable,* $\mathbb{Z}$*-valued, and* $\sum_{n=3}^{N} \mathbb{E}[g_n^{(N)}(X_n, X_{n+1})^2] \le 10^3 K^4 \sum_{n=3}^{N} u_n^2$*, where* u_n *are the structure constants of* f.

Proof Let $\left(Z_{n-2} \begin{smallmatrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{smallmatrix} X_{n+1}\right)$ be a random hexagon. By the definition of the structure constants,

$$\mathbb{E}\left[\sum_{n=3}^{N} \mathbb{E}\left(\Gamma\left(Z_{n-2} \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix} X_{n+1}\right)^2 \middle| Z_{n-2}, Z_{n-1}\right)\right] = \sum_{n=3}^{N} u_n^2.$$

Therefore, for every N there exist $z_n = z_n(N) \in \mathfrak{S}_n$ $(n = 1, \dots, N-2)$ such that

$$\sum_{n=3}^{N} \mathbb{E}\left(\Gamma\left(Z_{n-2} \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix} X_{n+1}\right)^2 \middle| Z_{n-2} = z_{n-2}, Z_{n-1} = z_{n-1}\right) \le \sum_{n=3}^{N} u_n^2.$$

We emphasize that z_n depends on N.

Let $c_n^{(N)} := f_{n-2}(z_{n-2}, z_{n-1})$. Let $a_n^{(N)}(x_n)$ be the (smallest) most likely value of

$$f_{n-2}(z_{n-2}, Y) + f_{n-1}(Y, x_n),$$

where Y has the bridge distribution of X_{n-1} conditioned on $X_{n-2} = z_{n-2}$ and $X_n = x_n$. The most likely value exists, and has probability bigger than $\delta_K := \frac{1}{5K}$, because $f_{n-2}(z_{n-2}, Y) + f_{n-1}(Y, x_n) \in [-2K, 2K] \cap \mathbb{Z}$.

Set $g_n^{(N)}(x_n, x_{n+1}) := f_n(x_n, x_{n+1}) + a_n^{(N)}(x_n) - a_{n+1}^{(N)}(x_{n+1}) - c_n^{(N)}$. Equivalently,

$$g_n^{(N)}(x_n, x_{n+1}) = -\Gamma\left(z_{n-2} \begin{matrix} z_{n-1} & y_n \\ y_{n-1} & x_n \end{matrix} x_{n+1}\right)$$

for the y_k which maximize the likelihood of the value $f_{k-1}(z_{k-1}, Y) + f_k(Y, x_{k+1})$, when Y has the bridge distribution of X_k given $X_{k-1} = z_{k-1}, X_{k+1} = x_{k+1}$.

Our task is to estimate $\sum_{n=3}^{N} \mathbb{E}[g_n^{(N)}(X_n, X_{n+1})^2]$. Define for this purpose the functions $h_n^{(N)} : \mathfrak{S}_n \times \mathfrak{S}_{n+1} \to \mathbb{R}$,

$$h_n^{(N)}(x_n, x_{n+1}) := \mathbb{E}\left[\Gamma\left(Z_{n-2} \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix} X_{n+1}\right)^2 \middle| \begin{matrix} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{matrix}\right]^{1/2}.$$

The plan is to show that:

(a) $\sum_{n=3}^{N} \mathbb{E}(h_n^{(N)}(X_n, X_{n+1})^2) \le \sum_{n=3}^{N} u_n^2$,

(b) If $h_n^{(N)}(x_n, x_{n+1}) < \delta_K$, then $g_n^{(N)}(x_n, x_{n+1}) = 0$,

(c) $\mathbb{E}(g_n^{(N)}(X_n, X_{n+1})^2) \le (6K)^2 \mathbb{P}[h_n^{(K)} \ge \delta_K] \le 36K^2 \delta_K^{-2} \mathbb{E}[h_n^{(N)}(X_n, X_{n+1})^2]$.

Part (a) is because of the choice of z_n. Part (c) follows from part (b), Chebyshev's inequality, and the estimate $\|g_n^{(N)}\|_\infty \le 6K$ (as is true for the balance of every hexagon). It remains to prove part (b).

Since f is integer-valued, either the balance of a hexagon is zero, or it has absolute value ≥ 1. This leads to the following inequality.

$$\mathbb{P}\left[\Gamma\left(Z_{n-2} \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix} X_{n+1}\right) \neq 0 \middle| \begin{matrix} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{matrix}\right]$$
$$\le \mathbb{E}\left[\Gamma\left(Z_{n-2} \begin{matrix} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{matrix} X_{n+1}\right)^2 \middle| \begin{matrix} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{matrix}\right] = h_n^{(N)}(x_n, x_{n+1})^2.$$

Thus, if $h_n^{(N)}(x_n, x_{n+1}) < \delta_K$, then

$$\mathbb{P}\left[\Gamma\left(Z_{n-2} \begin{array}{cc} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{array} X_{n+1}\right) = 0 \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array}\right] > 1 - \delta_K^2.$$

At the same time, by the structure of the hexagon measure, if

$$\Omega_n := \left\{\left(Z_{n-2} \begin{array}{cc} Z_{n-1} & Y_n \\ Y_{n-1} & X_n \end{array} X_{n+1}\right) : \begin{array}{l} f_{n-1}(Z_{n-1}, Y_n) + f_n(Y_n, X_{n+1}) = a_{n+1}^{(N)}(X_{n+1}) \\ f_{n-2}(Z_{n-2}, Y_{n-1}) + f_{n-1}(Y_{n-1}, X_n) = a_n^{(N)}(X_n) \end{array}\right\},$$

then $\mathbb{P}\left[\Omega_n \middle| \begin{array}{ll} Z_{n-2} = z_{n-2} & Z_{n-1} = z_{n-1} \\ X_n = x_n & X_{n+1} = x_{n+1} \end{array}\right] > \delta_K^2$.

If the sum of the probabilities of two events is bigger than one, then they must intersect. It follows that there exist y_{n-1}, y_n such that

- $a_n^{(N)}(x_n) = f_{n-2}(z_{n-2}, y_{n-1}) + f_{n-1}(y_{n-1}, x_n)$;
- $a_{n+1}^{(N)}(x_{n+1}) = f_{n-1}(z_{n-1}, y_n) + f_n(y_n, x_{n+1})$;
- $\Gamma\left(z_{n-2} \begin{array}{cc} z_{n-1} & y_n \\ y_{n-1} & x_n \end{array} x_{n+1}\right) = 0.$

By the definition of $g_n^{(N)}$, $g_n^{(N)}(x_n, x_{n+1}) = 0$, which proves part (b). □

Corollary 4.20 (Joint Reduction) *Let* f *be an additive functional of a uniformly elliptic Markov array* X *with row lengths* k_N*, such that* $|\mathsf{f}| \le K$ *a.s. If* $\xi \neq 0$ *and* $\sup_N U_N < \infty$*, then there is an additive functional* g *satisfying* (4.4)*, and* $\sum_{n=3}^{k_N} \|g_n^{(N)}\|_2^2 \le L U_N$*, with a constant* L *which only depends on* K *and* ξ*.*

Caution! $g_n^{(N)}$ depends on N, even if f is an additive functional of a Markov chain.

Proof Apply Lemma 4.16 to f, and then apply Lemma 4.19 to the resulting integer-valued additive functional $\frac{\xi \mathsf{g}}{2\pi}$. □

4.3.3 The Possible Values of the Co-range

We prove Theorem 4.3 in its version for Markov arrays: *The co-range of an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X *is equal to* $\mathbb{R}$ *when* f *is center-tight, and to* $\{0\}$ *or* $t\mathbb{Z}$ *with* $t \ge \pi/(6\text{ess sup}\,|\mathsf{f}|)$ *otherwise.*

Recall that the co-range is defined by

$$H := H(\mathsf{X}, \mathsf{f}) = \{\xi \in \mathbb{R} : \sup_N D_N(\xi) < \infty\}, \text{ where } D_N(\xi) = \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2.$$

Step 1. *H is a subgroup of $\mathbb{R}$.*

Proof of the Step. $H = -H$, because $d_n^{(N)}(-\xi) = d_n^{(N)}(\xi)$. H contains 0, because $d_n^{(N)}(0) = 0$. H is closed under addition, because if $\xi, \eta \in H$, then by Lemma 2.16,

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi + \eta)^2 \le 8\left[\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 + \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\eta)^2\right] < \infty.$$

Step 2. *If* f *is center-tight, then* $H = \mathbb{R}$.

Proof of the Step. By Corollary 3.9 and the center-tightness of f, $\sup_N \sum_{k=3}^{k_N} (u_k^{(N)})^2 < \infty$. By Lemma 2.16(c), $\sup_N \sum_{k=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty$ for all $\xi \in \mathbb{R}$.

Step 3. *If f is not center-tight, then $\exists t_0 > 0$ such that*

$$H \cap (-t_0, t_0) = \{0\}. \tag{4.13}$$

Proof of the Step. Let $K := \text{ess sup}\,|\mathsf{f}|$, then $|\Gamma(P)| \le 6K$ for a.e. hexagon P.

Fix $\tau_0 > 0$ such that $|e^{it} - 1|^2 \ge \frac{1}{2}t^2$ for all $|t| < \tau_0$, and let $t_0 := \tau_0(6K)^{-1}$. Then for all $|\xi| < t_0$, $|e^{i\xi\Gamma(P)} - 1|^2 \ge \frac{1}{2}\xi^2\Gamma(P)^2$ for all hexagons P. Taking the expectation over $P \in \text{Hex}(N, n)$, we obtain that

$$d_n^{(N)}(\xi)^2 \ge \frac{1}{2}\xi^2(u_n^{(N)})^2 \text{ for all } |\xi| < t_0, 1 \le n \le k_N, N \ge 1. \tag{4.14}$$

Now assume by way of contradiction that there is $0 \ne \xi \in H \cap (-t_0, t_0)$, then $\sup_N \sum_{n=3}^{k_N} (u_n^{(N)})^2 \le \frac{2}{\xi^2} \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 < \infty$. By Corollary 3.9, f is center-tight, in contradiction to our assumption.

Step 4. *If f is not center-tight, then $H = \{0\}$, or $H = t\mathbb{Z}$ with $t \ge \dfrac{\pi}{3\,\text{ess sup}\,|f|}$.*

Proof of the Step. By Steps 2 and 3, H is a proper closed subgroup of $\mathbb{R}$. So it must be equal to $\{0\}$ or $t\mathbb{Z}$, where $t > 0$. To see that $t \ge \dfrac{\pi}{3\,\text{ess sup}\,|f|}$, assume by contradiction that $t = \left(\dfrac{\pi}{3\,\text{ess sup}\,|f|}\right)\rho$ with $0 < \rho < 1$, and let

$$\kappa := \min\{|e^{i\gamma} - 1|^2/|\gamma|^2 : |\gamma| \le 2\pi\rho\} > 0.$$

Then $|t\Gamma(P)| \le 6t\,\text{ess sup}\,|f| = 2\pi\rho$ for every hexagon $P \in \text{Hex}\,(N, n)$, whence

$$d_n^{(N)}(t)^2 = \mathbb{E}_{m_{\text{Hex}}}(|e^{it\Gamma} - 1|^2) \ge \kappa\mathbb{E}_{m_{\text{Hex}}}(t^2\Gamma^2) = \kappa t^2(u_n^{(N)})^2.$$

So $D_N(\xi) \ge \kappa t^2 U_N$. But this is impossible: $t \in H$ so $\sup_N D_N < \infty$, whereas f is not center-tight so by Corollary 3.9, $\sup_N U_N = \infty$. □

4.3.4 Calculation of the Essential Range

We prove Theorem 4.4 in its version for Markov arrays: *For every a.s. uniformly bounded additive functional* f *on a uniformly elliptic Markov array* X, $G_{ess}(\mathsf{X}, \mathsf{f}) = \{0\}$ *when* $H(\mathsf{X}, \mathsf{f}) = \mathbb{R}$*;* $G_{ess}(\mathsf{X}, \mathsf{f}) = \frac{2\pi}{t}\mathbb{Z}$ *when* $H(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$*; and* $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$ *when* $H(\mathsf{X}, \mathsf{f}) = \{0\}$.

Lemma 4.21 *Suppose* f, g *are two a.s. uniformly bounded additive functionals on the same uniformly elliptic Markov array. If* f − g *is center-tight, then* f *and* g *have the same co-range.*

Proof By Corollary 3.9, if h = g − f is center-tight, then

$$\sup_N \sum_{n=3}^{k_N} u_n^{(N)}(\mathsf{h})^2 < \infty.$$

By Lemma 2.16,

$$\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, \mathsf{g})^2 \leq 8 \sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, \mathsf{f})^2 + 8\xi^2 \sup_N \sum_{n=3}^{k_N} u_n^{(N)}(\mathsf{h})^2.$$

So the co-range of f is a subset of the co-range of g. By symmetry they are equal. □

Proof of Theorem 4.4: As we saw in the previous section, the possibilities for the co-range are $\mathbb{R}$, $t\mathbb{Z}$ with $t \neq 0$, and $\{0\}$.

CASE 1: *The co-range is* $\mathbb{R}$. By Theorem 4.3, this can only happen if f is center-tight, in which case the essential range is $\{0\}$ because we may subtract f from itself.

CASE 2: *The co-range is* $t\mathbb{Z}$ *with* $t \neq 0$. We show that $G_{ess}(\mathsf{X}, \mathsf{f}) = \frac{2\pi}{t}\mathbb{Z}$.

By assumption, t is in the co-range, so $\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(t)^2 < \infty$. By the reduction lemma, f differs by a center-tight functional from a functional with algebraic range inside $\frac{2\pi}{t}\mathbb{Z}$. So

$$G_{ess}(\mathsf{X}, \mathsf{f}) \subset \frac{2\pi}{t}\mathbb{Z}.$$

Assume by way of contradiction that $G_{ess}(\mathsf{X}, \mathsf{f}) \neq \frac{2\pi}{t}\mathbb{Z}$, then there exists a center-tight h such that the algebraic range of

$$\mathsf{g} := \mathsf{f} - \mathsf{h}$$

is a subset of $\frac{2\pi\ell}{t}\mathbb{Z}$ for some integer $\ell > 1$. The structure constants of g must satisfy $d_n^{(N)}(\frac{t}{\ell}, \mathsf{g}) \equiv 0$, and therefore $\frac{t}{\ell}$ is in the co-range of g. By Lemma 4.21, $\frac{t}{\ell}$ is in the co-range of f, whence $\frac{t}{\ell} \in t\mathbb{Z}$. But this contradicts $\ell > 1$.

CASE 3: *The co-range is* $\{0\}$. We claim that the essential range is $\mathbb{R}$. Otherwise, there exists a center-tight h such that the algebraic range of $\mathsf{g} := \mathsf{f} - \mathsf{h}$ equals $t\mathbb{Z}$ with $t \in \mathbb{R}$. But this is impossible:

(a) If $t \neq 0$, then $d_n^{(N)}(\frac{2\pi}{t}, \mathsf{g}) = 0$ for all $3 \leq n \leq k_N$, $N \geq 1$, so the co-range of g contains $2\pi/t$. By Lemma 4.21, the co-range of f contains $2\pi/t$, in contradiction to the assumption that it is $\{0\}$.

(b) If $t = 0$, then the algebraic range of g is $\{0\}$, and by Lemma 4.15, the entries of g are all a.s. constant. So $\mathsf{f} \equiv \mathsf{h} + \mathsf{g}$ is center-tight, and by Theorem 4.3, the co-range of f is $\mathbb{R}$. But this contradicts our assumptions. □

4.3.5 Existence of Irreducible Reductions

We prove Theorem 4.5, in its version for arrays: *For every a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X*, there exists an irreducible functional* g *such that* f − g *is center-tight and* $G_{alg}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{f})$.

Proof. The essential range is a closed subgroup of $\mathbb{R}$, so $G_{ess}(\mathsf{X}, \mathsf{f}) = \{0\}, t\mathbb{Z}$ or $\mathbb{R}$.

(a) If $G_{ess}(\mathsf{X}, \mathsf{f}) = \{0\}$, then $H(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, and f is center-tight. So take $\mathsf{g} \equiv 0$.

(b) If $G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ with $t \neq 0$, then $H(\mathsf{X}, \mathsf{f}) = \xi\mathbb{Z}$ with $\xi := 2\pi/t$ (Theorem 4.4). So $\sup_N \sum_{n=3}^{k_N} d_n^{(N)}(\xi, \mathsf{f})^2 < \infty$. By the reduction lemma, there exists an additive functional g such that f − g is center-tight, and $G_{alg}(\mathsf{X}, \mathsf{g}) \subset t\mathbb{Z}$. Clearly, two additive functionals which differ by a center-tight functional have the same essential range. So $G_{ess}(\mathsf{X}, \mathsf{f}) = G_{ess}(\mathsf{X}, \mathsf{g}) \subset G_{alg}(\mathsf{X}, \mathsf{g}) \subset t\mathbb{Z} = G_{ess}(\mathsf{X}, \mathsf{f})$, and $G_{ess}(\mathsf{X}, \mathsf{f}) = G_{ess}(\mathsf{X}, \mathsf{g}) = G_{alg}(\mathsf{X}, \mathsf{g})$.

(c) If $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, take $\mathsf{g} := \mathsf{f}$. □

4.3.6 Characterization of Hereditary Additive Functionals

Proof of Theorem 4.13: Suppose f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array X, and consider the following conditions:

(1) f is hereditary: for all $\xi \notin H(\mathsf{X},\mathsf{f})$, $D_N(\xi) \xrightarrow[N\to\infty]{} \infty$;
(2) For all ξ, $(\liminf_{N\to\infty} D_N(\xi) < \infty \Rightarrow \limsup_{N\to\infty} D_N(\xi) < \infty)$;
(3) $H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) = H(\mathsf{X},\mathsf{f})$ for every sub-array X' of X;
(4) $G_{ess}(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) = G_{ess}(\mathsf{X},\mathsf{f})$ for every sub-array X' of X.

(1)⇒(2): Let $L_{\inf}(\xi) := \liminf D_N(\xi)$, and $L_{\sup}(\xi) := \limsup D_N(\xi)$. Assume (1), and suppose $L_{\inf}(\xi) < \infty$. Then $L_{\sup}(\xi) < \infty$, otherwise $\sup D_N(\xi) = \infty$, whence $\xi \notin H(\mathsf{X},\mathsf{f})$, whence by (1), $\lim D_N(\xi) = \infty$. But this contradicts $L_{\inf}(\xi) < \infty$.

(2)⇒(3): If $\xi \in H(\mathsf{X},\mathsf{f})$, then $\sup D_N(\xi,\mathsf{f}|_{\mathsf{X}'}) \le \sup D_N(\xi,\mathsf{f}) < \infty$, and $\xi \in H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'})$.

Conversely, if $\xi \in H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'})$, then $\liminf D_N(\xi) < \infty$, whence by (2) $\limsup D_N(\xi) < \infty$. Therefore $\sup D_N(\xi) < \infty$, and $\xi \in H(\mathsf{X},\mathsf{f})$.

(3)⇒(4) because the co-range determines the essential range (Theorem 4.4).

(4)⇒(1): Suppose $\xi \notin H(\mathsf{X},\mathsf{f})$, and assume by contradiction that $D_N(\xi) \not\to \infty$. Then $\exists N_\ell \uparrow \infty$ such that $\sup_\ell D_{N_\ell}(\xi) < \infty$. But this means that $\xi \in H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'})$ for the sub-array $\mathsf{X}' = \{X_n^{(N_\ell)} : 1 \le n \le k_{N_\ell} + 1, \ell \ge 1\}$, and we found a sub-array such that $H(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) \ne H(\mathsf{X},\mathsf{f})$. By Theorem 4.4, $G_{ess}(\mathsf{X}',\mathsf{f}|_{\mathsf{X}'}) \ne G_{ess}(\mathsf{X},\mathsf{f})$. □

Proof of Theorem 4.14: Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array X.

The first part of the theorem assumes that $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, and asserts that (X,f) is stably hereditary iff $G_{ess}(\mathsf{X}',\mathsf{g}|_{\mathsf{X}'}) = \mathbb{R}$ for all sub-arrays X', and all additive functionals $\mathsf{g} = \{(1+\varepsilon_N) f_n^{(N)}\}$, where $\varepsilon_N \to 0$.

(⇐): Taking $\varepsilon_N = 0$, and applying Theorem 4.13, we find that (X,f) is hereditary. So $D_N(\xi) \to \infty$ for all $\xi \ne 0$. To show that the convergence is uniform on compact subsets of $\mathbb{R} \setminus \{0\}$, it is sufficient to check that

$$\forall \xi \ne 0, \forall M > 0, \exists N_\xi, \delta_\xi > 0 \left(\begin{matrix} N > N_\xi \\ |\xi' - \xi| < \delta_\xi \end{matrix} \Rightarrow D_N(\xi') > M \right). \tag{4.15}$$

Assume by way of contradiction that (4.15) fails for some $\xi \ne 0$ and $M > 0$, then $\exists \xi_N \to \xi$ such that $D_N(\xi_N) \le M$. Let

$$\varepsilon_N := \frac{\xi_N}{\xi} - 1, \text{ and } \mathsf{g} := \{(1+\varepsilon_N) f_k^{(N)}\}.$$

Let $D_N(\xi,\mathsf{g})$ denote the structure constants of (X,g), then $\sup_N D_N(\xi,\mathsf{g}) = \sup_N D_N(\xi_N) \le M$, whence $\xi \in H(\mathsf{X},\mathsf{g})$. Thus $H(\mathsf{X},\mathsf{g}) \ne \{0\}$, whence $G_{ess}(\mathsf{X},\mathsf{g}) \ne \mathbb{R}$, a contradiction. (4.15) follows, and (X,f) is stably hereditary.

(⇒): Suppose (X,f) is stably hereditary, then $D_N(\xi) \to \infty$ uniformly on compact subsets of $\mathbb{R} \setminus \{0\}$, so $D_N(\xi,\mathsf{g}) \to \infty$ for all $\xi \ne 0$, $\mathsf{g} = \{(1+\varepsilon_N) f_k^{(N)}\}$, and $\varepsilon_N \to 0$. Thus $H(\mathsf{X}',\mathsf{g}|_{\mathsf{X}'}) = \{0\}$, and $G_{ess}(\mathsf{X}',\mathsf{g}|_{\mathsf{X}'}) = \mathbb{R}$ for all sub-arrays X' and such g. The proof of the first part of the theorem is complete.

The second part of the theorem assumes that $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$ or $\{0\}$, and asserts that f is stably hereditary iff it is hereditary.

It is sufficient to consider the case $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{Z}$: If $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$ with $t \ne 0$ we work with $t^{-1}\mathsf{f}$, and if $G_{ess}(\mathsf{X},\mathsf{f}) = \{0\}$ then $H(\mathsf{X},\mathsf{f}) = \mathbb{R}$ and the statement that $D_N(\xi) \to \infty$ on $H(\mathsf{X},\mathsf{f})^c$ (uniformly on compacts or not) holds vacuously.

By Theorem 4.5, we can write $\mathsf{f} = \mathsf{g} - \mathsf{h}$ where $G_{alg}(\mathsf{X},\mathsf{g}) = G_{ess}(\mathsf{X},\mathsf{g}) = G_{ess}(\mathsf{X},\mathsf{f})$, and h is a.s. uniformly bounded and center-tight. By Lemma 4.15, we can modify g and h by a suitable collection of uniformly bounded constants to arrange for g to be integer-valued: $\mathbb{P}[g^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) \in \mathbb{Z}] = 1$ for all n, N.

Choose an integer K such that ess sup $|\mathsf{g}| \le K$. Then the g-balance $\Gamma(P)$ of every hexagon $P \in \mathrm{Hex}(N, n)$ satisfies

$$\Gamma(P) \in \mathbb{Z} \cap [-6K, 6K].$$

Let $m_{\mathrm{Hex}}^{N,n}$ denote the hexagon measure on $\mathrm{Hex}(N, n)$ and define for every $\gamma \in \mathbb{Z} \cap [-6K, 6K]$,

$$\mu_N(\gamma) := \sum_{n=3}^{k_N} m_{\mathrm{Hex}}^{N,n}\{P \in \mathrm{Hex}(N, n) : \Gamma(P) = \gamma\}.$$

Since $|e^{i\xi\gamma} - 1|^2 = 4\sin^2 \frac{\xi\gamma}{2}$,

$$d_N^2(\xi, \mathsf{g}) = 4 \sum_{\gamma=-6K}^{6K} \mu_N(\gamma) \sin^2 \frac{\xi\gamma}{2}.$$

This expression shows that if $D_N(\xi, \mathsf{g}) \to \infty$ for some ξ, then $D_N(\eta, \mathsf{g}) \to \infty$ uniformly on an open neighborhood of this ξ. Thus, if (X, g) is hereditary, then (X, g) is stably hereditary. The converse statement is trivial.

By Lemma 4.21, $H(\mathsf{X}, \mathsf{g}) = H(\mathsf{X}, \mathsf{f})$. In addition, (X, g) is (stably) hereditary iff (X, f) is (stably) hereditary, because by Lemma 2.16 and Corollary 3.9,

$$D_N(\xi, \mathsf{f}) \ge \frac{1}{8} D_N(\xi, \mathsf{g}) - \frac{1}{8}\xi^2 \sup_n U_n(\mathsf{h}) = \frac{1}{8} D_N(\xi, \mathsf{g}) - \mathrm{const.}\xi^2$$
$$D_N(\xi, \mathsf{g}) \ge \frac{1}{8} D_N(\xi, \mathsf{f}) - \frac{1}{8}\xi^2 \sup_n U_n(\mathsf{h}) = \frac{1}{8} D_N(\xi, \mathsf{f}) - \mathrm{const.}\xi^2.$$

Therefore the equivalence of the hereditary and stable hereditary properties of (X, g) implies the equivalence of these properties for (X, f). □

4.4 Notes and References

In the stationary world, a center-tight cocycle is a coboundary (Schmidt [178]) and the problems discussed in this chapter reduce to the question how small can one make the range of a cocycle by subtracting from it a coboundary. The question appears naturally in the ergodic theory of group actions, because of its relation to the ergodic decomposition of skew-products [1, Ch. 8],[32],[178], and its relation to the structure of locally finite ergodic invariant measures for skew-products [7],[165],[175]. In the general setup of ergodic theory, minimal reductions such as in Theorem 4.5 are not always possible [128], but they do sometimes exist [165],[175].

The reduction lemma was proved for sums of independent random variables in [56]. For a version of Theorem 4.9 in this case, see [148].

The relevance of (ir)reducibility to the local limit theorem appears in a different form in the papers of Guivarc'h & Hardy [88], Aaronson & Denker [5], and Dolgopyat [56]. There "irreducibility" is expressed in terms of a condition which rules out non-trivial solutions for certain cohomological equations. We will meet this idea again when we discuss irreducibility in the context of homogeneous Markov chains (Theorem 8.9(3), [88]), and in the context of Markov chains in a random environment (Proposition 9.24).

It is more difficult to uncover the irreducibility condition in the probabilistic literature on the LLT for sums of independent random variable. Prokhorov [163] and Rozanov [169], for example, prove a LLT for independent $\mathbb{Z}$-valued random variables X_k assuming Lindeberg's condition (which is automatic for bounded random variables), $\sum \mathrm{Var}(X_k) = \infty$, and subject to an arithmetic condition on the distributions of X_k. For an interpretation of this condition in terms of the irreducibility conditions in this chapter, see §8.2. Other sufficient conditions such as those appearing in [144],[148],[190] can be analyzed in a similar way.

Chapter 5
The Local Limit Theorem in the Irreducible Case

Abstract We find the asymptotic behavior of $\mathbb{P}(S_N - z_N \in (a,b))$, assuming that $(z_N - \mathbb{E}(S_N))/\sqrt{\mathrm{Var}(S_N)}$ converges to a finite limit, and subject to the irreducibility condition: The algebraic range cannot be reduced by a center-tight modification.

5.1 Main Results

5.1.1 Local Limit Theorems for Markov Chains

In the next two theorems, we assume that f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov *chain* X, and we let $S_N = f_1(X_1, X_2) + \cdots + f_N(X_N, X_{N+1})$ and $V_N := \mathrm{Var}(S_N)$.

Recall that the **algebraic range** $G_{alg}(\mathsf{X}, \mathsf{f})$ is the smallest closed subgroup G with constants c_n such that $\mathbb{P}[f_n(X_n, X_{n+1}) - c_n \in G] = 1$ for all n. We call f **irreducible**, if there is no center-tight h such that $G_{alg}(\mathsf{X}, \mathsf{f} - \mathsf{h})$ is strictly smaller than $G_{alg}(\mathsf{X}, \mathsf{f})$.

Theorem 5.1 (Non-Lattice Case) *Suppose* f *is irreducible, with algebraic range* $\mathbb{R}$. *Then* $V_N \to \infty$, *and for every* $(a, b) \subset \mathbb{R}$ *and* $z_N, z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$,

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1 + o(1)]\frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}(b - a), \text{ as } N \to \infty. \tag{5.1}$$

Theorem 5.2 (Lattice Case) *Suppose* $t > 0$, f *is irreducible with algebraic range* $t\mathbb{Z}$, *and* $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$ *for all* N. *Then* $V_N \to \infty$, *and for all* $z_N \in \gamma_N + t\mathbb{Z}$ *and* $z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, *for all* $k \in \mathbb{Z}$,

$$\mathbb{P}[S_N - z_N = kt] = [1 + o(1)]\,\frac{e^{-z^2/2}t}{\sqrt{2\pi V_N}}, \text{ as } N \to \infty. \tag{5.2}$$

For a discussion of the necessity of the irreducibility assumption, see §6.1.3.
We can check the conditions of the theorems directly from the data of X and f, using the structure constants (2.26):

Lemma 5.3 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X. *Then*

(1) f *is irreducible with algebraic range* $\mathbb{R}$ *iff* $\sum d_n^2(\xi) = \infty$ *for all* $\xi \neq 0$.
(2) Fix $t > 0$, *then* f *is irreducible with algebraic range* $t\mathbb{Z}$ *iff* $\sum d_n^2(\xi) < \infty$ *for* $\xi \in (2\pi/t)\mathbb{Z}$ *and* $\sum d_n^2(\xi) = \infty$ *for* $\xi \notin (2\pi/t)\mathbb{Z}$.
(3) f *is irreducible with algebraic range* $\{0\}$ *iff there are constants* c_n *such that* $f_n(X_n, X_{n+1}) = c_n$ *a.s. for all* n.

Proof f is non-lattice and irreducible iff $G_{ess}(\mathsf{X}, \mathsf{f}) = G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$. By Theorem 4.3, this happens iff f has co-range $\{0\}$, which proves part (1). Part (2) is similar, and part (3) is a triviality. □

5.1.2 Local Limit Theorems for Markov Arrays

Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov *array* X with row lengths $k_N + 1$, and set $\mathsf{X} = \{X_n^{(N)}\}$, $\mathsf{f} = \{f_n^{(N)}\}$, $S_N = \sum_{i=1}^{k_N} f_i^{(N)}(X_i^{(N)}, X_{i+1}^{(N)})$, and $V_N := \mathrm{Var}(S_N)$.

The LLT for S_N may fail when $\mathsf{f}|_{\mathsf{X}'}$ has different essential range for different sub-arrays X'. To rule this out, we assume hereditary behavior, see §4.2.3.

Theorem 5.1'. *Suppose* f *is stably hereditary and irreducible, with algebraic range* $\mathbb{R}$*. Then* $V_N \to \infty$*, and for every* $(a,b) \subset \mathbb{R}$ *and* $z_N, z \in \mathbb{R}$ *s.t.* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N\to\infty]{} z$*,*

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1+o(1)]\frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}(b-a), \text{ as } N \to \infty. \tag{5.3}$$

Theorem 5.2'. *Suppose* $t > 0$ *and* f *is hereditary, irreducible, and with algebraic range* $t\mathbb{Z}$*. Suppose* $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$ *for all* N*. Then* $V_N \to \infty$*, and for all* $z_N \in \gamma_N + t\mathbb{Z}$ *and* $z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N\to\infty]{} z$*, for every* $k \in \mathbb{Z}$*,*

$$\mathbb{P}[S_N - z_N = kt] = [1+o(1)]\,\frac{e^{-z^2/2}t}{\sqrt{2\pi V_N}}, \text{ as } N \to \infty. \tag{5.4}$$

Whereas in the non-lattice case we had to assume that f is stably hereditary, in the lattice case it is sufficient to assume that f is hereditary. This is because the two assumptions are equivalent in the lattice case, see Theorem 4.14(2).

Again, it is possible to check the assumptions of the theorems using the structure constants (2.26):

Lemma 5.3'. *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov array* X *with row lengths* $k_N + 1$*. Then*

(1) f *is stably hereditary, irreducible, and with algebraic range* $\mathbb{R}$ *iff* $D_N \xrightarrow[N\to\infty]{} \infty$ *uniformly on compacts in* $\mathbb{R}\setminus\{0\}$*.*

(2) Fix $t > 0$*. Then* f *is hereditary and irreducible with algebraic range* $t\mathbb{Z}$*, iff* $\sup_N D_N(\xi) < \infty$ *for* $\xi \in \frac{2\pi}{t}\mathbb{Z}$*, and* $D_N(\xi) \xrightarrow[N\to\infty]{} \infty$ *for all* $\xi \notin \frac{2\pi}{t}\mathbb{Z}$*. In this case* f *is also stably hereditary.*

Proof As in the case of Markov chains, f is non-lattice and irreducible iff its co-range equals $\{0\}$. In this case, f is stably hereditary iff $\sum_{n=3}^{k_N} d_n^{(N)}(\xi)^2 \xrightarrow[N\to\infty]{} \infty$ uniformly on compact subsets of $\mathbb{R}\setminus\{0\}$. This proves part (1).

Part (2) is proved similarly, with the additional observation that by Theorem 4.14, in the lattice case, every hereditary additive functional is stably hereditary. □

5.1.3 Mixing Local Limit Theorems

Let f be an additive functional on a Markov array X with row lengths $k_N + 1$, and state spaces $\mathfrak{S}_n^{(N)}$. Let S_N and V_N be as in the previous section.

Theorem 5.4 (Mixing LLT) *Suppose* X *is uniformly elliptic, and* f *is irreducible, stably hereditary, and a.s. uniformly bounded. Let* $\mathfrak{A}_N \subset \mathfrak{S}_{k_N+1}^{(N)}$ *be measurable events such that* $\mathbb{P}[X_{k_N+1}^{(N)} \in \mathfrak{A}_N]$ *is bounded away from zero. Fix* $x_N \in \mathfrak{S}_1^{(N)}$*, and let* $\phi : \mathbb{R} \to \mathbb{R}$ *be an arbitrary continuous function with compact support.*

(1) **Non-lattice Case:** *Suppose* f *has algebraic range* $\mathbb{R}$*. Then for every* $z_N, z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$*,*

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}[\phi(S_N - z_N)|X_{k_N+1}^{(N)} \in \mathfrak{A}_N, X_1^{(N)} = x_N] = \frac{e^{-z^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\phi(u)du.$$

(2) **Lattice Case:** *Suppose* f *has algebraic range* $t\mathbb{Z}$ *(*$t > 0$*) and* $\mathbb{P}[S_N \in \gamma_N + t\mathbb{Z}] = 1$ *for all* N*. Then for every* $z_N \in \gamma_N + t\mathbb{Z}$ *and* $z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$*,*

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}[\phi(S_N - z_N)|X^{(N)}_{k_N+1} \in \mathfrak{A}_N, X^{(N)}_1 = x_N] = \frac{e^{-z^2/2}t}{\sqrt{2\pi}} \sum_{u\in\mathbb{Z}} \phi(tu).$$

To understand what this means, take $\phi \in C_c(\mathbb{R})$ such that $\|\phi - 1_{(a,b)}\|_1 \ll 1$.

Remark. The conditioning on $X^{(N)}_1$ can be removed, using Lemma 2.27.

In the next chapter, we will use *mixing* LLT for *irreducible* additive functionals to study *ordinary* LLT for *reducible* additive functionals. Recall that every reducible additive functional can be put in the form $\mathsf{f} = \mathsf{g}+\nabla\mathsf{a}+\mathsf{h}$, where h has summable variance, a is uniformly bounded, and g is irreducible. Assume for simplicity that $\mathsf{h} \equiv 0$. Then $S_N(\mathsf{f}) = S_N(\mathsf{g}) + a_{N+1}(X_{N+1}) - a_1(X_1)$. To pass from the LLT for $S_N(\mathsf{g})$ (which we know since g is irreducible) to the LLT for $S_N(\mathsf{f})$ (which we do not know because of the reducibility of f), we need to understand the *joint* distribution of $S_N(\mathsf{g})$, $a_1(X_1)$ and $a_{N+1}(X_{N+1})$. The mixing LLT helps to do that, since conditioned on X_{N+1} and X_1, $S_N(\mathsf{f}) - S_N(\mathsf{g})$ is a constant.

5.2 Proofs

5.2.1 Strategy of Proof

Our aim is to find the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (a,b)]$ as $N \to \infty$, and assuming that $\frac{z_N-\mathbb{E}(S_N)}{\sqrt{V_N}} \to z$. We will use Fourier analysis.

(I) Fourier-Analytic Formulation of the LLT. $\mathbb{P}[S_N - z_N \in (a,b)]$ can be written in terms of the characteristic functions $\Phi_N(\xi) := \mathbb{E}(e^{i\xi S_N})$ as follows:

- In the lattice case, say when S_N and z_N are integer valued, we write the indicator function $\delta_0(m) := 1_{\{0\}}(m)$ in the form $\delta_0(m) := \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{im\xi}d\xi$ $(m \in \mathbb{Z})$. So, by Fubini's theorem

$$\mathbb{P}[S_N - z_N{=}k]{=}\mathbb{E}[\delta_0(S_N - z_N - k)]{=}\mathbb{E}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\xi(S_N-z_N-k)}d\xi\right){=}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\xi z_N}e^{-i\xi k}\Phi_N(\xi)d\xi, \quad (5.5)$$

- In the non-lattice case, we put the indicator function $\phi_{a,b} = 1_{(a,b)}$ in the form

$$\phi_{a,b}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\xi}\widehat{\phi}_{a,b}(\xi)d\xi, \text{ for } \widehat{\phi}_{a,b}(\xi) = \frac{e^{-ia\xi} - e^{-ib\xi}}{i\xi} \text{ and } t \neq a, b$$

(this follows from the identity $\int_{-\infty}^{\infty}(e^{ip\xi}/\xi)d\xi = \mathrm{sgn}(p)\pi i$ for $p \in \mathbb{R}\setminus\{0\}$). So

$$\mathbb{P}[S_N{-}z_N{\in}(a,b)]{=}\mathbb{E}[\phi_{a,b}(S_N{-}z_N)]{=}\mathbb{E}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\xi(S_N-z_N)}\widehat{\phi}_{a,b}(\xi)d\xi\right){=}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi z_N}\widehat{\phi}_{a,b}(\xi)\Phi_N(\xi)d\xi, \quad (5.6)$$

provided a, b are not atoms of $S_N - z_N$, for any N. (Such a, b are dense in $\mathbb{R}$.)

To analyze these integrals, we need to control $\Phi_N(\xi)$ as $N \to \infty$. The integral (5.6) is much more difficult than (5.5): To understand (5.5) it is sufficient to control $\mathbb{E}(e^{i\xi S_N})$ on the *compact* interval $[-\pi, \pi]$. But to understand (5.6), we must control $\mathbb{E}(e^{i\xi S_N})$ on *all of* $\mathbb{R}$, and it is not sufficient for the error to be small in L^∞, it needs to be small in L^1. Getting such control on all of $\mathbb{R}$ is not easy.

Luckily, there is a way to circumvent this difficulty. As noted by Charles Stone, instead of calculating the asymptotic behavior of (5.6) for the *specific function* $\phi_{a,b}(t)$, it is sufficient to find the asymptotic behavior of (5.6) for *all L^1 functions ϕ whose Fourier transforms $\widehat{\phi}$ have compact supports*. We defer the precise statement to §5.2.3. At this point we just want to emphasize that thanks to Stone's trick, (5.6) can be replaced by the integral

$$\frac{1}{2\pi}\int_{-L}^{L} e^{-i\xi z_N}\widehat{\phi}(\xi)\Phi_N(\xi)d\xi \tag{5.7}$$

where $[-L, L] \supset \mathrm{supp}(\widehat{\phi})$, and we are back to the problem of estimating $\Phi_N(\xi)$ *uniformly on compacts*.

(II) What Does the CLT Say? In Chapter 3 we proved the CLT, and it is reasonable to ask what does this result say on the integral (5.7).

The answer is that the CLT gives the behavior of the part of the integral (5.7) located within distance $O(V_N^{-1/2})$ from the origin: For every R, no matter how large,

$$\frac{1}{2\pi}\int_{-R/\sqrt{V_N}}^{R/\sqrt{V_N}} e^{-i\xi z_N}\widehat{\phi}(\xi)\Phi_N(\xi)d\xi=\frac{1}{2\pi\sqrt{V_N}}\int_{-R}^{R} e^{-\frac{iz_N\xi}{\sqrt{V_N}}}\widehat{\phi}\left(\frac{\xi}{\sqrt{V_N}}\right)\Phi_N\left(\frac{\xi}{\sqrt{V_N}}\right)d\xi=\frac{1}{2\pi\sqrt{V_N}}\int_{-R}^{R}\widehat{\phi}\left(\frac{\xi}{\sqrt{V_N}}\right)e^{-i\xi\frac{z_N-\mathbb{E}(S_N)}{\sqrt{V_N}}}\mathbb{E}(e^{i\xi\frac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}})d\xi$$

$$\overset{(1)}{\sim}\frac{1}{2\pi\sqrt{V_N}}\int_{-R}^{R}\widehat{\phi}(0)e^{-i\xi z-(\xi^2/2)}d\xi \overset{(2)}{=} \frac{\widehat{\phi}(0)e^{-z^2/2}}{\sqrt{2\pi V_N}}\left(1+O(e^{-R})\right). \tag{5.8}$$

The first marked identity uses the continuity of $\widehat{\phi}$, the assumption $\frac{z_N-\mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, the CLT, and the bounded convergence theorem. The second marked identity uses the well-known formula $\int_{\mathbb{R}} e^{-i\xi z-\xi^2/2}d\xi = \sqrt{2\pi}e^{-z^2/2}$.

If ϕ is close to $1_{(a,b)}$ in L^1, then $\widehat{\phi}(0) = \int \phi dx \approx |a-b|$, so $\frac{\widehat{\phi}(0)e^{-z^2/2}}{\sqrt{2\pi V_N}} \approx \frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$. Therefore (5.8) gives the asymptotic predicted by the LLT.

(III) Showing that the Peripheral Contribution is Negligible. To prove the LLT, it remains to show that the peripheral contribution to (5.7), coming from the integral over $\{\xi \in [-L, L] : |\xi| > RV_N^{-1/2}\}$, is negligible, namely, it is $V_N^{-1/2}o_{R\to\infty}(1)$. This is the crux of the matter, the central mathematical difficulty in the proof.

Since $\widehat{\phi}$ is bounded, the peripheral contribution is less than a constant times $\int_{R/\sqrt{V_N}}^{L}|\Phi_N(\xi)|d\xi + \int_{-L}^{-R/\sqrt{V_N}}|\Phi_N(\xi)|d\xi$. There are two things to worry about:

(1) **The Behavior Close to Zero:** $\Phi_N(0) = 1$, therefore $|\Phi(\xi)|$ *must* be large on a neighborhood of $\pm R/\sqrt{V_N}$.

(2) **The Behavior Away from Zero:** $\Phi_N(\xi)$ *could* be large away from zero, due to approximate arithmetic structures in the distribution of S_N. For example, if $1-\epsilon$ of the mass of S_N is located within distance ϵ from a coset of $(2\pi/\xi)\mathbb{Z}$, then $|\Phi_N(\xi)| = 1 - O(\epsilon)$.

We address these issues using two key estimates. The first, Proposition 5.7, says that for some positive constants c_1, c_2,

$$|\Phi_N(\xi)| \le c_1 \exp\left[-c_2 D_N(\xi)\right], \tag{5.9}$$

where $D_N(\xi)$ are the structure constants from (2.26). The second, Corollary 5.10, says that if $|\Phi_N(\xi)|$ is large at some value $\widetilde{\xi}_N$ then $|\Phi_N(\widetilde{\xi}_N + t)|$ drops very fast to zero as t moves away from zero. For example, in the special case $\widetilde{\xi}_N = 0$, Cor. 5.10 says that $\exists\widetilde{\delta}, \widehat{\varepsilon}, \widetilde{C} > 0$ such that for every N,

$$|\Phi_N(\xi)| \le \widetilde{C}e^{-\widehat{\varepsilon}V_N\xi^2} \quad (|\xi| < \widetilde{\delta}). \tag{5.10}$$

In particular, $\displaystyle\int_{-\widetilde{\delta}}^{-R/\sqrt{V_N}}|\Phi_N(t)|dt + \int_{R/\sqrt{V_N}}^{\widetilde{\delta}}|\Phi_N(t)|dt = V_N^{-1/2}o_{R\to\infty}(1)$, which takes care of the peripheral behavior near zero.

Corollary 5.10 also leads to the crucial estimate (5.30): $\exists\widetilde{\delta}_0 > 0$ so that for each interval I of length at most $\widetilde{\delta}_0$, if $A_N(I) := |\log\|\Phi_N\|_{L^\infty(I)}|$, then

$$\|\Phi_N\|_{L^1(I)} \le \frac{\text{const.}}{\sqrt{V_N A_N}}. \tag{5.11}$$

Note that by (5.9), $A_N(I) \geq \text{const} \sup_{\xi \in I} D_N(\xi) - \text{const}$.

For uniformly elliptic Markov chains (and stably hereditary arrays), $D_N(\xi) \to \infty$ uniformly on compacts outside the co-range of (X, f). So for any closed interval I of length at most $\widetilde{\delta}_0$ outside $H(\mathsf{X}, \mathsf{f})$, $A_N(I) \to \infty$, and the right-hand-side of (5.11) is $V_N^{-1/2} o(1)$. This takes care of the behavior away from $H(\mathsf{X}, \mathsf{f})$.

We now employ the assumption that (X, f) is irreducible: In the non-lattice case $H(\mathsf{X}, \mathsf{f}) = \{0\}$; in the lattice case, when the algebraic range is $\mathbb{Z}$, $H(\mathsf{X}, \mathsf{f}) = 2\pi\mathbb{Z}$ — but the domain of integration in (5.5) is $[-\pi, \pi]$. In both cases there can be no problematic ξ, except for $\xi = 0$, with which we have already dealt.

Standing Assumptions and Notation for the Remainder of the Chapter: We fix an additive functional $\mathsf{f} = \{f_n^{(N)}\}$ on a Markov array $\mathsf{X} = \{X_n^{(N)}\}$ with row lengths $k_N + 1$, state spaces $\mathfrak{S}_n^{(N)}$, and transition probabilities $\pi_{n,n+1}^{(N)}(x, \mathrm{d}y)$.

We assume throughout that ess sup $|\mathsf{f}| < K < \infty$, and that X is uniformly elliptic with ellipticity constant ϵ_0.

Recall that this means that $\pi_{n,n+1}^{(N)}(x, \mathrm{d}y) = p_n^{(N)}(x, y)\mu_{n+1}^{(N)}(\mathrm{d}y)$, where

$$0 \leq p_n^{(N)}(x, y) < \epsilon_0^{-1} \text{ and } \int p_n^{(N)}(x, y) p_{n+1}^{(N)}(y, z) \mu_{n+1}^{(N)}(\mathrm{d}y) > \epsilon_0.$$

There is no loss of generality in assuming that $\mu_k^{(N)}(E) = \mathbb{P}(X_k^{(N)} \in E)$ for $k \geq 3$, see Corollary 2.9.

5.2.2 Characteristic Function Estimates

It is convenient to use the **characteristic functions of S_N conditioned on X_1**:

$$\Phi_N(x, \xi) := \mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}\xi S_N}\right) \equiv \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\xi S_N} | X_1^{(N)} = x\right).$$

$$\Phi_N(x, \xi | \mathfrak{A}) = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\xi S_N} | X_{k_N+1}^{(N)} \in \mathfrak{A}, X_1^{(N)} = x\right) := \frac{\mathbb{E}_x(\mathrm{e}^{\mathrm{i}\xi S_N} 1_{\mathfrak{A}}(X_{k_N+1}^{(N)}))}{\mathbb{P}_x(X_{k_{N+1}}^{(N)} \in \mathfrak{A}_{k_N+1}^{(N)})}.$$

Here $x \in \mathfrak{S}_1^{(N)}$, $\xi \in \mathbb{R}$, $\mathbb{E}_x(\cdot) = \mathbb{E}(\,\cdot\,| X_1^{(N)} = x)$, and $\mathfrak{A} \subset \mathfrak{S}_{k_N+1}^{(N)}$ are measurable.

We write these functions in terms of **Nagaev's perturbation operators**: $\mathcal{L}_{n,\xi}^{(N)} : L^\infty(\mathfrak{S}_{n+1}^{(N)}) \to L^\infty(\mathfrak{S}_n^{(N)})$, defined for $1 \leq n \leq k_N$ and $N \in \mathbb{N}$, by

$$\left(\mathcal{L}_{n,\xi}^{(N)} v\right)(x) := \int_{\mathfrak{S}_{n+1}^{(N)}} p_n^{(N)}(x, y) \mathrm{e}^{\mathrm{i}\xi f_n^{(N)}(x,y)} v(y) \mu_{n+1}^{(N)}(\mathrm{d}y) \equiv \mathbb{E}(\mathrm{e}^{\mathrm{i}\xi f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)})} v(X_{n+1}^{(N)}) | X_n^{(N)} = x).$$

Lemma 5.5 (Nagaev) *Let* $1(\cdot) \equiv 1$, *then the following identities hold:*

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\xi S_N} v(X_{k_N+1}^{(N)}) \middle| X_1^{(N)} = x\right) = \left(\mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \cdots \mathcal{L}_{k_N,\xi}^{(N)} v\right)(x), \tag{5.12}$$

$$\Phi_N(x, \xi) = \left(\mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \cdots \mathcal{L}_{k_N,\xi}^{(N)} 1\right)(x), \tag{5.13}$$

$$\Phi_N(x, \xi | \mathfrak{A}) = \frac{\left(\mathcal{L}_{1,\xi}^{(N)} \mathcal{L}_{2,\xi}^{(N)} \cdots \mathcal{L}_{N,\xi}^{(N)} 1_{\mathfrak{A}}\right)(x)}{\mathbb{P}_x[X_{k_N+1}^{(N)} \in \mathfrak{A}]}. \tag{5.14}$$

Proof $\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_N} v(X_{k_N+1}^{(N)}) | X_1^{(N)} = x) = \mathbb{E}[\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_N} v(X_{k_N+1}^{(N)}) | X_1^{(N)}, X_2^{(N)}) | X_1^{(N)} = x]$

$$= \int p_1^{(N)}(x,y)e^{i\xi f_1^{(N)}(x,y)}\mathbb{E}(e^{i\xi\sum_{n=2}^{k_N} f_n^{(N)}} v|X_2^{(N)}=y)\mu_2^{(N)}(dy).$$

Proceeding by induction, we obtain (5.12), and (5.12) implies (5.13) and (5.14). □

Let $\|\cdot\|$ denote the operator norm, i.e. $\|\mathcal{L}_{n,\xi}^{(N)}\| = \sup\left\{\|\mathcal{L}_{n,\xi}^{(N)}v\|_\infty : \|v\|_\infty \le 1\right\}$.

Lemma 5.6 $\|\mathcal{L}_{n,\xi}^{(N)}\| \le 1$*, and there is a positive constant* $\widetilde{\varepsilon}$ *which only depends on* ϵ_0 *such that for all* $N \ge 1$ *and* $5 \le n \le k_N$*,*

$$\left\|\mathcal{L}_{n-4,\xi}^{(N)}\mathcal{L}_{n-3,\xi}^{(N)}\mathcal{L}_{n-2,\xi}^{(N)}\mathcal{L}_{n-1,\xi}^{(N)}\mathcal{L}_{n,\xi}^{(N)}\right\| \le e^{-\widetilde{\varepsilon}d_n^{(N)}(\xi)^2}.$$

Proof It is clear that $\|\mathcal{L}_{n,\xi}^{(N)}\| \le 1$. We will present the operator $\mathcal{L}^{(N)} := \mathcal{L}_{n-4,\xi}^{(N)}\mathcal{L}_{n-3,\xi}^{(N)}\mathcal{L}_{n-2,\xi}^{(N)}\mathcal{L}_{n-1,\xi}^{(N)}\mathcal{L}_{n,\xi}^{(N)}$ as an integral operator, and study the kernel.

Henceforth we fix N, and drop the superscripts $^{(N)}$. The variables x_i, z_i will always denote points in $\mathfrak{S}_i = \mathfrak{S}_i^{(N)}$, and $\int \varphi(z_i)dz_i := \int_{\mathfrak{S}_i} \varphi(z_i)\mu_i^{(N)}(dz_i)$. Let

$p(x_k,\ldots,x_m) := \prod_{i=k}^{m-1} p_i(x_i,x_{i+1})$, $f(x_k,\ldots,x_m) := \sum_{i=k}^{m-1} f_i(x_i,x_{i+1})$, and

$$L(x_{n-4},z_{n+1}) := \iiiint p(x_{n-4},z_{n-3},\ldots,z_{n+1})e^{i\xi f(x_{n-4},z_{n-3},\ldots,z_{n+1})}dz_{n-3}\cdots dz_n.$$

Then $(\mathcal{L}v)(x_{n-4}) = \int [L(x_{n-4},z_{n+1})v(z_{n+1})]dz_{n+1}$, and it follows that

$$\|\mathcal{L}v\|_\infty \le \|v\|_\infty \sup_{x_{n-4}\in\mathfrak{S}_{n-4}} \int |L(x_{n-4},z_{n+1})|dz_{n+1} \tag{5.15}$$

$$\le \|v\|_\infty \sup_{x_{n-4}\in\mathfrak{S}_{n-4}} \iint dz_{n-2}dz_{n+1}\left[|K_n(z_{n-2},z_{n+1})| \times \int p(x_{n-4},z_{n-3},z_{n-2})dz_{n-3}\right], \tag{5.16}$$

where $K_n(z_{n-2},z_{n+1}) := \iint p(z_{n-2},z_{n-1},z_n,z_{n+1})e^{i\xi f(z_{n-2},z_{n-1},z_n,z_{n+1})}dz_{n-1}dz_n$.

Claim: *Let* $p(z_{n-2}\to z_{n+1}) := \iint p(z_{n-2},z_{n-1},z_n,z_{n+1})dz_{n-1}dz_n$*, then*

$$|K_n(z_{n-2},z_{n-1})| \le p(z_{n-2}\to z_{n+1}) - \frac{1}{4}p(z_{n-2}\to z_{n+1})\mathbb{E}\left(|e^{i\xi\Gamma\left(X_{n-2}\,{X_{n-1}\atop Y_{n-1}}\,{X_n\atop Y_n}\,X_{n+1}\right)}-1|^2\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right), \tag{5.17}$$

where $\{Y_n\}$ *is an independent copy of* $\{X_n\}$ *and* Γ *is the balance* (2.25).

Proof of the Claim. Set $\widetilde{K}_n(z_{n-2},z_{n+1}) := \dfrac{K_n(z_{n-2},z_{n+1})}{p(z_{n-2}\to z_{n+1})}$.

Looking at identity (2.20) for the bridge probabilities $\mathbb{P}(\,\cdot\,|X_{n-2}=z_{n-2},X_{n+1}=z_{n+1})$, we find that

$$\widetilde{K}_n(z_{n-2},z_{n+1}) \equiv \mathbb{E}\left(e^{i\xi\sum_{k=n-2}^n f_k(X_k,X_{k+1})}\middle|\begin{matrix}X_{n-2}=z_{n-2}\\X_{n+1}=z_{n+1}\end{matrix}\right),\quad \overline{\widetilde{K}_n(z_{n-2},z_{n+1})} \equiv \mathbb{E}\left(e^{-i\xi\sum_{k=n-2}^n f_k(Y_k,Y_{k+1})}\middle|\begin{matrix}Y_{n-2}=z_{n-2}\\Y_{n+1}=z_{n+1}\end{matrix}\right).$$

If $X_{n-2}=Y_{n-2}$ and $X_{n+1}=Y_{n+1}$, then

$$e^{i\xi\sum_{k=n-2}^n f_k(X_k,X_{k+1})}e^{-i\xi\sum_{k=n-2}^n f_k(Y_k,Y_{k+1})} = e^{i\xi\Gamma\left(X_{n-2}\,{X_{n-1}\atop Y_{n-1}}\,{X_n\atop Y_n}\,X_{n+1}\right)}.$$

Multiplying the identities for $\widetilde{K}_n$ and $\overline{\widetilde{K}_n}$, denoting $P := \left(X_{n-2}\,{X_{n-1},\atop Y_{n-1}}\,{X_n\atop Y_n}\,X_{n+1}\right)$ and recalling that X,Y are independent copies, we arrive at the following consequence:

$$|\widetilde{K}_n(z_{n-2},z_{n+1})|^2 = \mathbb{E}\left(e^{i\xi\Gamma\left(X_{n-2}\,{}^{X_{n-1}}_{Y_{n-1}}\,{}^{X_n}_{Y_n}\,X_{n+1}\right)}\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\ X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right) = \mathbb{E}\left(\cos(\xi\Gamma(P))\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\ X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right),$$

$$= 1-\frac{1}{2}\mathbb{E}\left(|e^{i\xi\Gamma(P)}-1|^2\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\ X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right),\quad \text{because } \cos\alpha = 1-\tfrac{1}{2}|e^{i\alpha}-1|^2.$$

Since $\sqrt{1-t}\le 1-\frac{t}{2}$ for all $0\le t\le 1$, (5.17) follows.

Now that we proved the claim, we substitute (5.17) in (5.16). The result is a bound of the form $\int |L(x_{n-4},z_{n+1})|dz_{n+1}\le \text{I}-\text{II}$, where

$$\begin{aligned}\text{I} &:= \iint dz_{n-2}dz_{n+1}\left[p(z_{n-2}\to z_{n+1})\int p(x_{n-4},z_{n-3},z_{n-2})dz_{n-3}\right]\\ &\equiv \int\cdots\int p(x_{n-4},z_{n-3},z_{n-2},z_{n-1},z_n,z_{n+1})dz_{n+1}dz_n dz_{n-1}dz_{n-2}dz_{n-3};\end{aligned}$$

$$\begin{aligned}\text{II} := \frac{1}{4}\iint dz_{n-2}dz_{n+1}\Big[p(z_{n-2}\to z_{n+1})\times\int p(x_{n-4},z_{n-3},z_{n-2})dz_{n-3}\times\\ \times\,\mathbb{E}\left(|e^{i\xi\Gamma\left(X_{n-2}\,{}^{X_{n-1}}_{Y_{n-1}}\,{}^{X_n}_{Y_n}\,X_{n+1}\right)}-1|^2\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\ X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right)\Big].\end{aligned}$$

Clearly I = 1. We will now show that II $\ge$ const. $d_n^{(N)}(\xi)^2$.

First, $p(z_{n-2}\to z_{n+1})\ge\epsilon_0$ by (2.19), and $\int p(x_{n-4},z_{n-3},z_{n-2})dz_{n-3}\ge\epsilon_0$ by uniform ellipticity. So

$$\text{II}\ge\frac{\epsilon_0^2}{4}\iint\mathbb{E}\left(|e^{i\xi\Gamma\left(X_{n-2}\,{}^{X_{n-1}}_{Y_{n-1}}\,{}^{X_n}_{Y_n}\,X_{n+1}\right)}-1|^2\middle|\begin{matrix}X_{n-2}=Y_{n-2}=z_{n-2}\\ X_{n+1}=Y_{n+1}=z_{n+1}\end{matrix}\right)dz_{n-2}dz_{n+1}.$$

Recalling the definitions of the bridge probabilities, we obtain the following:

$$\begin{aligned}\text{II}&\ge\frac{\epsilon_0^2}{4}\iiiint\!\!\iint |e^{i\xi\Gamma\left(z_{n-2}\,{}^{x_{n-1}}_{y_{n-1}}\,{}^{x_n}_{y_n}\,z_{n+1}\right)}-1|^2\frac{p(z_{n-2},x_{n-1},x_n,z_{n+1})dx_{n-1}dx_n}{p(z_{n-2}\to z_{n+1})}\frac{p(z_{n-2},y_{n-1},y_n,z_{n+1})dy_{n-1}dy_n}{p(z_{n-2}\to z_{n+1})}dz_{n-2}dz_{n+1}\\ &\overset{(2.23)}{\equiv}\frac{\epsilon_0^2}{4}\int_{\text{Hex}(N,n)}|e^{i\xi\Gamma(P)}-1|^2 m'_{\text{Hex}}(dP)\overset{(2.24)}{\ge}\frac{\epsilon_0^6}{4}\int_{\text{Hex}(N,n)}|e^{i\xi\Gamma(P)}-1|^2 m_{\text{Hex}}(dP),\ge\frac{\epsilon_0^6}{4}d_n^{(N)}(\xi)^2.\end{aligned}$$

In summary, $\int|L(x_{n-4},z_{n+1})|dz_{n+1}\le \text{I}-\text{II}\le 1-\frac{1}{4}\epsilon_0^6 d_n^{(N)}(\xi)^2$. By (5.15), $\|\mathcal{L}\|\le 1-\widetilde{\epsilon}d_n^{(N)}(\xi)^2$, where $\widetilde{\epsilon}:=\frac{1}{4}\epsilon_0^6$. Since $1-t\le e^{-t}$, we are done. □

Proposition 5.7 *Let $\widetilde{\varepsilon}>0$ be the constant in Lemma 5.6.*

(1) $\exists C>0$ independent of N such that for all N,

$$|\Phi_N(x,\xi)|\le Ce^{-\frac{1}{5}\widetilde{\varepsilon}D_N(\xi)}. \tag{5.18}$$

(2) $\forall\overline{\delta}>0\ \exists C(\overline{\delta})>0$ such that if $\mathbb{P}[X_{k_N+1}^{(N)}\in\mathfrak{A}]\ge\overline{\delta}$, then

$$|\Phi_N(x,\xi|\mathfrak{A})|\le C(\overline{\delta})e^{-\frac{1}{5}\widetilde{\varepsilon}D_N(\xi)}. \tag{5.19}$$

Proof $D_N(\xi)\equiv\sum_{n=3}^{k_N}d_n^{(N)}(\xi)^2=\sum_{j=0}^{4}D_{j,N}$, where $D_{j,N}(\xi)=\sum_{\substack{3\le n\le k_N\\ n\equiv j \bmod 5}}d_n^{(N)}(\xi)^2$. Applying Lemma 5.6 iteratively, we obtain $|\Phi_N(x,\xi)|\le Ce^{-\widetilde{\varepsilon}\max\left(D_{0,N},\dots,D_{4,N}\right)}\le Ce^{-\frac{1}{5}\widetilde{\varepsilon}D_N(\xi)}$, whence (5.18).

If $\mathbb{P}(X_{k_N+1}^{(N)}\in\mathfrak{A})\ge\overline{\delta}$ then $|\Phi_N(x,\xi|\mathfrak{A})|\le\overline{\delta}^{-1}\|\mathcal{L}_{1,\xi}^{(N)}\mathcal{L}_{2,\xi}^{(N)}\dots\mathcal{L}_{k_N,\xi}^{(N)}1_{\mathfrak{A}}\|_\infty$, see (5.14). Continuing as before, we obtain (5.19). □

The next result says that if $u_n^{(N)}$ is big, then $d_n^N(\cdot)$ cannot be small at two nearby points. Recall the standing assumption ess sup $\|f_n^{(N)}\|_\infty \le K$.

Lemma 5.8 $\exists\widetilde{\delta} = \widetilde{\delta}(K) > 0$ *such that if* $|\delta| \le \widetilde{\delta}$*, then for all* $3 \le n \le k_N$*,*

$$d_n^{(N)}(\xi+\delta)^2 \ge \frac{2}{3}\delta^2(u_n^{(N)})^2 - 2|\delta|u_n^{(N)}d_n^{(N)}(\xi). \tag{5.20}$$

Proof Fix a hexagon $P \in \mathrm{Hex}(N,n)$, and let $\mathfrak{u}_n := \Gamma(P)$ and $\mathfrak{d}_n(\xi) := |e^{i\xi\mathfrak{u}_n} - 1|$.

$$\begin{aligned}\mathfrak{d}_n^2(\xi+\delta) &= |e^{i(\xi+\delta)\mathfrak{u}_n} - 1|^2 = 2[1-\cos((\xi+\delta)\mathfrak{u}_n)] = 2[1-\cos(\xi\mathfrak{u}_n)\cos(\delta\mathfrak{u}_n) + \sin(\xi\mathfrak{u}_n)\sin(\delta\mathfrak{u}_n)]\\ &= 2[(1-\cos(\xi\mathfrak{u}_n))\cos(\delta\mathfrak{u}_n) + (1-\cos(\delta\mathfrak{u}_n)) + \sin(\xi\mathfrak{u}_n)\sin(\delta\mathfrak{u}_n)].\end{aligned} \tag{5.21}$$

Suppose $|\delta| < \widetilde{\delta} < \frac{\pi}{12K}$, then $|\delta\mathfrak{u}_n| < \frac{\pi}{2}$, so $\cos(\delta\mathfrak{u}_n) \ge 0$. Make $\widetilde{\delta}$ even smaller to guarantee that $0 \le |t| \le 6K\widetilde{\delta} \Rightarrow \frac{1}{3}t^2 \le 1 - \cos t \le t^2$. Then,

$$\begin{aligned}\mathfrak{d}_n^2(\xi+\delta) &\ge 2\big[(1-\cos(\delta\mathfrak{u}_n)) - |\sin(\xi\mathfrak{u}_n)\sin(\delta\mathfrak{u}_n)|\big] \ge 2\big(\tfrac{1}{3}\delta^2\mathfrak{u}_n^2 - |\delta\mathfrak{u}_n|\sqrt{1-\cos^2(\xi\mathfrak{u}_n)}\big)\\ &= 2\Big(\tfrac{1}{3}\delta^2\mathfrak{u}_n^2 - |\delta\mathfrak{u}_n|\sqrt{(1-\cos(\xi\mathfrak{u}_n))(1+\cos(\xi\mathfrak{u}_n))}\Big) \ge 2\Big(\tfrac{1}{3}\delta^2\mathfrak{u}_n^2 - |\delta\mathfrak{u}_n|\sqrt{2(1-\cos(\xi\mathfrak{u}_n))}\Big)\\ &= \tfrac{2}{3}\delta^2\mathfrak{u}_n^2 - 2|\delta\mathfrak{u}_n||e^{i\xi\mathfrak{u}_n} - 1| \ge \tfrac{2}{3}\delta^2\mathfrak{u}_n^2 - 2|\delta\mathfrak{u}_n|\mathfrak{d}_n(\xi).\end{aligned}$$

Integrating on $P \in \mathrm{Hex}(N,n)$, and using the Cauchy-Schwarz inequality to estimate the second term, gives the result. □

Proposition 5.9 *Let* $\widetilde{\delta}$ *be the constant from Lemma 5.8. There are* $\widehat{\varepsilon}, \widehat{c}, C, M > 0$*, which only depend on* ϵ_0 *and* K*, such that if* $V_N > M$*, then for all* ξ *and* $|\delta| < \widetilde{\delta}$

$$|\Phi_N(x,\xi+\delta)| \le C\exp\left(-\widehat{\varepsilon}V_N\delta^2 + \widehat{c}|\delta|\sqrt{V_N D_N(\xi)}\right). \tag{5.22}$$

Proof $U_n \equiv \sum_{k=3}^{k_N}(u_k^{(N)})^2$. By Lemma 5.8 and the Cauchy-Schwarz inequality, $D_N(\xi+\delta) \ge \frac{2}{3}\delta^2 U_N - 2|\delta|\sqrt{U_N D_N(\xi)}$.

Theorem 3.6 says that there are two constants C_1, C_2 which only depend on ϵ_0 and K such that $C_1^{-1}U_N - C_2 \le V_N \le C_1 U_N + C_2$. This implies that $V_N > 2C_2 \Longrightarrow \dfrac{U_N}{2C_1} \le V_N \le 2C_1 U_N$. Thus there are constants $\widehat{\varepsilon}_1, \widehat{c}_1 > 0$, which only depend on ϵ_0 and K, such that for all N such that $V_N > 2C_2$, $D_N(\xi+\delta) \ge \widehat{\varepsilon}_1\delta^2 V_N - \widehat{c}_1|\delta|\sqrt{V_N D_N(\xi)}$. The proposition now follows from (5.18). □

Given a compact interval $I \subset \mathbb{R}$, let

$$A_N(I) := -\log\sup\{|\Phi_N(x,\xi)| : (x,\xi) \in \mathfrak{S}_1^{(N)} \times I\}. \tag{5.23}$$

Now choose some pair $(\widetilde{x}_N, \widetilde{\xi}_N) \in \mathfrak{S}_1^{(N)} \times I$ such that $A_N(I) \le -\log|\Phi_N(\widetilde{x}_N, \widetilde{\xi}_N)| \le A_N(I) + \ln 2$.

So $|\Phi(\widetilde{x}_N, \widetilde{\xi}_N)| \ge \frac{1}{2}e^{-A_N(I)} = \frac{1}{2}\sup|\Phi_N(\cdot,\cdot)|$ on $\mathfrak{S}_1^{(N)} \times I$. Then:

Corollary 5.10 *For each* $\overline{\delta}$ *there are constants* $\widetilde{C}, \widehat{\varepsilon}, \overline{c} > 0$ *which only depend on* ϵ_0 *and* K*, such that for every compact interval* I *with length* $|I| \le \widetilde{\delta}$*, for every measurable set* $\mathfrak{A} \subset \mathfrak{S}_{k_N+1}^{(N)}$ *with measure* $\mu_{k_N+1}^{(N)}(\mathfrak{A}) \ge \overline{\delta}$*, for every* $(x,\xi) \in \mathfrak{S}_1^{(N)} \times I$*, and for all* N*,*

$$\begin{aligned}|\Phi_N(x,\xi)| &\le \widetilde{C}\exp\left(-\widehat{\varepsilon}V_N(\xi-\widetilde{\xi}_N)^2 + \overline{c}|\xi-\widetilde{\xi}_N|\sqrt{V_N A_N(I)}\right);\\ |\Phi_N(x,\xi|\mathfrak{A})| &\le \widetilde{C}\exp\left(-\widehat{\varepsilon}V_N(\xi-\widetilde{\xi}_N)^2 + \overline{c}|\xi-\widetilde{\xi}_N|\sqrt{V_N A_N(I)}\right).\end{aligned}$$

Proof Let $\widetilde{\delta}$ be the constant from Lemma 5.8. Let I be an interval such that $|I| \le \widetilde{\delta}$, and fix some $(x,\xi) \in \mathfrak{S}_1^{(N)} \times I$. Choose $(\widetilde{x}_N, \widetilde{\xi}_N)$ as above.

By (5.18), $e^{-A_N(I)} \le 2|\Phi_N(\widetilde{x}_N, \widetilde{\xi}_N)| \le \text{const.}e^{-\frac{1}{5}\widetilde{\varepsilon}D_N(\widetilde{\xi}_N)}$, therefore there is a global constant C' such that $D_N(\widetilde{\xi}) \le C'A_N(I) + C'$.

Fix $M, C, \widehat{\varepsilon}$ and $\widehat{c}$ as in Proposition 5.9. Let $\alpha := \max\{1, M\widetilde{\delta}^2, 2C'(2\widehat{c}/\widehat{\varepsilon})^2\}$.

Suppose first that $V_N(\xi - \widetilde{\xi}_N)^2 \le \alpha$. $|\Phi_N(x,\xi)|$ and $|\Phi_N(x,\xi|\mathfrak{A})|$ are less than or equal to 1, and $\exp(-\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2) \ge \exp(-\widehat{\varepsilon}\alpha)$. Therefore the corollary holds with any $\widehat{\varepsilon}, \overline{c}, \widetilde{C} > 0$ such that $\widetilde{C} > \exp(\widehat{\varepsilon}\alpha)$.

Suppose now that $V_N(\xi - \widetilde{\xi}_N)^2 > \alpha$. Since $|\xi - \widetilde{\xi}_N| \le \widetilde{\delta}$, $V_N \ge M$, and (5.22) holds. (5.22) with $\widetilde{\xi}_N$ instead of ξ and with $\delta := \xi - \widetilde{\xi}_N$ says that $|\Phi_N(x,\xi)| \le C\exp\left(-\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2 + \widehat{c}|\widetilde{\xi}_N - \xi|\sqrt{V_N D_N(\widetilde{\xi}_N)}\right)$.

- If $A_N(I) > 1$, then $D_N(\widetilde{\xi}) \le C'A_N(I) + C' \le 2C'A_N(I)$, and
$$|\Phi_N(x,\xi)| \le C\exp\left(-\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2 + \widehat{c}\sqrt{2C'}|\widetilde{\xi}_N - \xi|\sqrt{V_N A_N(I)}\right).$$
- If $A_N(I) \le 1$, then $D_N(\widetilde{\xi}_N) \le 2C'$, and
$$\begin{aligned}|\Phi_N(x,\xi)| &\le C\exp\left(-\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2 + \widehat{c}\sqrt{2C'}|\widetilde{\xi}_N - \xi|\sqrt{V_N}\right)\\ &\le C\exp\left(-\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2 + \widehat{c}\sqrt{2C'}\frac{V_N(\widetilde{\xi}_N - \xi)^2}{\sqrt{\alpha}}\right), \quad \because \sqrt{V_N(\widetilde{\xi}_N - \xi)^2} \ge \sqrt{\alpha}\\ &\le C\exp\left(-\frac{1}{2}\widehat{\varepsilon}V_N(\xi - \widetilde{\xi}_N)^2\right), \quad \because \alpha \ge 2C'(2\widehat{c}/\widehat{\varepsilon})^2.\end{aligned}$$

Thus the corollary holds with $\frac{1}{2}\widehat{\varepsilon}$ replacing $\widehat{\varepsilon}$, $\overline{c} := \widehat{c}\sqrt{2C_1}$, and $\widetilde{C} := Ce^{\widehat{\varepsilon}\alpha}$. The second estimate has a similar proof. □

5.2.3 The LLT via Weak Convergence of Measures

In this section we give the mathematical background needed to justify Stone's trick from §5.2.1. Let $C_c(\mathbb{R})$ denote the space of real-valued continuous functions on $\mathbb{R}$, with compact support. Such functions are bounded and uniformly continuous, and they can all be approximated uniformly by piecewise constant functions. It follows that if the LLT asymptotic expansion (5.1) holds for all intervals (a,b), then

$$\underbrace{\sqrt{2\pi V_N}\mathbb{E}[\phi(S_N - z_N)]}_{\mu_N(\phi)} \xrightarrow[N\to\infty]{} \underbrace{e^{-z^2/2}\int \phi(t)dt}_{\mu_z(\phi)} \text{ for all } \phi \in C_c(\mathbb{R}). \tag{5.24}$$

Conversely, (5.24) implies (5.1): To see this, apply (5.24) with $\phi_n, \psi_n \in C_c(\mathbb{R})$ such that $\phi_n \le 1_{[a,b]} \le \psi_n$, and $\int |\phi_n(t) - \psi_n(t)|dt \to 0$.

A **Radon measure** on $\mathbb{R}$ is a positive Borel measure on $\mathbb{R}$, which may be infinite, but which assigns finite mass to every compact set. By the Riesz representation theorem, $\mu_N(\phi) = \int \phi d\mu_N$ and $\mu_z(\phi) = \int \phi d\mu_z$, for some Radon measures μ_N, μ_z.

Let m_N and m be Radon measures on $\mathbb{R}$. We say that m_n converges to m **weakly** (or **vaguely**) , if $\int \phi dm_N \to \int \phi dm$ for all $\phi \in C_c(\mathbb{R})$. In this case we write $m_N \xrightarrow[N\to\infty]{w} m$.

Since the non-lattice LLT (5.1) is equivalent to (5.24), it can be restated as saying that $\mu_N \xrightarrow{w} \mu_z$, with μ_N and μ_z as above. The other LLT have similar reformulations. For example, the lattice mixing LLT is equivalent to $\mu_N \xrightarrow{w} \mu_z$, where $\mu_N(\phi) := \sqrt{V_N}\mathbb{E}[\phi(S_N - z_N)|X_{k_N+1}^{(N)} \in \mathfrak{A}_N, X_1^{(N)} = x_N]$, and $\mu_z := \frac{e^{-z^2/2}t}{\sqrt{2\pi}}\sum_{u\in\mathbb{Z}}\delta_{tu}$.

Weak convergence is defined using test functions in $C_c(\mathbb{R})$, but we claim that it can also be checked using test functions in $\mathcal{K} := \left\{\phi : \mathbb{R} \to \mathbb{C} : \int |\phi(u)| \mathrm{d}u < \infty, \widehat{\phi} \in C_c(\mathbb{R})\right\}$, $(\widehat{\phi} :=$ Fourier transform of $\phi)$ (even though $C_c(\mathbb{R}) \cap \mathcal{K} = \{0\}$). We note for future reference that by the Fourier inversion formula, every $\phi \in \mathcal{K}$ is uniformly bounded and continuous.

Lemma 5.11 (Breiman) *Let m_N and m be Radon measures on $\mathbb{R}$. Suppose $\exists$ strictly positive $\phi_0 \in \mathcal{K}$ such that $\int \phi_0 \mathrm{d}m < \infty$. If for all $\phi \in \mathcal{K}$, $\int_{\mathbb{R}} \phi \mathrm{d}m_N \xrightarrow[N\to\infty]{} \int_{\mathbb{R}} \phi \mathrm{d}m$, then $\int_{\mathbb{R}} \psi \mathrm{d}m_N \xrightarrow[N\to\infty]{} \int_{\mathbb{R}} \psi \mathrm{d}m$ for all $\psi \in C_c(\mathbb{R})$.*

Proof Let ϕ_0 be as in the statement, and suppose $\int \phi \mathrm{d}m_N \to \int \phi \mathrm{d}m$ for all $\phi \in \mathcal{K}$.

If $\int_{\mathbb{R}} \phi_0 \mathrm{d}m = 0$, then $m \equiv 0$, and $\int_{\mathbb{R}} \phi_0 \mathrm{d}m_N \to 0$. Since ϕ_0 is positive and continuous, for every $\psi \in C_c(\mathbb{R})$ there exists $\varepsilon > 0$ such that $\varepsilon|\psi| \le \phi_0$ on $\mathbb{R}$. It follows that $\int_{\mathbb{R}} \psi \mathrm{d}m_N \to 0$ for all $\psi \in C_c(\mathbb{R})$.

Suppose now that $\int_{\mathbb{R}} \phi_0 \mathrm{d}m > 0$. Then $\int_{\mathbb{R}} \phi_0 \mathrm{d}m_N \to \int_{\mathbb{R}} \phi_0 \mathrm{d}m$, and $\int_{\mathbb{R}} \phi_0 \mathrm{d}m_N > 0$ for all N large enough.

For such N, we construct the following *probability* measures: $\mathrm{d}\mu_N := \dfrac{\phi_0 \mathrm{d}m_N}{\int \phi_0 \mathrm{d}m_N}$ and $\mathrm{d}\mu := \dfrac{\phi_0 \mathrm{d}m}{\int \phi_0 \mathrm{d}m}$.

$\forall t \in \mathbb{R}$, $\phi_t(u) := \mathrm{e}^{\mathrm{i}tu}\phi_0(u)$ belongs to $\mathcal{K}$, because $\widehat{\phi_t}(\xi) = \widehat{\phi_0}(\xi - t)$. So $\displaystyle\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tu}\mu_N(\mathrm{d}u) \xrightarrow[N\to\infty]{} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tu}\mu(\mathrm{d}u)$.

Equivalently, the characteristic functions of the random variables W_N with distribution $\mathbb{P}[W_N < a] = \mu_N((-\infty, a))$ converge to the characteristic function of the random variable W with distribution $\mathbb{P}[W < a] := \mu((-\infty, a))$. By Lévy's continuity theorem, $W_N \xrightarrow[n\to\infty]{\text{dist}} W$. It follows that $\int_{\mathbb{R}} G(u)\mu_N(\mathrm{d}u) \to \int_{\mathbb{R}} G(u)\mu(\mathrm{d}u)$ for every bounded continuous function on $\mathbb{R}$.

Looking at the special case $G(u) = \psi(u)/\phi_0(u)$ with $\psi \in C_c(\mathbb{R})$, we obtain $\int \psi \mathrm{d}m_N \to \int \psi \mathrm{d}m$ for all $\psi \in C_c(\mathbb{R})$. □

To apply the lemma to the proof of the LLT, we will need to find a strictly positive $\phi_0 \in \mathcal{K}$ so that $\int \phi_0 \mathrm{d}m$ is finite for the measure m which represents the limit. This is the purpose of the next Lemma. Parts (1) and (2) are needed for the LLT in this chapter, and part (3) will be used in the next chapter.

Lemma 5.12 *There exists a strictly positive $\phi_0 \in \mathcal{K}$ so that $\int \phi_0 \mathrm{d}m_i < \infty$, for $i \in \{1, 2, 3\}$, where*

(1) m_1 is Lebesgue's measure on $\mathbb{R}$; (2) m_2 is the counting measure on $t\mathbb{Z}$;

(3) m_3 is the measure representing the functional $m(\phi) = \sum_{k\in\mathbb{Z}} \mathbb{E}[\phi(k\delta + \mathfrak{F})]$ on $C_c(\mathbb{R})$, where δ is a positive constant, and $\mathfrak{F}$ is bounded random variable.

Proof Let $\psi_a(x) := (\frac{\sin ax}{ax})^2$, extended continuously to zero by $\psi_a(0) := 1$. This function is non-negative and absolutely integrable. To see that $\widehat{\psi_a}$ has compact support, we argue as follows: The Fourier transform of $1_{[-a,a]}$ is proportional to $\sin ax/ax$. Therefore the Fourier transform of $1_{[-a,a]} * 1_{[-a,a]}$ is proportional to ψ_a. Applying the inverse Fourier transform to ψ_a, we find that $\widehat{\psi_a}$ is proportional to $1_{[-a,a]} * 1_{[-a,a]}$, a function supported on $[-2a, 2a]$. Thus $\psi_a \in \mathcal{K}$. But ψ_a has zeroes at $\pi k/a$, $k \in \mathbb{Z} \setminus \{0\}$. To get a strictly positive element of $\mathcal{K}$, we take $\phi_0 := \psi_1 + \psi_{\sqrt{2}}$. Since $\phi_0(x) = O(|x|^{-2})$ as $|x| \to \infty$, $\int \phi_0 dm_i < \infty$. □

5.2.4 The LLT in the Irreducible Non-Lattice Case

We give the proof for arrays (Theorem 5.1'). Theorem 5.1 on chains follows, because every additive functional on a Markov chain is stably hereditary (Theorem 4.9).

We begin by proving that $V_N \xrightarrow[N\to\infty]{} \infty$. Otherwise $\liminf V_N < \infty$, and $\exists N_\ell \uparrow \infty$ such that $\mathrm{Var}(S_{N_\ell}) = O(1)$. Let X' denote the sub-array with rows $\mathsf{X}'^{(\ell)} = \mathsf{X}^{(N_\ell)}$. By Theorem 3.8, $\mathsf{f}|_{\mathsf{X}'}$ is center-tight, whence $G_{ess}(\mathsf{X}', \mathsf{f}|_{\mathsf{X}'}) = \{0\}$. But $G_{ess}(\mathsf{X}, \mathsf{f}) = G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, because f is irreducible and non-lattice. So $G_{ess}(\mathsf{X}', \mathsf{f}|_{\mathsf{X}'}) \neq G_{ess}(\mathsf{X}, \mathsf{f})$, in contradiction to the assumption that (X, f) is stably hereditary, see Theorem 4.13.

Next we fix $z_N \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, and show that for every non-empty interval (a, b), for every choice of $x_1^{(N)} \in \mathfrak{S}_1^{(N)}$ $(N \ge 1)$,

$$\mathbb{P}_{x_1^{(N)}}[S_N - z_N \in (a,b)] \sim \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N}}(b-a), \text{ as } N \to \infty. \tag{5.25}$$

By Lemma 5.11, we can prove (5.25) by showing that for every $\phi \in L^1(\mathbb{R})$ whose Fourier transform $\widehat{\phi}(\xi) := \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi u}\phi(u)\mathrm{d}u$ has compact support,

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}_{x_1^{(N)}}[\phi(S_N - z_N)] = \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\phi(u)\mathrm{d}u. \tag{5.26}$$

Fix $\phi \in L^1$ such that $\mathrm{supp}(\widehat{\phi}) \subseteq [-L, L]$. By the Fourier inversion formula, $\mathbb{E}_{x_1^{(N)}}(\phi(S_N - z_N)) = \frac{1}{2\pi}\int_{-L}^{L}\widehat{\phi}(\xi)\Phi_N(x_1^{(N)}, \xi)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi$. So (5.26) is equivalent to

$$\lim_{N\to\infty}\sqrt{V_N}\cdot\frac{1}{2\pi}\int_{-L}^{L}\widehat{\phi}(\xi)\Phi_N(x_1^{(N)},\xi)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi = \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\widehat{\phi}(0). \tag{5.27}$$

Below, we give a proof of (5.27).[1]

Divide $[-L, L]$ into segments I_j so that I_0 is centered at zero, and all segments have length less than or equal to $\widetilde{\delta}$, where $\widetilde{\delta}$ is given by Lemma 5.8. Let $J_{j,N} := \frac{1}{2\pi}\int_{I_j}\widehat{\phi}(\xi)\Phi_N(x_1^{(N)},\xi)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi$.

Claim 1:

$$\sqrt{V_N}J_{0,N} \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\widehat{\phi}(0). \tag{5.28}$$

Proof of the Claim. Fix $R > 0$. Since $I_0 \ni 0$, $A_N(I_0) = 0$. By Corollary 5.10, $|\Phi_N(x,\xi)| \le \widetilde{C}\exp(-\widehat{\varepsilon}\, V_N\xi^2)$ on $[-\widetilde{\delta},\widetilde{\delta}]$. So given $\varepsilon > 0$, there is $R > 0$ such that $\left|\frac{\sqrt{V_N}}{2\pi}\int_{\{\xi\in I_0:|\xi|>R/\sqrt{V_N}\}}\widehat{\phi}(\xi)\Phi_N(x_1^{(N)},\xi)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi\right| \le \varepsilon$.
Changing variables $\xi = s/\sqrt{V_N}$, we get $\frac{\sqrt{V_N}}{2\pi}\int\limits_{[|\xi|\le R/\sqrt{V_N}]}\widehat{\phi}(\xi)\Phi_N\left(x_1^{(N)},\xi\right)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi = \frac{1}{2\pi}\int\limits_{[|s|\le R]}\widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right)\mathbb{E}_{x_1^{(N)}}\left(\mathrm{e}^{\mathrm{i}s\frac{S_N - z_N}{\sqrt{V_N}}}\right)\mathrm{d}s$.

By Dobrushin's CLT for arrays (Theorem 3.10), $\frac{S_N - z_N}{\sqrt{V_N}}$ converges in distribution w.r.t. $\mathbb{P}_{x_1^{(N)}}$ to the normal distribution with mean $-z$ and variance 1. This implies that $\mathbb{E}_{x_1^{(N)}}\left(\mathrm{e}^{\mathrm{i}s\frac{S_N - z_N}{\sqrt{V_N}}}\right) \xrightarrow[N\to\infty]{} \mathrm{e}^{-\mathrm{i}sz - s^2/2}$ uniformly on compacts. Since $\widehat{\phi}$ is bounded and continuous at zero,

$$\frac{\sqrt{V_N}}{2\pi}\int\limits_{|\xi|\le R/\sqrt{V_N}}\widehat{\phi}(\xi)\Phi_N(x_1^{(N)},\xi)\mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{d}\xi = \frac{\widehat{\phi}(0)}{2\pi}\int_{-R}^{R}\mathrm{e}^{-\mathrm{i}sz}\mathrm{e}^{-s^2/2}\mathrm{d}s + o_{N\to\infty}(1).$$

Since this is true for all R, and $\int_{\mathbb{R}}\mathrm{e}^{-\mathrm{i}sz}\mathrm{e}^{-s^2/2}\mathrm{d}s = \sqrt{2\pi}\mathrm{e}^{-z^2/2}$, we have (5.28).

Claim 2: $\sqrt{V_N}J_{j,N} \xrightarrow[N\to\infty]{} 0$ *for* $j \neq 0$.

Proof of the Claim. Since $\widehat{\phi}$ is bounded, it is sufficient to show that

$$\sqrt{V_N}\int_{I_j}|\Phi(x_1^{(N)},\xi)|\mathrm{d}\xi \to 0. \tag{5.29}$$

[1] We note for future reference that this proof works for all ϕ such that $\widehat{\phi}$ is bounded, continuous at zero, and has compact support, including $\phi(u) = \frac{\sin(\pi u)}{\pi u}$ (whose Fourier transform is proportional to $1_{[-\pi,\pi]}$), which does not belong to L^1.

Recall that $A_N(I_j) = -\log\sup|\Phi_N(\cdot,\cdot)|$ on $\mathfrak{S}_1^{(N)} \times I_j$, and $(\widetilde{x}_{j,N}, \widetilde{\xi}_{j,N})$ are points where this supremum is achieved up to factor 2. Set $A_{j,N} := A_N(I_j)$.

Take large R and split I_j into two regions, $I'_{j,N} := \left\{\xi \in I_j : |\xi - \widetilde{\xi}_{j,N}| \le R\sqrt{\frac{A_{j,N}}{V_N}}\right\}$, $I''_{j,N} := I_j \setminus I'_{j,N}$. Split $\int_{I_j} |\Phi(x_1^{(N)}, \xi)|d\xi$ into two integrals $J'_{j,N}$, $J''_{j,N}$ accordingly.

- On $I'_{j,N}$, $|\Phi_N(x_1^{(N)}, \xi)| \le e^{-A_{j,N}}$ and $|I'_{j,N}| \le 2R\sqrt{\frac{A_{j,N}}{V_N}}$, so $\sqrt{V_N}|J'_{j,N}| \le 2R\sqrt{A_{j,N}}e^{-A_{j,N}}$.
- On $I''_{j,N}$, we have the following estimate, by Corollary 5.10, provided that $R\widehat{\varepsilon} > 2\overline{c}$.

$$|\Phi_N(x_1^{(N)}, \xi)| \le \widetilde{C}\exp\left(-\widehat{\varepsilon}V_N|\xi - \widetilde{\xi}_{j,N}|R\sqrt{\frac{A_{j,N}}{V_N}} + \overline{c}|\xi - \widetilde{\xi}_{j,N}|\sqrt{V_N A_{j,N}}\right) \le \widetilde{C}\exp\left(-\frac{\widehat{\varepsilon}}{2}|\xi - \widetilde{\xi}_{j,N}|\sqrt{A_{j,N}V_N}\right).$$

Hence $\sqrt{V_N}J''_{j,N} \le \sqrt{V_N}\widetilde{C}\int_{-\infty}^{\infty} e^{-\frac{\widehat{\varepsilon}}{2}|s|\sqrt{A_{j,N}V_N}}ds = O\left(A_{j,N}^{-\frac{1}{2}}\right)$. Combining these estimates, we obtain that

$$\sqrt{V_N}\|\Phi_N(x_1^{(N)}, \cdot)\|_{L^1(I_j)} \le 2R\sqrt{A_{j,N}}\, e^{-A_{j,N}} + O\left(A_{j,N}^{-1/2}\right) = O\left(A_{j,N}^{-1/2}\right). \tag{5.30}$$

We now employ the assumptions of the theorem: Firstly, f is irreducible with algebraic range $\mathbb{R}$, therefore the co-range of f is $\{0\}$ (Theorem 4.4). Secondly, f is stably hereditary, therefore $D_N(\xi) \xrightarrow[N\to\infty]{} \infty$ uniformly on compacts in $\mathbb{R} \setminus \{0\}$. By (5.18), $\Phi_N(x_1^{(N)}, \xi) \to 0$ uniformly on compacts in $\mathbb{R} \setminus \{0\}$, whence $A_{j,N} \to \infty$ as $N \to \infty$ for each $j \neq 0$. Thus (5.30) implies (5.29), and Claim 2.

Remark. Notice that (5.30) does not require the irreducibility assumption or the hereditary assumption. It holds for all uniformly elliptic arrays.

Claims 1 and 2 imply (5.27), and (5.27) implies (5.25) by Lemma 5.11. This proves the LLT for initial distributions concentrated at single points.

To deduce the theorem for arbitrary initial distributions, we can either appeal to Lemma 2.27, or prove the following claim and then integrate:

Claim 3: (5.25) *holds uniformly with respect to the choice of* $\{x_n^{(N)}\}$.

Proof of the Claim. Otherwise, there exist $\varepsilon > 0$ and $N_k \to \infty$ with $y_1^{(N_k)}$ such that $\mathbb{P}_{y_1^{(N_k)}}[S_{N_k} - z_{N_k} \in (a,b)]/\frac{e^{-z^2/2}(b-a)}{\sqrt{2\pi V_{N_k}}} \notin [e^{-\varepsilon}, e^{\varepsilon}]$. But this contradicts (5.25) for a sequence $\{x_1^{(N)}\}$ such that $x_1^{(N_k)} = y_1^{(N_k)}$. □

5.2.5 The LLT in the Irreducible Lattice Case

We give the proof in the context of arrays (Theorem 5.2'). Suppose X is a uniformly elliptic array, and f is an additive functional on X which is a.s. uniformly bounded, hereditary, irreducible, and with algebraic range $t\mathbb{Z}$ with $t > 0$. Without loss of generality, $t = 1$, otherwise work with $t^{-1}\mathsf{f}$. By Lemma 4.15, there are constants $c_n^{(N)}$ such that $f_n^{(N)}(X_n^{(N)}, X_{n+1}^{(N)}) - c_n^{(N)} \in \mathbb{Z}$ a.s. We may assume without loss of generality that $c_n^{(N)} = 0$, otherwise we work with $\mathsf{f} - \mathsf{c}$. So $S_N \in \mathbb{Z}$ a.s. for every $N \ge 1$.

We will show that for every sequence of integers z_N such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, and for every $x_1^{(N)} \in \mathfrak{S}_n^{(N)}$,

$$\mathbb{P}_{x_1^{(N)}}(S_N = z_N) = [1 + o(1)]\frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}, \text{ as } N \to \infty. \tag{5.31}$$

As in the non-lattice case, once we prove (5.31) for all choices of $\{x_1^{(N)}\}$, it automatically follows that (5.31) holds uniformly in $\{x_1^{(N)}\}$. Integrating over $\mathfrak{S}_1^{(N)}$ gives (5.4) with $k = 0$. For general k, take $z'_N := z_N + k$.

The assumptions on f imply that $V_N \xrightarrow[N\to\infty]{} \infty$, see the proof of Theorem 5.1'.

As we explained in §5.2.1, to prove (5.31) it is sufficient to show that

$$\lim_{N\to\infty} \sqrt{V_N} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_N(x_1^{(N)}, \xi) \mathrm{e}^{-\mathrm{i}\xi z_N} \mathrm{d}\xi = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}. \tag{5.32}$$

Notice that (5.32) is (5.27) in the case $\phi(u) = \frac{\sin(\pi u)}{\pi u}$, $\widehat{\phi}(\xi) = \text{const.}1_{[-\pi,\pi]}(\xi)$. It can be proved in almost exactly the same way.

Here is a sketch of the proof. One divides $[-\pi, \pi]$ into segments I_j of length less than the $\widetilde{\delta}$ of Lemma 5.8, so that one of the intervals contains zero in its interior.

The contribution of the interval which contains zero is asymptotic to $\frac{1}{\sqrt{2\pi V_N}} \mathrm{e}^{-z^2/2}$. This is shown as in Claim 1 of the preceding proof.

The remaining intervals are bounded away from $2\pi\mathbb{Z}$. Since f is irreducible with algebraic range $\mathbb{Z}$, $H(\mathsf{X}, \mathsf{f}) = 2\pi\mathbb{Z}$. Lattice hereditary f are stably hereditary, therefore $D_N(\xi) \xrightarrow[N\to\infty]{} 0$ uniformly on compacts in $\mathbb{R} \setminus 2\pi\mathbb{Z}$. Arguing as in the proof of Claim 2 of the preceding proof, one shows that the contribution of the intervals which do not contain zero is $o(1/\sqrt{V_N})$. □

5.2.6 Mixing LLT

The proof is very similar to the proof of the previous local limit theorems, except that it uses $\Phi(x, \xi|\mathfrak{A})$ instead of $\Phi(x, \xi)$. We outline the proof in the non-lattice case, and leave the lattice case to the reader.

Suppose X is a uniformly elliptic Markov array, and that f is a.s. uniformly bounded, stably hereditary, irreducible and with algebraic range $\mathbb{R}$. Let $\mathfrak{A}_N \in \mathfrak{S}_{k_N+1}^{(N)}$ be measurable sets such that $\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N) > \delta > 0$, and fix $x_N \in \mathfrak{S}_1^{(N)}$. Suppose $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$. As before, $V_N \to \infty$, and as explained in §5.2.3, it is enough to show that for every $\phi \in L^1(\mathbb{R})$ such that $\text{supp}(\widehat{\phi}) \subset [-L, L]$,

$$\lim_{N\to\infty} \sqrt{V_N} \cdot \frac{1}{2\pi} \int_{-L}^{L} \widehat{\phi}(\xi) \Phi_N(x_N, \xi|\mathfrak{A}_N) \mathrm{e}^{-\mathrm{i}\xi z_N} \mathrm{d}\xi = \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}} \widehat{\phi}(0).$$

Divide $[-L, L]$ as before into intervals I_j of length $\le \widetilde{\delta}$, where $\widetilde{\delta}$ is given by Lemma 5.8 and where I_0 is centered at zero. Let $J_{j,N} := \dfrac{1}{2\pi} \displaystyle\int_{I_j} \widehat{\phi}(\xi) \Phi_N(x_N, \xi|\mathfrak{A}_N) \mathrm{e}^{-\mathrm{i}\xi z_N} \mathrm{d}\xi$.

Claim 1: $\sqrt{V_N} J_{0,N} \xrightarrow[N\to\infty]{} (2\pi)^{-\frac{1}{2}} \mathrm{e}^{-z^2/2} \widehat{\phi}(0)$.

Proof of the Claim: Fix $R > 0$. As before, applying Corollary 5.10 with $A_N = 0$ we conclude that for each $\varepsilon > 0$ there is $R > 0$ such that $\left| \sqrt{V_N} \displaystyle\int_{\{\xi \in I_0 : |\xi| > R/\sqrt{V_N}\}} \widehat{\phi}(\xi) \Phi_N(x_N, \xi|\mathfrak{A}_N) \mathrm{e}^{-\mathrm{i}\xi z_N} \mathrm{d}\xi \right| \le \varepsilon$.

Next the change of variables $\xi = s/\sqrt{V_N}$ gives

$$\begin{aligned} &\sqrt{V_N} \int_{\{\xi \in I_0 : |\xi| \le R/\sqrt{V_N}\}} \widehat{\phi}(\xi) \Phi(x_N, \xi|\mathfrak{A}_N) \mathrm{e}^{-\mathrm{i}\xi z_N} \mathrm{d}\xi = \int_{-R}^{R} \widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right) \mathbb{E}_{x_N} \left(\mathrm{e}^{\mathrm{i}s \frac{S_N - z_N}{\sqrt{V_N}}} \middle| X_{k_N+1}^{(N)} \in \mathfrak{A}_N \right) \mathrm{d}s \\ &= \frac{1}{\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N)} \int_{-R}^{R} \widehat{\phi}\left(\frac{s}{\sqrt{V_N}}\right) \mathbb{E}_{x_N} \left(\mathrm{e}^{\mathrm{i}s \frac{S_N - z_N}{\sqrt{V_N}}} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)}) \right) \mathrm{d}s. \end{aligned} \tag{5.33}$$

We analyze the expectation in the integrand. Take $1 \le r_N \le k_N$ such that $r_N \to \infty$ and $r_N/\sqrt{V_N} \to 0$, and let

$$S_N^* := \sum_{j=1}^{k_N - r_N - 1} f_j^{(N)}(X_j^{(N)}, X_{j+1}^{(N)}) \equiv S_N - \sum_{j=k_N - r_N}^{k_N} f_j^{(N)}(X_j^{(N)}, X_{j+1}^{(N)}).$$

Since ess sup $|\mathsf{f}| < \infty$, $\|S_N - S_N^*\|_\infty = o(\sqrt{V_N})$, and so

$$\mathbb{E}_{x_N}\left(e^{is\frac{S_N - z_N}{\sqrt{V_N}}} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})\right) = \mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})\right) + o(1) = \mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}} \mathbb{E}\left(1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})|X_1^{(N)}, \ldots, X_{k_n - r_N}^{(N)}\right)\right) + o(1)$$

$$= \mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}} \mathbb{E}\left(1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})|X_{k_n - r_N}^{(N)}\right)\right) + o(1) \text{ by the Markov property.}$$

By (2.11), $\left\|\mathbb{E}(1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})|X_{k_n-r_N}^{(N)}) - \mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N)\right\|_\infty \le \text{const.}\theta^{r_N}$, where $0 < \theta < 1$. Since $r_N \to \infty$, we get

$$\mathbb{E}_{x_N}\left(e^{is\frac{S_N - z_N}{\sqrt{V_N}}} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})\right) = \mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}}\right)\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N) + o(1).$$

Since $r_N = o(\sqrt{V_N})$, $\dfrac{S_N^* - z_N}{\sqrt{V_N}} = \dfrac{S_N - z_N}{\sqrt{V_N}} + o(1) = \dfrac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} - z + o(1)$.
By Dobrushin's CLT (applied to the array with the transition probabilities of X, and the initial distributions $\pi^{(N)} := \delta_{x_N}$), we get $\mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}}\right) \sim \mathbb{E}_{x_N}\left(e^{is\frac{S_N - z_N}{\sqrt{V_N}}}\right) \xrightarrow[N\to\infty]{} e^{-isz - s^2/2}$. So

$\mathbb{E}_{x_N}\left(e^{is\frac{S_N^* - z_N}{\sqrt{V_N}}} 1_{\mathfrak{A}_N}(X_{k_N+1}^{(N)})\right) = e^{-isz-s^2/2}\mathbb{P}(X_{k_N+1}^{(N)} \in \mathfrak{A}_N) + o(1)$. Substituting this in (5.33) gives the claim.

Claim 2: $\sqrt{V_N} J_{j,N} \xrightarrow[N\to\infty]{} 0$ *for* $j \neq 0$.

This is similar to Claim 2 in §5.2.4. Instead of (5.18), use (5.19). □

5.3 Notes and References

For the history of the local limit theorem, see the end of Chapter 1.

The first statement of the LLT in terms of weak convergence of Radon measures is due to Shepp [183]. Stone's trick (the reduction to test functions with Fourier transforms with compact support) appears in [191]. This method was further developed and clarified by Breiman. Lemma 5.11 and its proof, are taken from Breiman's book [17, Chapter 10, §2].

The method of perturbation operators is due to Nagaev [149], who used it to prove central and local limit theorems for homogeneous Markov chains (see §8.4). Guivarc'h & Hardy used this method to prove LLT for Birkhoff sums generated by dynamical systems [88]. See also [168] where Rousseau-Egele used similar ideas in the study of interval maps. Hafouta & Kifer [93], Hafouta [89, 90], and Dragičević, Froyland, & González-Tokman [64], used this technique to prove the local limit theorem in different but related non-homogeneous settings.

The terminology "mixing LLT" is due to Rényi [166]. Mixing LLT have numerous applications including mixing of special flows [88], homogenization [60], and the study of skew-products [58, 57]. Mixing LLT for additive functionals of (stationary) Gibbs-Markov processes in the Gaussian and in the stable domains of attraction were proved by Aaronson & Denker [5]. Guivarc'h & Hardy noted the relevance of Mixing LLT to the study of reducible additive functionals, in the homogeneous case [88]. In the next chapter, we will use the mixing LLT to study the inhomogeneous reducible case.

Chapter 6
The Local Limit Theorem in the Reducible Case

Abstract We prove the local limit theorem for $\mathbb{P}(S_N - z_N \in (a,b))$ when $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{\mathrm{Var}(S_N)}}$ converges to a finite limit and f is reducible. In the reducible case, the asymptotic behavior of $\mathbb{P}(S_N - z_N \in (a,b))$ is not universal, and it depends on $f_n(X_n, X_{n+1})$. The dependence is strong for small intervals, and weak for large intervals.

6.1 Main Results

6.1.1 Heuristics and Warm Up Examples

Recall that an additive functional is called **reducible** if $\mathsf{f} = \mathsf{g} + \mathsf{c}$ where c is center-tight, and the algebraic range of g is strictly smaller than the algebraic range of f. By the results of Chapter 4, if $\mathrm{Var}(S_N(\mathsf{f})) \to \infty$, X is uniformly elliptic, and f is a.s. bounded, then we can choose g to be irreducible. In this case, $S_N(\mathsf{f}) = S_N(\mathsf{g}) + S_N(\mathsf{c})$, where $\mathrm{Var}(S_N(\mathsf{g})) \sim \mathrm{Var}(S_N(\mathsf{f})) \to \infty$, $\mathrm{Var}(S_N(\mathsf{c})) = O(1)$, and $S_N(\mathsf{g})$ satisfies the lattice local limit theorem. *But the contribution of $S_n(\mathsf{c})$ cannot be neglected.* In this chapter we give the corrections to the LLT needed to take $S_n(\mathsf{c})$ into account. Before stating our results in general, we discuss two simple examples which demonstrate some of the possible effects of $S_N(\mathsf{c})$.

Example 6.1 (Simple Random Walk with Continuous First Step) Suppose $\{X_n\}_{n\geq 1}$ are independent real-valued random variables, where X_1 has some distribution $\mathfrak{F}$, and X_i $(i\geq 2)$ are equal to 0, 1 with equal probabilities.

Suppose $0 \leq \mathfrak{F} < 1$, $\mathbb{E}[\mathfrak{F}] = \frac{1}{2}$, the distribution of $\mathfrak{F}$ has a density, and $\mathfrak{F}$ is *not* uniformly distributed on $[0,1]$. Let $\mu_{\mathfrak{F}}(\mathrm{d}x)$ denote the probability measure on $\mathbb{R}$ associated with the distribution of $\mathfrak{F}$.

Let $S_N := X_1 + \cdots + X_N$, then $S_N = S_N(\mathsf{f})$, where $f_n(x,y) := x$. Since the distribution of $\mathfrak{F}$ has a density, (X,f) has algebraic range $\mathbb{R}$. The following decomposition shows that (X,f) is reducible:

$$\mathsf{f} = \mathsf{g} + \mathsf{c}, \text{ where } g_n(x,y) := \begin{cases} 0 & n = 1 \\ x & n \geq 2 \end{cases} \text{ and } c_n(x,y) := \begin{cases} x & n = 1 \\ 0 & n \geq 2. \end{cases}$$

Clearly, g is irreducible with essential range $\mathbb{Z}$, and c is center-tight. $S_N(\mathsf{g})$, $S_N(\mathsf{c})$ are independent; $S_N(\mathsf{c}) \sim \mathfrak{F}$; $S_N(\mathsf{g})$ has the binomial distribution $B(\frac{1}{2}, N-1)$; and

$$S_N = (\underbrace{X_2 + \cdots + X_N}_{S_N(\mathsf{g})}) + \underbrace{X_1}_{S_N(\mathsf{c})}.$$

So S_N has distribution $\mu_{\mathfrak{F}} * B(\frac{1}{2}, N-1)$. This distribution has a density, which we denote by $p_N(x)\mathrm{d}x$.

(A) The Scaling Limit of $p_N(x)\mathrm{d}x$ is not Haar's Measure: Fix $z_N{:=}\mathbb{E}(S_N){=}N/2$ and let $V_N{:=}\mathrm{Var}(S_N){\sim}N/4$. The measure $m_N := p_N(x)\mathrm{d}x$ determines a positive functional on $C_c(\mathbb{R})$, and for every $\phi \in C_c(\mathbb{R})$ and N even,

$$\int \phi(x - z_N)p_N(x)\mathrm{d}x = \mathbb{E}[\phi(S_N - z_N)] = \mathbb{E}[\phi(S_N(\mathsf{g}) + S_N(\mathsf{c}) - z_N)] = \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(\mathfrak{F} + m - z_N)]\mathbb{P}[S_N(\mathsf{g}) = m]$$

$$= \sum_{m=0}^{N-1} \binom{N-1}{m} \frac{1}{2^{N-1}} \mathbb{E}[\phi(\mathfrak{F} + m - z_N)] = \frac{1}{2^{N-1}} \sum_{m=0}^{N-1} \binom{N-1}{m} \psi(m - \tfrac{N}{2}), \text{ where } \psi(m) := \mathbb{E}[\phi(\mathfrak{F} + m)]$$

$$= \frac{1}{2^{N-1}} \sum_{m=-N/2}^{N/2-1} \binom{N-1}{m + N/2} \psi(m) \sim \frac{1}{\sqrt{2\pi V_N}} \sum_{m\in\mathbb{Z}} \psi(m), \sim \frac{1}{\sqrt{2\pi V_N}} \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(\mathfrak{F} + m)], \text{ as } N \to \infty.$$

This also holds for N odd. Thus the distribution of $S_N - z_N$ converges weakly to zero, "at a rate of $1/\sqrt{V_N}$," and if we inflate it by $\sqrt{V_N}$ then it converges weakly to

$$\lambda := \mu_{\mathfrak{F}} * \frac{1}{\sqrt{2\pi}}(\text{the counting measure on } \mathbb{Z}).$$

By the assumptions on $\mathfrak{F}$, the scaling limit λ is not a Haar measure on a closed subgroup of $\mathbb{R}$. This is different from the irreducible case, when the scaling limit is a Haar measure on $G_{ess}(\mathsf{X}, \mathsf{f})$.

(B) Non-Standard Limit for $\sqrt{2\pi V_N}\mathbf{P}[S_N - \mathbf{E}(S_N) \in (a, b)]$: Fix $a, b \in \mathbb{R} \setminus \mathbb{Z}$ such that $|a - b| > 1$. Let $z_N := \mathbb{E}(S_N)$. The previous calculation with $\phi_i \in C_c(\mathbb{R})$ such that $\phi_1 \leq 1_{(a,b)} \leq \phi_2$ gives

$$\sqrt{2\pi V_N}\mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N\to\infty]{} \sum_{m\in\mathbb{Z}} \mathbb{E}[1_{(a,b)}(m + \mathfrak{F})]. \tag{6.1}$$

This is different from the limit in the irreducible non-lattice LLT (Theorem 5.1),

$$\sqrt{2\pi V_N}\mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N\to\infty]{} |a - b|; \tag{6.2}$$

and the limit in the irreducible lattice LLT with range $\mathbb{Z}$ (Theorem 5.2):

$$\sqrt{2\pi V_N}\mathbb{P}[S_N - z_N \in (a, b)] \xrightarrow[N\to\infty]{} \sum_{m\in\mathbb{Z}} 1_{(a,b)}(m). \tag{6.3}$$

(C) Robustness for Large Intervals: Although different, the limits in (6.1)–(6.3) are nearly the same, as $|a - b| \to \infty$. This is because all three limits belong to $[|a - b| - 2, |a - b| + 2]$.

Example 6.1 is very special in that $S_n(\mathsf{g}), S_N(\mathsf{c})$ are independent. Nevertheless, we will see below that (A), (B), (C) are general phenomena, which also happen when $S_N(\mathsf{g}), S_N(\mathsf{h})$ are strongly correlated.

The following example exhibits another common pathology:

Example 6.2 (Gradient Perturbation of the Lazy Random Walk) Suppose X_n, Y_n are independent random variables such that $X_n = -1, 0, +1$ with equal probabilities, and Y_n are uniformly distributed in $[0, 1/2]$. Let $\mathsf{X} = \{(X_n, Y_n)\}_{n\geq 1}$.

- The additive functional $g_n((x_n, y_n); (x_{n+1}, y_{n+1})) = x_n$ generates the lazy random walk on $\mathbb{Z}$, $S_N(\mathsf{g}) = X_1 + \cdots + X_N$. It is irreducible, and satisfies the lattice LLT.
- The additive functional $c_n((x_n, y_n), (x_{n+1}, y_{n+1})) = y_{n+1} - y_n$ is center-tight, and $S_N(\mathsf{c}) = Y_{N+1} - Y_1$.
- The sum $\mathsf{f} = \mathsf{g} + \mathsf{c}$ is reducible, with algebraic range $\mathbb{R}$ (because of c) and essential range $\mathbb{Z}$ (because of g). It generates the process
$$S_N(\mathsf{f}) = S_N(\mathsf{g}) + Y_{N+1} - Y_1.$$

Observe that $S_N(\mathsf{f})$ lies in a "random coset" $b_N + \mathbb{Z}$, where $b_N = Y_{N+1} - Y_1$.

Since the distribution of b_N is continuous, $\mathbb{P}[S_N - z_N = k] = 0$ for all $z_N, k \in \mathbb{Z}$, and the standard lattice LLT fails. To deal with this, we must "shift" $S_N - z_N$ back to $\mathbb{Z}$. This leads to the following (correct) statement: For all $z \in \mathbb{R}$ and $z_N \in \mathbb{Z}$ such that $\frac{z_N}{\sqrt{V_N}} \to z$, for all $k \in \mathbb{Z}$, $\mathbb{P}[S_N - z_N - b_N = k] \sim \dfrac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N}}$. Notice that the shift $b_N = Y_{N+1} - Y_1$ is *random.*

6.1.2 The LLT in the Reducible Case

Theorem 6.3 *Let $\mathsf{X} = \{X_n\}$ be a uniformly elliptic Markov chain, and let f be a reducible a.s. uniformly bounded additive functional with essential range $\delta(\mathsf{f})\mathbb{Z}$, where $\delta(\mathsf{f}) \neq 0$. Then there are bounded random variables $b_N = b_N(X_1, X_{N+1})$ and $\mathfrak{F} = \mathfrak{F}(X_1, X_2, \ldots)$ with the following properties:*

(1) For every $z \in \mathbb{R}$ *and* $z_N \in \delta(\mathsf{f})\mathbb{Z}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$*, for every* $\phi \in C_c(\mathbb{R})$*,*

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}\left[\phi(S_N - z_N - b_N)\right] = \frac{\delta(\mathsf{f})e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(m\delta(\mathsf{f}) + \mathfrak{F})].$$

(2) In addition, for every sequence of measurable sets $\mathfrak{A}_{N+1} \subset \mathfrak{S}_{N+1}$ *such that* $\mathbb{P}[X_{N+1} \in \mathfrak{A}_{N+1}]$ *is bounded below,*

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}\left[\phi(S_N - z_N - b_N)|X_1 = x, X_{N+1} \in \mathfrak{A}_{N+1}\right] = \frac{\delta(\mathsf{f})e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(m\delta(\mathsf{f}) + \mathfrak{F})|X_1 = x].$$

(3) $\|b_N\|_\infty \le 9\delta(\mathsf{f})$*, and* $\mathfrak{F} \in [0, \delta(\mathsf{f}))$*.*

Remark. We can remove the dependence of b_N and $\mathfrak{F}$ on X_1, and the conditioning on $X_1 = x$ in (2), using Lemma 2.27.

Theorem 6.3 may seem abstruse at first reading, and it is worthwhile to explain it in more detail.

• $\mathbb{E}\left[\phi(S_N - z_N - b_N)\right]$, as an element of $C_c(\mathbb{R})^*$, represents the Radon measure

$$m_N(E) = \mathbb{P}[S_N - z_N - b_N(X_1, X_{N+1}) \in E] \quad (E \in \mathscr{B}(\mathbb{R})).$$

This is the distribution of S_N, after a shift by $z_N + b_N(X_1, X_{N+1})$. The deterministic shift by z_N cancels most of the drift $\mathbb{E}(S_N)$; The random shift b_N brings S_N back to $\delta(\mathsf{f})\mathbb{Z}$. To understand why we need b_N, see Example 6.2.

• The limit

$$\mathscr{A}(\phi) := \delta(\mathsf{f}) \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(m\delta(\mathsf{f}) + \mathfrak{F})], \tag{6.4}$$

is also an element of $C_c(\mathbb{R})^*$. It represents the Radon measure $\mu_{\mathscr{A}} := \mu_{\mathfrak{F}} * m_{\delta(\mathsf{f})}$, where $\mu_{\mathfrak{F}}(E) = \mathbb{P}(\mathfrak{F} \in E)$ and $m_{\delta(\mathsf{f})} := \delta(\mathsf{f}) \times$ counting measure on $\delta(\mathsf{f})\mathbb{Z}$.

• Theorem 6.3(1) says that $m_N \to 0$ weakly at rate $1/\sqrt{V_N}$, and gives the scaling limit $\sqrt{V_N} m_N \xrightarrow[N\to\infty]{w} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \mu_{\mathscr{A}}$. In particular, for all $a < b$ such that $\mathfrak{F}$ has no atoms in $\{a, b\} + \delta(\mathsf{f})\mathbb{Z}$, and for all $z \in \mathbb{R}$ and $z_N \in \delta(\mathsf{f})\mathbb{Z}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, we have $\mathbb{P}[S_N - z_N - b_N \in (a, b)] = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \mu_{\mathscr{A}}((a, b))$.

• On one hand, $\mu_{\mathscr{A}}(a, b) \sim |a - b|$ as $|a - b| \to \infty$ (Lemma 6.10). But for small $(a, b) \subset [0, \delta(\mathsf{f}))$, $\mu_{\mathscr{A}}(a, b) = \mathbb{P}[\mathfrak{F} \in (a, b) \bmod \delta(\mathsf{f})\mathbb{Z}]$, which could be quite different from $|a - b|$.

In summary, Theorem 6.3(1) gives the necessary "correction" to the classical LLT (5.1) in the reducible case. This correction is significant for intervals with length $\le \delta(\mathsf{f})$, and becomes less and less significant, as $|a - b|/\delta(\mathsf{f}) \to \infty$.

Theorem 6.3(2) is a "mixing" version of part (1), in the sense of §5.1.3. Such results are particularly useful in the reducible setup for the following reason. The random shift $b_N(X_1, X_{N+1})$ is sometimes a nuisance, and it is tempting to turn it into a deterministic quantity by conditioning on X_1, X_{N+1}. We would have liked to say that part (1) survives such conditioning, but we cannot. The best we can say in general is that part (1) remains valid under conditioning of the form $X_1 = x_1, X_{N+1} \in \mathfrak{A}_{N+1}$ *provided* $\mathbb{P}(X_{N+1} \in \mathfrak{A}_{N+1})$ *is bounded below.* This the content of part (2). For an example how to use such a statement, see §6.2.3.

The LLT in this section is only stated for Markov *chains*. The reason is that to construct $\mathfrak{F}$, we need the joint distribution of $(X_1, X_2, \ldots)$, which is not defined for Markov arrays. For (weaker) results for arrays, see §6.2.5.

6.1.3 Irreducibility as a Necessary Condition for the Mixing LLT

We can use Theorem 6.3 to clarify the necessity of the irreducibility condition for the "classical" LLT expansions in Theorems 5.1 and 5.4. We begin with an example showing that strictly speaking, irreducibility is not necessary:

Example 6.4 Take $S_N = X_1 + X_2 + X_3 + \cdots + X_N$ where X_i are independent, $X_1 \equiv 0$, X_2 is uniformly distributed on $[0,1]$, and $X_3, X_4, \ldots = 0, 1$ with equal probabilities.

Conditioned on X_1 and X_N, S_N is the sum of a constant plus two independent random variables, one, $\mathbf{U}$, is uniformly distributed on $[0,1]$, and the other, $\mathbf{B}_N$, has the binomial distribution $B(\frac{1}{2}, N-2)$. It is not difficult to see, using the De-Moivre-Laplace LLT for the binomial distribution, that for every $z, z_N \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, $\forall \phi \in C_c(\mathbb{R})$, $\mathbb{E}[\phi(S_N - \mathbb{E}(S_N))|X_1, X_N] \sim \dfrac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \sum_{m \in \mathbb{Z}} \mathbb{E}[\phi(m + \mathbf{U})] \overset{!}{=} \dfrac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \int \phi(u) du.$

So the "classical" non-lattice mixing LLT (Theorem 5.4) holds, even though our additive functional is reducible, with algebraic range $\mathbb{R}$ and essential range $\mathbb{Z}$.

Of course, the identity marked by (!) is a "coincidence," due to the particular choice of X_2. If we change X_2, it need not be valid anymore.

The next result says that irreducibility *is* a necessary condition for the mixing LLT, provided we impose the mixing LLT not just for (X, f), but also for all $(\mathsf{X}', \mathsf{f}')$ obtained from (X, f) by deleting finitely many terms.

Let f be an additive functional on a Markov chain X. Denote the state spaces of X by $\mathfrak{S}_n$, and write $\mathsf{X} = \{X_n\}_{n \geq 1}$, $\mathsf{f} = \{f_n\}_{n \geq 1}$. A sequence of events $\mathfrak{A}_k \subset \mathfrak{S}_k$ is called **regular** if $\mathfrak{A}_k$ are measurable, and $\mathbb{P}(X_n \in \mathfrak{A}_n)$ is bounded away from zero.

- We say that (X, f) satisfies the **mixing non-lattice LLT** if $V_N := \mathrm{Var}(S_N) \to \infty$, and for every regular sequence of events $\mathfrak{A}_n \in \mathscr{B}(\mathfrak{S}_n)$, $x \in \mathfrak{S}_1$, for all $z_N, z \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, and for each non-empty interval (a,b), $\mathbb{P}_x\Big(S_N - z_N \in (a,b) | X_{N+1} \in \mathfrak{A}_{N+1}\Big) = [1 + o(1)] \dfrac{e^{-z^2/2}}{\sqrt{2\pi V_N}} |a - b|$ as $N \to \infty$.
- Fix $t > 0$. We say that (X, f) satisfies the **mixing mod t LLT**, if for every regular sequence of events $\mathfrak{A}_n \in \mathscr{B}(\mathfrak{S}_n)$, $x \in \mathfrak{S}_1$, and for every non-empty interval (a,b) with length less than t,

$$\mathbb{P}_x(S_N \in (a,b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N \to \infty]{} \frac{|a-b|}{t}. \tag{6.5}$$

Theorem 6.5 *Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Given m, let $(\mathsf{X}_m, \mathsf{f}_m) := (\{X_n\}_{n \geq m}, \{f_n\}_{n \geq m})$. The following are equivalent:*
(1) f is irreducible with algebraic range $\mathbb{R}$; (2) $(\mathsf{X}_m, \mathsf{f}_m)$ satisfy the mixing non-lattice LLT for all m;
(3) $(\mathsf{X}_m, \mathsf{f}_m)$ satisfy the mixing mod t LLT for all m and t.

6.1.4 Universal Bounds for Prob$[S_N - z_N \in (a,b)]$

So far, we have considered the problem of finding $\mathbb{P}[S_N - z_N \in (a,b)]$ up to asymptotic equivalence, and subject to assumptions on the algebraic and essential range. We now consider the problem of estimating these probabilities up to a bounded multiplicative error, but only assuming that $V_N \to \infty$.

We already saw that the predictions of the LLT for *large* intervals (a,b) are nearly the same in all non center-tight cases, reducible or irreducible, lattice or non-lattice. Therefore it is reasonable to expect universal lower and upper bounds, for all sufficiently large intervals, and without further assumptions on the arithmetic structure of the range. The question is how large is "sufficiently large."

Define the **graininess constant** of (X, f) to be

$$\delta(\mathsf{f}) := \begin{cases} t & G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}, \quad t > 0 \\ 0 & G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R} \\ \infty & G_{ess}(\mathsf{X}, \mathsf{f}) = \{0\}. \end{cases} \tag{6.6}$$

By Corollary 4.6, if (X, f) is uniformly elliptic and $V_N \to \infty$, then $\delta(\mathsf{f}) \leq 6 \mathrm{ess\ sup}\, |\mathsf{f}|$.

We certainly cannot expect universal lower and upper bounds for intervals with length smaller than $\delta(\mathsf{f})$, because such intervals may fall in the gaps of the support of $S_N - z_N$. However, universal bounds do apply as soon as $|a-b| > \delta(\mathsf{f})$:

Theorem 6.6 *Suppose* f *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*. Then for every interval* (a,b) *of length* $L > \delta(\mathsf{f})$*, for all* $\epsilon > 0$ *and* $z_N, z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$*, for all* N *large enough,*

$$\mathbb{P}(S_N - z_N \in (a,b)) \le \frac{\mathrm{e}^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\left(1 + \frac{21\delta(\mathsf{f})}{L} + \epsilon\right), \tag{6.7}$$

$$\mathbb{P}(S_N - z_N \in (a,b)) \ge \frac{\mathrm{e}^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\left(1 - \frac{\delta(\mathsf{f})}{L} - \epsilon\right). \tag{6.8}$$

In addition, if $0 < \delta(\mathsf{f}) < \infty$ *and* $k\delta(f) \lneqq L \lneqq (k+1)\delta(f)$*,* $k \in \mathbb{N}$*, then*

$$\left(\frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N}}\right) k\delta(f) \lesssim \mathbb{P}(S_N - z_N \in (a,b)) \lesssim \left(\frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N}}\right)(k+1)\delta(f). \tag{6.9}$$

Here $A_N \lesssim B_N$ *means that* $\limsup_{N\to\infty}(A_N/B_N) \le 1$.

Notice that the theorem makes no assumptions on the irreducibility of f, although it does become vacuous in the center-tight case, when $\delta(\mathsf{f}) = \infty$. Note also that the bounds in this theorem are sharp in the limit $L \to \infty$.

Theorem 6.6 is an easy corollary of Theorem 6.3, see §6.2.4, but this is an overkill. §6.2.5 gives another proof of similar estimates, for intervals of length $L > 2\delta(\mathsf{f})$, which does not require the full force of Theorem 6.3, and which also applies to arrays. There we will also see that anti-concentration inequalities similar to (6.7) hold without any assumptions on z_N.

6.2 Proofs

6.2.1 Characteristic Functions in the Reducible Case

Throughout this section we assume that $\mathsf{X} = \{X_n\}$ is a uniformly elliptic Markov *chain* with state spaces $\mathfrak{S}_n$, marginals $\mu_n(E) := \mathbb{P}(X_n \in E)$, and transition probabilities $\pi_{n,n+1}(x, \mathrm{d}y) = p_n(x,y)\mu_{n+1}(\mathrm{d}y)$, with the uniform ellipticity condition, and ellipticity constant ϵ_0. Given $\varphi \in L^\infty(\mathfrak{S}_n \times \mathfrak{S}_{n+1})$, we let $\mathbb{E}(\varphi) := \mathbb{E}[\varphi(X_n, X_{n+1})]$ and $\sigma(\varphi) := \sqrt{\mathrm{Var}(\varphi(X_n, X_{n+1}))}$.

As explained in §5.2.1, it is possible to express the LLT probabilities in terms of

$$\Phi_N(x,\xi|\mathfrak{A}_{N+1}) = \mathbb{E}_x(\mathrm{e}^{\mathrm{i}\xi S_N}|X_{N+1} \in \mathfrak{A}_{N+1}) \equiv \frac{\mathbb{E}_x(\mathrm{e}^{\mathrm{i}\xi S_N}1_{\mathfrak{A}_{N+1}})}{\mathbb{E}_x(1_{\mathfrak{A}_{N+1}})}.$$

It will transpire that $\Phi_N(x,\xi|\mathfrak{A}_{N+1})$ decays fast to zero for ξ bounded away from the co-range $H(\mathsf{X},\mathsf{f})$. What matters is the behavior for ξ within distance $O\left(\frac{1}{\sqrt{V_N}}\right)$ from $H(\mathsf{X},\mathsf{f})$. Our aim in this section is to estimate $\Phi_N(x,\xi|\mathfrak{A}_{N+1})$ for such ξ.

We begin with an estimate for *arrays.* Suppose $\widehat{C}, K > 0$, $\epsilon \in (0,1)$ and f is an *array* of real-valued functions $f_n^{(N)} \in L^\infty(\mathfrak{S}_n \times \mathfrak{S}_{n+1})$ $(1 \le n \le N)$ as follows:

(I) $\mathbb{E}(f_n^{(N)}) = 0$ and ess sup $|\mathsf{f}| < K$.

(II) Let $S_N := \sum_{n=1}^{N} f_n^{(N)}(X_n, X_{n+1})$ and $V_N := \mathrm{Var}(S_N)$, then

$$V_N \to \infty \quad \text{and} \quad \frac{1}{V_N}\sum_{n=1}^{N}\sigma^2(f_n^{(N)}) \leq \widehat{C}. \tag{6.10}$$

(III) $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$, where

(a) $\mathbb{F}_n^{(N)}$ are measurable, integer valued, and $\operatorname{ess\,sup}|\mathbb{F}| < K$.

(b) $h_n^{(N)}$ are measurable functions such that $\mathbb{E}(h_n^{(N)}) = 0$, $\operatorname{ess\,sup}|\mathsf{h}| < K$ and $\sum_{n=1}^{N}\sigma^2(h_n^{(N)}) \leq \epsilon$.

(c) $c_n^{(N)}$ are constants. Necessarily $|c_n^{(N)}| \leq 3K$ and $c_n^{(N)} = -\mathbb{E}(\mathbb{F}_n^{(N)})$. Let $c^{(N)} := \sum_{n=1}^{N} c_n^{(N)}$.

We are *not* assuming that $\mathbb{E}(\mathbb{F}_n^{(N)}) = 0$: $\mathbb{F}_n^{(N)}$ are integer-valued, and we do not wish to destroy this by subtracting the mean.

Lemma 6.7 *Suppose (I), (II) and (III). For every $\overline{K} > 0$ and $m \in \mathbb{Z}$, there are $\overline{C}, \overline{N} > 0$ such that for every $N > \overline{N}$, $|s| \leq \overline{K}$, $x \in \mathfrak{S}_1$, and $v_{N+1} \in L^\infty(\mathfrak{S}_{N+1})$ with $\|v_{N+1}\|_\infty \leq 1$,*

$$\mathbb{E}_x\left(e^{i\left(2\pi m + \frac{s}{\sqrt{V_N}}\right)S_N} v_{N+1}(X_{N+1})\right) = e^{2\pi i m c^{(N)} - (s^2/2)}\mathbb{E}(v_{N+1}(X_{N+1})) + \eta_N(x), \text{ where } \mathbb{E}(|\eta|) \leq \overline{C}\left[\sum_{n=1}^{N}\sigma^2(h_n^{(N)})\right]^{1/2} \leq \overline{C}\sqrt{\epsilon}.$$

Proof We will use Nagaev's perturbation operators, as in §5.2.2.

Throughout this proof, we fix the value of N, and drop the superscripts $^{(N)}$. For example $c_k^{(N)} = c_k$ and $c^{(N)} = c$.

Define $\mathcal{L}_{n,\xi} : L^\infty(\mathfrak{S}_{n+1}) \to L^\infty(\mathfrak{S}_n)$ by $(\mathcal{L}_{n,\xi}u)(x) = \int_{\mathfrak{S}_{n+1}} p_n(x,y)e^{i\xi f_n(x,y)}u(y)\mu_{n+1}(dy)$, $\mathcal{L}_n := \mathcal{L}_{n,0}$.

Let $\xi = \xi(m,s) := 2\pi m + \frac{s}{\sqrt{V_N}}$, with $m \in \mathbb{Z}$ fixed and $|s| \leq \overline{K}$. Since $\mathbb{F}_n(x,y) \in \mathbb{Z}$,

$$e^{i\xi f_n} = \exp[2\pi i m\mathbb{F}_n + \tfrac{is}{\sqrt{V_N}}\mathbb{F}_n + i\xi c_n + i\xi h_n] = e^{2\pi i m c_n} e^{\frac{is}{\sqrt{V_N}}(\mathbb{F}_n + c_n) + i\xi h_n}.$$

We split $e^{-2\pi i m c_n}\mathcal{L}_{n,\xi} = \overline{\mathcal{L}}_{n,\xi} + \widehat{\mathcal{L}}_{n,\xi} + \widetilde{\mathcal{L}}_{n,\xi}$, where

$$\left(\overline{\mathcal{L}}_{n,\xi}u\right)(x) = \int_{\mathfrak{S}_{n+1}} p_n(x,y)e^{\frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x,y)+c_n)}u(y)\mu_{n+1}(dy),$$
$$\left(\widehat{\mathcal{L}}_{n,\xi}u\right)(x) = i\xi\int_{\mathfrak{S}_{n+1}} p_n(x,y)h_n(x,y)u(y)\mu_{n+1}(dy), \text{ and}$$
$$\left(\widetilde{\mathcal{L}}_{n,\xi}u\right)(x) = \int_{\mathfrak{S}_{n+1}} p_n(x,y)\left[e^{i\xi h_n + \frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x,y)+c_n)} - e^{\frac{is}{\sqrt{V_N}}(\mathbb{F}_n(x,y)+c_n)} - i\xi h_n(x,y)\right]u(y)\mu_{n+1}(dy).$$

Preparatory Estimates 1: $\exists C_1(\overline{K}, m) > 1$ *such that for every $|s| \leq \overline{K}$ and $n \geq 1$,*

$$\left\|\mathcal{L}_{n,\xi}\right\| := \left\|\mathcal{L}_{n,\xi}\right\|_{L^\infty\to L^\infty} \leq 1, \tag{6.11}$$
$$\left\|\mathcal{L}_{n,\xi}\right\|_{L^1\to L^\infty} \leq C_1(\overline{K}, m), \tag{6.12}$$
$$\left\|\overline{\mathcal{L}}_{n,\xi}\right\| := \left\|\overline{\mathcal{L}}_{n,\xi}\right\|_{L^\infty\to L^\infty} \leq 1, \tag{6.13}$$
$$\left\|\widehat{\mathcal{L}}_{n,\xi}\right\|_{L^\infty\to L^1} \leq C_1(\overline{K}, m)\sigma(h_n), \tag{6.14}$$
$$\left\|\widetilde{\mathcal{L}}_{n,\xi}\right\|_{L^\infty\to L^1} \leq C_1(\overline{K}, m)\left[\sigma^2(h_n) + \frac{\sigma(h_n)\sigma(f_n)}{\sqrt{V_N}}\right]. \tag{6.15}$$

Proof of the Step. The first three estimates are trivial.

To see (6.14), we note that $\widehat{\mathcal{L}}_{n,\xi}$ is an integral operator whose kernel has absolute value $|\mathrm{i}\xi p_n(x,y)h_n(x,y)| \leq \epsilon_0^{-1}|\xi||h_n(x,y)|$. So $\|\widehat{\mathcal{L}}_{n,\xi}\|_{L^\infty\to L^1} \leq \epsilon_0^{-1}|\xi|\|h_n\|_{L^1} \leq \epsilon_0^{-1}(2\pi|m| + \overline{K}V_N^{-1/2})\|h_n\|_{L^2}$, and (6.14) follows from the identity $\|h_n\|_{L^2} \equiv \sigma(h_n)$, and the assumption $V_N \to \infty$.

Similarly, $\widetilde{\mathcal{L}}_{n,\xi}$ is an integral operator with kernel $\widetilde{k}(x,y)$ such that

$$\begin{aligned}|\widetilde{k}(x,y)| &\leq \epsilon_0^{-1}|\mathrm{e}^{\mathrm{i}s\frac{\mathbb{F}_n(x,y)+c_n}{\sqrt{V_N}}}(\mathrm{e}^{\mathrm{i}\xi h_n}-1) - \mathrm{i}\xi h_n| = \epsilon_0^{-1}|\mathrm{e}^{\mathrm{i}s\frac{\mathbb{F}_n(x,y)+c_n}{\sqrt{V_N}}}(\mathrm{i}\xi h_n + O(\xi^2 h_n^2)) - \mathrm{i}\xi h_n| \\ &= \epsilon_0^{-1}|\mathrm{e}^{\mathrm{i}s\frac{\mathbb{F}_n(x,y)+c_n}{\sqrt{V_N}}} - 1||\xi h_n| + O\left(h_n^2\right) = O(\tfrac{1}{\sqrt{V_N}}|h_n(\mathbb{F}_n + c_n)|) + O(h_n^2).\end{aligned}$$

The big Oh's are uniform on compact sets of ξ.

So, uniformly on compact sets of ξ,

$$\|\widetilde{\mathcal{L}}_{n,\xi}\|_{L^\infty\to L^1} = O(V_N^{-1/2})\mathbb{E}(|h_n(\mathbb{F}_n + c_n)|) + O(\|h_n\|_2^2) = O(V_N^{-1/2})\|h_n\|_2\|\mathbb{F}_n + c_n\|_2 + O(\|h_n\|_2^2).$$

By (III), $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$, so

$$\begin{aligned}\|\widetilde{\mathcal{L}}_{n,\xi}\|_{L^\infty\to L^1} &= O(V_N^{-1/2})\|h_n\|_2\|f_n - h_n\|_2 + O(\|h_n\|_2^2) = O(V_N^{-1/2})\|h_n\|_2(\|f_n\|_2 + \|h_n\|_2) + O(\|h_n\|_2^2) \\ &= O\left(\frac{\|h_n\|_2\|f_n\|_2}{\sqrt{V_N}} + \|h_n\|_2^2\right) = O\left(\frac{\sigma(h_n)\sigma(f_n)}{\sqrt{V_N}} + \sigma^2(h_n)\right),\end{aligned}$$

as claimed in (6.15).

Preparatory Estimates 2. *There is a constant $C_1'(\overline{K},m)$ such that for all k,*

$$\|\widehat{\mathcal{L}}_{k,\xi}(\overline{\mathcal{L}}_{k+1,\xi}\overline{\mathcal{L}}_{k+2,\xi} - \mathcal{L}_{k+1}\mathcal{L}_{k+2})\|_{L^\infty\to L^1} \leq C_1'(\overline{K},m)\frac{\sigma(h_k)}{\sqrt{V_N}}(\sigma(f_{k+1}) + \sigma(f_{k+2}) + \sigma(h_{k+1}) + \sigma(h_{k+2}))\,. \quad (6.16)$$

Next, let $\varphi_k := \mathcal{L}_k\widehat{\mathcal{L}}_{k+1,\xi}1$, then

$$\mathbb{E}[\varphi_k(X_k)] = 0\,,\ \|\varphi_k\|_\infty \leq C_1'(\overline{K},m)^2\sigma(h_{k+1}). \quad (6.17)$$

Proof. Call the operator on the LHS of (6.16) $\mathcal{K}$. Then $(\mathcal{K}u)(x) = \iiint k(x,y,z,w)u(w)\mu_{k+1}(\mathrm{d}y)\mu_{k+2}(\mathrm{d}z)\mu_{k+3}(\mathrm{d}w)$ where $k(x,y,z,w)$ equals $p_k(x,y)p_{k+1}(y,z)p_{k+2}(z,w)\cdot \mathrm{i}\xi h_k(x,y)\left(\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}(\mathbb{F}_{k+1}(y,z)+\mathbb{F}_{k+2}(z,w)+c_{k+1}+c_{k+2})} - 1\right)$.

Recall from (III) that $\mathbb{F} + \mathsf{c} = \mathsf{f} - \mathsf{h}$. Therefore

$$|k(x,y,z,w)| \leq Cp_k(x,y)p_{k+1}(y,z)p_{k+2}(z,w)V_N^{-1/2}|h_k(x,y)|(|f_{k+1}(y,z)| + |f_{k+2}(z,w)| + |h_{k+1}(y,z)| + |h_{k+2}(z,w)|),$$

and the constant depends only on $K, \overline{K}$ and m.

Clearly, $\|\mathcal{K}\|_{L^\infty\to L^1} \leq \iiiint |k(x,y,z,w)|\mu_k(\mathrm{d}x)\mu_{k+1}(\mathrm{d}y)\mu_{k+2}(\mathrm{d}z)\mu_{k+3}(\mathrm{d}w)$. Integrating and applying the Cauchy-Schwarz inequality, we arrive at (6.16).

We continue to (6.17). By definition, $(\mathcal{L}_k u)(X_k) = \mathbb{E}(u(X_{k+1})|X_k)$. So

$$\mathbb{E}(\varphi_k) = \mathbb{E}((\widehat{\mathcal{L}}_{k+1,\xi}1)(X_{k+1})) = \mathrm{i}\xi\iint p_{k+1}(x,y)h_{k+1}(x,y)\mu_{k+1}(\mathrm{d}x)\mu_{k+2}(\mathrm{d}y).$$

Thus $\mathbb{E}(\varphi_k) = \mathrm{i}\xi\mathbb{E}(h_{k+1}) = 0$. Next, by (6.12) and (6.14),

$$\|\varphi_k\|_\infty \leq \|\mathcal{L}_k\|_{L^1\to L^\infty}\|\widehat{\mathcal{L}}_{k+1,\xi}\|_{L^\infty\to L^1} \leq C_1(\overline{K},m)^2\sigma(h_{k+1}).$$

Preparatory Estimates 3: *Let $\mathcal{L}_{k,\xi}' := \mathcal{L}_{k,\xi} - \mathrm{e}^{2\pi\mathrm{i}mc_k}\mathcal{L}_k$. There is a constant $C_1''(\overline{K},m)$ such that for all k,*

$$\|\mathcal{L}'_{k,\xi}\|_{L^\infty\to L^1} \le C_1''\left(\frac{\sigma(f_k)}{\sqrt{V_N}}+\sigma(h_k)\right), \tag{6.18}$$

$$\|\mathcal{L}'_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}1\|_1 \le C_1''\sigma(h_k)\left(\frac{\sigma(f_{k-1})}{\sqrt{V_N}}+\sigma(h_{k-1})\right). \tag{6.19}$$

Proof of the Step. $\mathcal{L}'_{k,\xi}$ is an integral operator with kernel

$$\mathfrak{l}(x,y) := \mathrm{e}^{2\pi \mathrm{i} m c_k} p_k(x,y)\left(\mathrm{e}^{\mathrm{i}\xi f_k(x,y)-2\pi\mathrm{i} m c_k}-1\right) = \mathrm{e}^{2\pi\mathrm{i} m c_k} p_k(x,y)\left(\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}(f_k(x,y)-h_k(x,y))+\mathrm{i}\xi h_k(x,y)}-1\right)$$

because by (III), $\mathrm{i}\xi f_k - 2\pi\mathrm{i} m c_k = \frac{\mathrm{i}s}{\sqrt{V_N}}(f_k-h_k)+\mathrm{i}\xi h_k \bmod 2\pi\mathbb{Z}$. Thus $|\mathfrak{l}(x,y)| \le C p_k(x,y)\left(\frac{|f_k(x,y)|}{\sqrt{V_n}}+|h_k(x,y)|\right)$.

(6.18) can now be shown as in the proof of (6.16).

The bound (6.19) has a similar proof, which we leave to the reader.

Preparatory Estimate 4: *The following estimates hold uniformly in $x \in \mathfrak{S}_1$:*

$$\mathbb{E}(S_N|X_1 = x) = \mathbb{E}(S_N) + O(1) = O(1)\,,\ \operatorname{Var}(S_N|X_1 = x) = V_N + O(1). \tag{6.20}$$

Proof of the Step. By (2.11), $|\mathbb{E}(S_N|X_1 = x) - \mathbb{E}(S_N)| = O(1)$, and the estimate on $\mathbb{E}(S_N|X_1 = x)$ follows from the assumption that $\mathbb{E}(f_n) = 0$ for all n.

Let $\mathsf{Y}_x := (\mathsf{X}$ conditioned on $X_1 = x)$. These Markov chains are all uniformly elliptic with the ellipticity constant of X. So, uniformly in x, for all $1 \le i < j \le N$,

$$\mathbb{E}_x(f_i^2) = \mathbb{E}(f_i^2) + O(\theta^i), \quad \mathbb{E}_x(f_i f_j) = \mathbb{E}_x(f_i)\mathbb{E}_x(f_j) + O(\theta^{j-i}) = [\mathbb{E}(f_i) + O(\theta^i)][\mathbb{E}(f_j) + O(\theta^j)] + O(\theta^{j-i})$$

$$\overset{!}{=} \mathbb{E}(f_i)\mathbb{E}(f_j) + O(\theta^{j-i}) = \mathbb{E}(f_i f_j) + O(\theta^{j-i})\ \ (\because \mathbb{E}(f_i) = \mathbb{E}(f_j) = 0).$$

$$\mathbb{E}_x(f_i f_j) = \mathbb{E}(g_{ij}|X_1 = x),\ \text{where } g_{ij} := \mathbb{E}(f_i f_j|X_{i+1},\ldots,X_1)$$

$$\overset{!}{=} \mathbb{E}(g_{ij}) + O(\theta^i) = \mathbb{E}(f_i f_j) + O(\theta^i)\ \ (\because g_{ij} \text{ depends only on } X_i, X_{i+1}).$$

Therefore $|\mathbb{E}_x(f_i f_j) - \mathbb{E}(f_i f_j)|=O(\min\{\theta^i,\theta^{j-i}\}\}) = O(\theta^{i/2}\theta^{(j-i)/2})$. Summing over $1 \le i, j \le N$, we obtain

$$|\mathbb{E}_x(S_N^2) - \mathbb{E}(S_N^2)| \le \sum |\mathbb{E}_x(f_i f_j) - \mathbb{E}(f_i f_j)| = O(1).$$

Since $\mathbb{E}(S_N) = 0$ and $\mathbb{E}_x(S_N) = O(1)$, $\operatorname{Var}(S_N|X_1 = x) = \operatorname{Var}(S_N) + O(1)$.

We are ready to start the proof of the lemma. By (5.12), $\mathbb{E}_x[\mathrm{e}^{\mathrm{i}\xi S_N} v_{N+1}(X_{N+1})] = (\mathcal{L}_{1,\xi}\mathcal{L}_{2,\xi}\cdots\mathcal{L}_{N,\xi}v_{N+1})(x)$.

The decomposition $\mathrm{e}^{-2\pi\mathrm{i} m c_n}\mathcal{L}_{n,\xi} = \overline{\mathcal{L}}_{n,\xi} + \widehat{\mathcal{L}}_{n,\xi} + \widetilde{\mathcal{L}}_{n,\xi}$ implies that

$$\mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}\xi S_N} v_{N+1}(X_{N+1})\right) = \mathrm{e}^{2\pi i m c}\left(\overline{\Phi}_N(x,\xi) + \widehat{\Phi}_N(x,\xi) + \widetilde{\Phi}_N(x,\xi)\right) \tag{6.21}$$

where $c = c^{(N)} = c_1 + \cdots + c_N$, and

$$\overline{\Phi}_N(x,\xi) := \left(\overline{\mathcal{L}}_{1,\xi}\cdots\overline{\mathcal{L}}_{N,\xi}v_{N+1}\right)(x),$$

$$\widetilde{\Phi}_N(x,\xi) := \sum_{k=1}^{N-1}\mathrm{e}^{-2\pi\mathrm{i} m(c_1+\cdots+c_{k-1})}\left(\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widetilde{\mathcal{L}}_{k,\xi}\overline{\mathcal{L}}_{k+1,\xi}\cdots\overline{\mathcal{L}}_{N,\xi}v_{N+1}\right)(x),$$

$$\widehat{\Phi}_N(x,\xi) := \sum_{k=1}^{N-1}\mathrm{e}^{-2\pi\mathrm{i} m(c_1+\cdots+c_{k-1})}\left(\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}\overline{\mathcal{L}}_{k+1,\xi}\cdots\overline{\mathcal{L}}_{N,\xi}v_{N+1}\right)(x).$$

We analyze $\overline{\Phi}_N$, $\widetilde{\Phi}_N$ and $\widehat{\Phi}_N$ separately.

Claim 1: *For every $m \in \mathbb{Z}$, $\left|\overline{\Phi}_N(x,\xi) - \mathrm{e}^{-s^2/2}\mathbb{E}_x(v_{N+1}(X_{N+1}))\right| \xrightarrow[N\to\infty]{} 0$ in L^1 uniformly in s on $\{s \in \mathbb{R} : |s| \leq \overline{K}\}$, and $v_{N+1} \in \{v \in L^\infty(\mathfrak{S}_{N+1}) : \|v\| \leq 1\}$.*

Proof of the Claim: $\overline{\Phi}_N(x,\xi) = \mathbb{E}_x\left(\exp\left(\mathrm{i}s\frac{\sum_{k=1}^N \mathbb{F}_k + c}{\sqrt{V_N}}\right)v_{N+1}(X_{N+1})\right)$, and $\mathbb{E}\left(\sum_{k=1}^N \mathbb{F}_n\right) = -c$. Fix $1 \leq r \leq N$. Using the decomposition $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$, we find that

$$\frac{1}{\sqrt{V_N}}\Big(\sum_{k=1}^N \mathbb{F}_k + c\Big) = \frac{1}{\sqrt{V_N}}S_{N-r} + \frac{1}{\sqrt{V_N}}\Big(O(r) - \sum_{k=1}^N h_k\Big), \text{ where } S_{N-r} := \sum_{i=1}^{N-r-1} f_i^{(N)}.$$

By assumption III(b), the L^2 norm of the second summand is $O(1/\sqrt{V_N})$. Therefore the second term converges to 0 in probability as $N \to \infty$, and

$$\overline{\Phi}_N(x,\xi) = \mathbb{E}_x\left(\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}S_{N-r}}v_{N+1}(X_{N+1})\right) + o(1). \tag{6.22}$$

The rate of convergence to 0 of the error term in (6.22) depends on r and m, but is uniform when $|s| \leq \overline{K}$ and $\|v_{N+1}\|_\infty \leq 1$. Next, we study the main term:

$$\begin{aligned}
&\mathbb{E}_x\left(\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}S_{N-r}}v_{N+1}(X_{N+1})\right) = \mathbb{E}_x\left[\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}S_{N-r}}\mathbb{E}_x\left(v_{N+1}(X_{N+1})\middle|X_1,\ldots,X_{N-r}\right)\right]\\
&= \mathbb{E}_x\left[\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}S_{N-r}}\mathbb{E}_x\left(v_{N+1}(X_{N+1})\middle|X_{N-r}\right)\right] \text{(Markov property)}\\
&\overset{(2.11)}{=} \mathbb{E}_x\left(\mathrm{e}^{\frac{\mathrm{i}s}{\sqrt{V_N}}S_{N-r}}[\mathbb{E}_x(v_{N+1}(X_{N+1})) + O(\theta^r)]\right) = \mathbb{E}_x(\mathrm{e}^{\mathrm{i}sS_{N-r}/\sqrt{V_N}})\mathbb{E}_x(v_{N+1}(X_{N+1})) + O(\theta^r),
\end{aligned} \tag{6.23}$$

and the big Oh is uniform in x and in $\|v_{N+1}\|_\infty$ (because X conditioned on $X_1 = x$ has the same ellipticity constant for all x).

Fix r. By Dobrushin's CLT for X conditioned on $X_1 = x$, $\frac{S_{N-r} - \mathbb{E}_x(S_{N-r})}{\sqrt{\mathrm{Var}_x(S_{N-r})}}$ converges in distribution to the standard normal distribution. Equation (6.20) and another mixing argument shows that

$$\mathbb{E}_x(S_{N-r}) = \mathbb{E}(S_{N-r}) = O(1) \text{ and } \mathrm{Var}_x(S_{N-r}) = V_N + O(1).$$

Therefore, conditioned on x, $\frac{S_{N-r}}{\sqrt{V_N}}$ converges in distribution to the standard normal distribution. In particular, it is tight, and $\sup_{|s|\leq\overline{K}} |\mathbb{E}_x(\mathrm{e}^{\mathrm{i}sS_{N-r}/\sqrt{V_N}}) - \mathrm{e}^{-s^2/2}| \xrightarrow[N\to\infty]{} 0$ as $N \to \infty$, for all x.

The claim follows from this, (6.22), and (6.23), and the observation that a uniformly bounded sequence of functions which tends to zero pointwise, tends to zero in L^1.

Claim 2. *There exists $C_2(\overline{K}, m)$ such that for all $|s| \leq \overline{K}$ and $\|v_{N+1}\|_\infty \leq 1$, $\|\widetilde{\Phi}_N(x,\xi)\|_{L^1} \leq C_2(\overline{K}, m)\sqrt{\varepsilon}$.*
Proof of the Claim: We begin with the obvious estimate

$$\|\widetilde{\Phi}_N(x,\xi)\|_1 \leq \|\widetilde{\mathcal{L}}_{1,\xi}\|_{L^\infty\to L^1}\left\|\overline{\mathcal{L}}_{2,\xi}\right\|\cdots\left\|\overline{\mathcal{L}}_{N,\xi}\right\| + \sum_{k=2}^{N-1}\left(\left\|\mathcal{L}_{1,\xi}\right\|\cdots\left\|\mathcal{L}_{k-1,\xi}\widetilde{\mathcal{L}}_{k,\xi}\right\|\left\|\overline{\mathcal{L}}_{k+1,\xi}\right\|\cdots\left\|\overline{\mathcal{L}}_{N,\xi}\right\|\right).$$

Suppose $|s| \leq \overline{K}$. By (6.13) and (6.15), the first summand is bounded above by $C_1(\overline{K}, m)\left[\sigma(h_1)^2 + \frac{\sigma(h_1)\sigma(f_1)}{\sqrt{V_N}}\right]$.

Next, (6.12) and (6.15) give $\left\|\mathcal{L}_{k-1,\xi}\widetilde{\mathcal{L}}_{k,\xi}\right\| \leq C_1(\overline{K}, m)^2\left[\sigma(h_k)^2 + \frac{\sigma(h_k)\sigma(f_k)}{\sqrt{V_N}}\right]$. The other norms are no larger than one.

Therefore $\|\widetilde{\Phi}_N(x,\xi)\|_1 \le C_1(\overline{K},m)^2 \sum_{k=1}^{N-1}\left[\sigma(h_k)^2 + \frac{\sigma(h_k)\sigma(f_k)}{\sqrt{V_N}}\right]$. By the Cauchy-Schwarz inequality, and assumptions II and III(b),

$$\|\widetilde{\Phi}_N(x,\xi)\|_1 \le C_1(\overline{K},m)^2 \sum_{k=1}^{N-1} \sigma(h_k)^2 + C_1(\overline{K},m)^2 \sqrt{\sum_{k=1}^{N-1} \sigma^2(h_k) \cdot \frac{1}{V_N} \sum_{k=1}^{N-1} \sigma^2(f_k)} \le \text{const.}(\epsilon + \sqrt{\epsilon}).$$

Claim 3. *There exists $C_3(\overline{K},m)$ such that for all $|s| \le \overline{K}$ and $\|v_{N+1}\|_\infty \le 1$, $\|\widehat{\Phi}_N(x,\xi)\|_1 \le C_3(\overline{K},m)\sqrt{\varepsilon}$.*

Proof of the Claim. Fix N, $v_{N+1} \in L^\infty(\mathfrak{S}_{N+1})$ such that $\|v_{N+1}\|_\infty \le 1$, and let $\phi_k(\cdot) := (\overline{\mathcal{L}}_{k,\xi} \cdots \overline{\mathcal{L}}_{N,\xi})v_{N+1}$. Observe that $\|\phi_k\|_\infty \le 1$, and

$$\|\widehat{\Phi}_N(x,\xi)\|_1 \le \sum_{k=1}^{N} \|\mathcal{L}_{1,\xi} \cdots \mathcal{L}_{k-1,\xi} \widehat{\mathcal{L}}_{k,\xi} \phi_{k+1}\|_1. \tag{6.24}$$

We will decompose $\phi = \zeta_k' + \zeta_k'' + \eta_k$, and then estimate the contribution to (6.24) from ζ_k', ζ_k'' and η_k.

Step 1 (Decomposition): *We can decompose $\phi_k = \zeta_k' + \zeta_k'' + \eta_k$ so that $\eta_k := \mathbb{E}(\phi_k(X_k))$ (a constant function), and for all $|s| \le \overline{K}$, there exist $\widehat{C}_0, \widehat{K}_0 > 0$ and $0 < \widehat{\theta}_0 < 1$ such that for all $k = 1, \ldots, N-2$*

$$\|\zeta_k'\|_\infty \le \widehat{\theta}_0^2 \|\zeta_{k+2}'\|_\infty + \widehat{K}_0 \|\zeta_{k+2}''\|_1, \tag{6.25}$$

$$\|\zeta_k''\|_1 \le \widehat{C}_0 \left(\frac{\sigma(f_k) + \sigma(f_{k+1}) + \sigma(h_k) + \sigma(h_{k+1})}{\sqrt{V_N}} \right), \tag{6.26}$$

$$\mathbb{E}(\zeta_k'(X_k)) = \mathbb{E}(\zeta_k''(X_k)) = 0. \tag{6.27}$$

Proof of the Step. Let $\zeta_k := \phi_k - \eta_k$, a function with mean zero.

By construction, $\phi_k = \left(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi}\right) \phi_{k+2}$, therefore

$$\phi_k = (\mathcal{L}_k \mathcal{L}_{k+1})\, \eta_{k+2} + (\mathcal{L}_k \mathcal{L}_{k+1})\, \zeta_{k+2} + \left(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}\right) \phi_{k+2}.$$

Observe that $\mathcal{L}_k 1 = 1$, so $(\mathcal{L}_k \mathcal{L}_{k+1})\, \eta_{k+2} = \eta_{k+2}$. This leads to the decomposition

$$\zeta_k = \underbrace{(\mathcal{L}_k \mathcal{L}_{k+1})\, \zeta_{k+2}}_{\zeta_k'} + \underbrace{\left(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}\right) \phi_{k+2} + \eta_{k+2} - \eta_k}_{\zeta_k''}.$$

Then $\zeta_k = \zeta_k' + \zeta_k''$, and the pieces ζ_k', ζ_k'' satisfy the following recursion:

$$\begin{aligned} \zeta_k' &= (\mathcal{L}_k \mathcal{L}_{k+1})\, \zeta_{k+2}' + (\mathcal{L}_k \mathcal{L}_{k+1})\, \zeta_{k+2}'', \\ \zeta_k'' &= \left(\overline{\mathcal{L}}_{k,\xi} \overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k \mathcal{L}_{k+1}\right) \phi_{k+2} + \eta_{k+2} - \eta_k. \end{aligned} \tag{6.28}$$

Notice that ζ_k', ζ_k'' have zero means. Indeed in our setup, $\mu_j(E) = \mathbb{P}(X_j \in E)$, whence by the identities $\zeta_k' = \mathcal{L}_k \mathcal{L}_{k+1} \zeta_{k+2}$ and $(\mathcal{L}_k u)(x) = \mathbb{E}(u(X_{k+1})|X_k = x)$,

$$\int \zeta_k' \mathrm{d}\mu_k = \mathbb{E}(\zeta_k'(X_k)) = \mathbb{E}[\mathbb{E}(\mathbb{E}(\zeta_{k+2}(X_{k+2})|X_{k+1})|X_k)] = \mathbb{E}(\zeta_{k+2}(X_{k+2})) = 0,$$

and $\mathbb{E}(\zeta_k'') = \mathbb{E}(\zeta_k) - \mathbb{E}(\zeta_k') = 0 - 0 = 0$.

To prove the estimates on $\|\zeta_k'\|_\infty$, we first make the following general observations. If $\psi_{k+2} \in L^\infty(\mathfrak{S}_{k+2})$, then $(\mathcal{L}_k \mathcal{L}_{k+1} \psi_{k+2})\,(x) = \int \widetilde{p}(x,z) \psi_{k+2}(z) \mu_{k+2}(\mathrm{d}z)$, where $\widetilde{p}(x,z) = \int_{\mathfrak{S}_{k+1}} p_k(x,y) p_{k+1}(y,z) \mu_{k+1}(\mathrm{d}y)$.

By uniform ellipticity, $\widetilde{p} \geq \epsilon_0$ so we can decompose $\widetilde{p}_k = \epsilon_0 + (1-\epsilon_0)\widetilde{q}_k$ where $\widetilde{q}_k$ is a probability density. Hence, if ψ_{k+2} has zero mean, then

$$(\mathcal{L}_k\mathcal{L}_{k+1}\psi_{k+2})(x)=\epsilon_0\int \psi_{k+2}\mathrm{d}\mu_{k+2} + (1-\epsilon_0)\int \widetilde{q}_k(x,y)\psi_{k+2}(y)\mu_{k+2}(\mathrm{d}y)=(1-\epsilon_0)\int \widetilde{q}_k(x,y)\psi_{k+2}(y)\mu_{k+2}(\mathrm{d}y).$$

Thus $\|\mathcal{L}_k\mathcal{L}_{k+1}\psi_{k+2}\|_\infty \leq (1-\epsilon_0)\|\psi_{k+2}\|_\infty$. By (6.28),

$$\begin{aligned}\|\zeta_k'\|_\infty &= \|(\mathcal{L}_k\mathcal{L}_{k+1})\zeta_{k+2}' + (\mathcal{L}_k\mathcal{L}_{k+1})\zeta_{k+2}''\|_\infty \leq (1-\epsilon_0)\|\zeta_{k+2}'\|_\infty + \|\mathcal{L}_k\mathcal{L}_{k+1}\zeta_{k+2}''\|_\infty\\ &\leq (1-\epsilon_0)\|\zeta_{k+2}'\|_\infty + \|\mathcal{L}_k\|_{L^\infty\to L^\infty}\|\mathcal{L}_{k+1}\|_{L^1\to L^\infty}\|\zeta_{k+2}''\|_1 \leq (1-\epsilon_0)\|\zeta_{k+2}'\|_\infty + C_1(\overline{K},m)\|\zeta_{k+2}''\|_1,\end{aligned}$$

by (6.11) and (6.12). This proves (6.25).

Next we analyze $\|\zeta_k''\|_1$. Since ζ_k'' has zero mean and $\eta_{k+2}-\eta_k$ is constant, we can write $\zeta_k'' = \widehat{\zeta}_k'' - \mathbb{E}(\widehat{\zeta}_k'')$ with $\widehat{\zeta}_k'' := \left(\overline{\mathcal{L}}_{k,\xi}\overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k\mathcal{L}_{k+1}\right)\phi_{k+2}$. Observe that $\left(\overline{\mathcal{L}}_{k,\xi}\overline{\mathcal{L}}_{k+1,\xi}u - \mathcal{L}_{k+1}\mathcal{L}_{k+2}u\right)(x) = \iint k(x,y,z)u(z)\mu_{k+1}(\mathrm{d}y)\mu_{k+2}(\mathrm{d}z)$, where

$$\begin{aligned}|k(x,y,z)| &\leq \text{const.}p_k(x,y)p_{k+1}(y,z)\frac{|s|}{\sqrt{V_N}}|\mathbb{F}_k(x,z)+\mathbb{F}_{k+1}(z,y)+c_k+c_{k+1}|\\ &\leq \frac{\text{const.}}{\sqrt{V_N}}p_k(x,y)p_{k+1}(y,z)\left(|f_k(x,z)|+|f_{k+1}(z,y)|+|h_k(x,z)|+|h_{k+1}(z,y)|\right),\end{aligned}$$

because $\mathbb{F}+\mathsf{c}=\mathsf{f}-\mathsf{h}$.

Next, $\|\zeta_k''\|_1 \leq 2\|\widehat{\zeta}_k''\|_1 \leq 2\iiint |k(x,y,z)|\mu_k(\mathrm{d}x)\mu_{k+1}(\mathrm{d}y)\mu_{k+2}(\mathrm{d}z)\|\phi_{k+2}\|_\infty$. The estimate for $|k(x,y,z)|$ and the Cauchy-Schwarz inequality now lead to (6.26).

Step 2 (Contribution of ζ'): *For some constant which depends only on $K, \overline{K}, m$,*

$$\sum_k \left\|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}')\right\|_{L^1} \leq \text{const.}\sqrt{\varepsilon}. \tag{6.29}$$

Proof of the Step. Using (6.25) and (6.26), it is easy to see by induction that for some constant $\overline{C}$, $\|\zeta_k'\|_\infty$ is bounded above by

$$\begin{aligned}&\overline{C}\Big(\widehat{\theta}_0^{2\lfloor\frac{N-k}{2}\rfloor} + \sum_{r=1}^{\lfloor\frac{N-k}{2}\rfloor-1}\frac{\widehat{\theta}_0^{2r}\left(\sigma(f_{k+2r})+\sigma(f_{k+2r+1})+\sigma(h_{k+2r})+\sigma(h_{k+2r+1})\right)}{\sqrt{V_N}}\Big)\\ &\leq \overline{C}\widehat{\theta}_0^{-1}\Big(\widehat{\theta}_0^{N-k} + \sum_{r=1}^{N-k}\frac{\widehat{\theta}_0^r}{\sqrt{V_N}}(\sigma(f_{k+r})+\sigma(h_{k+r}))\Big).\end{aligned}$$

Since $\mathcal{L}_{j,\xi}$ are contractions and $\|\widehat{\mathcal{L}}_{k,\xi}\|_{L^\infty\to L^1} \leq C_1(\overline{K},m)\sigma(h_k)$, this implies that

$$\sum_k\|\mathcal{L}_{1,\xi}\ldots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(\zeta_{k+1}')\|_{L^1} \leq \overline{C}C_1(\overline{K},m)\left[\sum_r\widehat{\theta}_0^r\sum_k\sigma(h_k)\frac{\sigma(f_{k+r})+\sigma(h_{k+r})}{\sqrt{V_N}} + \sum_k\sigma(h_k)\widehat{\theta}_0^{N-k}\right].$$

As in the proof of Claim 2, it follows from the Cauchy Schwartz inequality, (6.10), and assumption III(b) that the sum over k is $O(\sqrt{\varepsilon})$. Hence (6.29).

Step 3 (Contribution of η_k): *For some constant which depends only on $\overline{K}, m, K$ and ϵ_0,*

$$\sum_k \left\|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}\eta_{k+1}\right\|_1 \leq \text{const.}\sqrt{\varepsilon}. \tag{6.30}$$

Proof of the Step. Recall that η_{k+1} is a constant function, with value $\mathbb{E}(\phi_{k+1})$. Since $\|\phi_{k+1}\|_\infty \le 1$, $|\eta_{k+1}| \le 1$. Therefore (after possibly rescaling η_k) it suffices to prove the step when $\eta_k \equiv 1$. Split $\mathcal{L}_{n,\xi} = e^{2\pi i m c_n}\mathcal{L}_n + \mathcal{L}'_{n,\xi}$, then for all $k \ge 2$,

$$\begin{aligned}\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(1) &= e^{2\pi i m(c_1+\cdots+c_{k-1})}\mathcal{L}_1\cdots\mathcal{L}_{k-1}\widehat{\mathcal{L}}_{k,\xi}(1)\\ &+\sum_{j=1}^{k-1} e^{2\pi i m(c_{j+1}+\cdots+c_{k-1})}\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{j-1,\xi}\mathcal{L}'_{j,\xi}\mathcal{L}_{j+1}\cdots\mathcal{L}_{k-1}\widehat{\mathcal{L}}_{k,\xi}(1).\end{aligned} \tag{6.31}$$

The first expression on the RHS of (6.31), when $k = 1$, equals $\widehat{\mathcal{L}}_{k,\xi}1$, and has norm $\|\widehat{\mathcal{L}}_{k,\xi}1\|_1 \le C_1(\overline{K}, m)\sigma(h_k)$, by (6.14).

For $k \ge 2$, we use (6.17):

$$\|\mathcal{L}_1\cdots\mathcal{L}_{k-1}\widehat{\mathcal{L}}_{k,\xi}1\|_1 = \|\mathcal{L}_1\cdots\mathcal{L}_{k-2}\varphi_{k-1}\|_1 \le \big\|\mathbb{E}(\varphi_{k-1}(X_{k-1})|X_1)\big\|_\infty.$$

By (2.11) and (6.17) the last expression is bounded by $C\theta^k\sigma(h_k)$, for some constant C which depends only on $\overline{K}, m, \epsilon_0$, and some $0 < \theta < 1$ which depends only on ϵ_0. So for all $k \ge 1$, $\|\mathcal{L}_1\cdots\mathcal{L}_{k-1}\widehat{\mathcal{L}}_{k,\xi}1\|_1 \le \text{const.}\theta^k\sigma(h_k)$.

The second term on the RHS of (6.31) is not zero only for $k \ge 2$. If $k = 2$, it has L^1 norm $\|\mathcal{L}'_{1,\xi}\widehat{\mathcal{L}}_{2,\xi}1\|_1 \le C''_1(\overline{K}, m)\sigma(h_2)\left(\dfrac{\sigma(f_1)}{\sqrt{V_N}} + \sigma(h_1)\right)$, by (6.19). If $k \ge 3$, then it has L^1-norm bounded by

$$\begin{aligned}&\sum_{j=1}^{k-1}\|\mathcal{L}_{1,\xi}\|\cdots\|\mathcal{L}_{j-2,\xi}\|\cdot\|\mathcal{L}_{j-1,\xi}\|_{L^1\to L^\infty}\|\mathcal{L}'_{j,\xi}\|_{L^\infty\to L^1}\text{const.}\theta^{k-j}\sigma(h_k)\\ &\le \text{const.}\sum_{j=1}^{k-1}\left(\frac{\sigma(f_j)}{\sqrt{V_k}} + \sigma(h_j)\right)\sigma(h_k)\theta^{k-j},\ \text{see (6.11), (6.12), (6.17), (6.18), (2.11).}\end{aligned}$$

Thus $\sum_k \left\|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(1)\right\|_1$ is bounded by

$$\begin{aligned}&C\sum_{k=1}^N\Big(\sigma(h_k)\theta^k + \sigma(h_k)\sum_{j=1}^k\theta^{k-j}\Big(\frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j)\Big)\Big) \le C\sqrt{\sum_{k=1}^N\sigma^2(h_k)}\sqrt{\sum_{k=1}^N\theta^{2k}} + C\sum_{r=0}^{N-1}\theta^r\sum_{j=1}^N\Big(\frac{\sigma(f_j)}{\sqrt{V_N}} + \sigma(h_j)\Big)\sigma(h_{j+r})\\ &\le C\sqrt{\epsilon} + C\sum_{r=0}^{N-1}\theta^r\left[\left(\sqrt{\sum_{j=1}^N\frac{\sigma^2(f_j)}{V_N}} + \sqrt{\sum_{j=1}^N\sigma^2(h_j)}\right)\sqrt{\sum_{j=1}^N\sigma^2(h_{j+r})}\right] \le C\sqrt{\epsilon},\ \text{and C depends only on } \overline{K}, m, \text{and } \mathsf{X}.\end{aligned}$$

STEP 4 (CONTRIBUTION OF ζ''_k): *For some constant depending only on $\overline{K}, m, K, \epsilon_0$,*

$$\sum_k\left\|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(\zeta''_{k+1})\right\|_{L^1} \le \text{const.}\sqrt{\varepsilon}. \tag{6.32}$$

Proof of the Step. Recall that $\zeta''_{k+1} \equiv \left(\overline{\mathcal{L}}_{k,\xi}\overline{\mathcal{L}}_{k+1,\xi} - \mathcal{L}_k\mathcal{L}_{k+1}\right)\phi_{k+2} + (\eta_{k+2} - \eta_k)$. The term $\eta_{k+2} - \eta_2$ is a constant function with size at most two, and its contribution can be controlled as in the previous step. Let $\widehat{\zeta}''_{k+1} := \zeta''_{k+1} - (\eta_{k+2} - \eta_k)$, then:

$$\begin{aligned}&\|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(\widehat{\zeta}''_{k+1})\|_1 \le \|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-1,\xi}\widehat{\mathcal{L}}_{k,\xi}(\widehat{\zeta}''_{k+1})\|_\infty\\ &\le \|\mathcal{L}_{1,\xi}\cdots\mathcal{L}_{k-2,\xi}\|_{L^\infty\to L^\infty}\cdot\|\mathcal{L}_{k-1,\xi}\|_{L^1\to L^\infty}\|\widehat{\mathcal{L}}_{k,\xi}(\widehat{\zeta}''_{k+1})\|_{L^\infty\to L^1}\\ &\overset{!}{\le} 1\cdot C_1(\overline{K}, m)\cdot C'_1(\overline{K}, m)\frac{\sigma(h_k)}{\sqrt{V_N}}(\sigma(f_{k+1}) + \sigma(f_{k+2}) + \sigma(h_{k+1}) + \sigma(h_{k+2})),\end{aligned}$$

see (6.11), (6.12), and (6.16).

Now we sum over k, and apply the Cauchy-Schwarz inequality and assumptions II and III. Step 4 follows.

Claim 3 follows from Steps 1–4.

Lemma 6.7 follows from Claims 1–3 and (6.21). □

We now return to the proof of Theorem 6.3. Let $\mathsf{f} = \{f_n\}$ be a non center-tight a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain $\mathsf{X} = \{X_n\}$ with state spaces $\mathfrak{S}_n$ and marginals $\mu_n(E) = \mathbb{P}(X_n \in E)$. In particular, $V_N \to \infty$. Without loss of generality, $\delta(\mathsf{f}) = 1$, $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{Z}$, and $\mathbb{E}(f_n) = 0$ for all n.

By the reduction lemma (Lemma 4.16), $\mathsf{f} = \mathbb{F} + \nabla \mathsf{a} + \mathsf{h} + \mathsf{c}$, where $G_{alg}(\mathsf{X}, \mathbb{F}) = \mathbb{Z}$, h has summable variances, $\mathsf{c} = \{c_n\}$ are constants, and $\mathbb{F}$, a, h, c are a.s. uniformly bounded. We make the following additional assumptions:

(IV) $\mathbb{F}_n$ are integer-valued;
(V) $\mathbb{E}(h_n) = 0$, $c_n = -\mathbb{E}(\mathbb{F}_n)$;
(VI) $\mathsf{a} \equiv 0$. In particular, $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$.

Assumptions (IV) and (V) can be arranged using $G_{alg}(\mathsf{X}, \mathbb{F}) = \mathbb{Z}$ and $\mathbb{E}(f_n) = 0$, by trading constants with c_n. Assumption (VI) is a genuine assumption. Let

$$c(N) := -\sum_{k=1}^{N} \mathbb{E}[\mathbb{F}_k(X_k, X_{k+1})]. \tag{6.33}$$

By Theorem 3.12 and the uniform ellipticity assumption, $\mathfrak{H}(X_1, X_2, \dots) := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$ converges almost surely. Here it is essential that X be a Markov chain and not just a Markov array.

Proposition 6.8 *Suppose (IV), (V) and (VI). Let $v_{N+1} \in L^\infty(\mathfrak{S}_{N+1})$ be non-negative functions such that $\|v_{N+1}\|_\infty \neq 0$, and for some $\overline{\delta} > 0$,*

$$\mathbb{E}(v_{N+1}(X_{N+1})) \geq \overline{\delta} \|v_{N+1}\|_\infty. \tag{6.34}$$

Then for all $m \in \mathbb{Z}$, $s \in \mathbb{R}$ and $x \in \mathfrak{S}_1$,

$$\frac{\mathbb{E}_x \left(\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_N}})S_N} v_{N+1}(X_{N+1}) \right)}{\mathbb{E}(v_{N+1}(X_{N+1}))} = \mathrm{e}^{2\pi \mathrm{i} m c(N) - \frac{s^2}{2}} \mathbb{E}_x \left(\mathrm{e}^{2\pi m \mathrm{i} \mathfrak{H}} \right) + o(1), \tag{6.35}$$

as $N \to \infty$, where the $o(\cdot)$ term converges to 0 uniformly when $|m + is|$ are bounded, v_{N+1} are bounded, and (6.34) *holds.*

Proof Since the LHS of (6.35) remains unchanged upon multiplying v_{N+1} by a constant, we may assume that $\|v_{N+1}\|_\infty = 1$.

Fix $\epsilon > 0$ small and r so large that $\sum_{k=r}^{\infty} \mathrm{Var}(h_k) < \epsilon$. Fix N. Applying Lemma 4.19 to $\{\mathbb{F}_n\}_{n=r}^{N}$, we obtain the following decomposition:

$$\mathbb{F}_n(x_n, x_{n+1}) = \mathfrak{a}_{n+1}^{(N)}(x_{n+1}) - \mathfrak{a}_n^{(N)}(x_n) + \mathfrak{c}_n^{(N)} + \widetilde{f}_n^{(N)}(x_n, x_{n+1})$$

where $\mathfrak{c}_n^{(N)}$ are bounded integers, and $\mathfrak{a}_n^{(N)}(\cdot)$, $\widetilde{f}_n^{(N)}(\cdot, \cdot)$ are uniformly bounded measurable integer-valued functions such that $\sum_{n=r}^{N} \|\widetilde{f}_n^{(N)}\|_2^2 = O\left(\sum_{n=r}^{N} u_n^2(\mathbb{F})\right)$.

There is no loss of generality in assuming that $\mathfrak{a}_{N+1}^{(N)} = \mathfrak{a}_r^{(N)} = 0$, otherwise replace $\widetilde{f}_r^{(N)}(x, y)$ by $\widetilde{f}_r^{(N)}(x, y) - \mathfrak{a}_r^{(N)}(x)$, and $\widetilde{f}_N^{(N)}(x, y)$ by $\widetilde{f}_N^{(N)}(x, y) + \mathfrak{a}_{N+1}^{(N)}(y)$. Then

$$\sum_{n=r}^{N} \mathbb{F}_n = \sum_{n=r}^{N} (\mathfrak{c}_n^{(N)} + \widetilde{f}_n^{(N)}). \tag{6.36}$$

This and the identity $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$ gives

$$S_N - S_{r-1} = \sum_{n=r}^{N} f_n = \sum_{n=r}^{N} \left[\mathfrak{c}_n^{(N)} + \widetilde{f}_n^{(N)} + h_n + c_n\right] = \sum_{n=r}^{N} \left[\widetilde{f}_n^{(N)} + h_n - \mathbb{E}(\widetilde{f}_n^{(N)} + h_n)\right], \quad \because \mathbb{E}(S_N - S_r) = 0. \tag{6.37}$$

Let g denote the array with rows $g_n^{(N)} := \widetilde{f}_n^{(N)} + h_n - \mathbb{E}(\widetilde{f}_n^{(N)} + h_n)$ $(n = r, \dots, N)$, $N > r$. We claim that g satisfies assumptions (I)–(III) of Lemma 6.7. (I) is clear, and (III) holds by choice of r and because $\widetilde{f}_n^{(N)}$ is integer-valued. Next we check (II):

$$\sum_{n=1}^{N} \sigma^2(g_n^{(N)}) = \sum_{n=1}^{N} \sigma^2(\widetilde{f}_n^{(N)} + h_n) = \sum_{n=1}^{N} \left[\sigma^2(\widetilde{f}_n^{(N)}) + \sigma^2(h_n + 2\mathrm{Cov}(\widetilde{f}_n^{(N)}, h_n)\right] \leq \sum_{n=1}^{N} \left[\sigma^2(\widetilde{f}_n^{(N)}) + \sigma^2(h_n) + 2\sigma(\widetilde{f}_n^{(N)})\sigma(h_n)\right]$$

$$\leq 2\sum_{n=1}^{N} \left[\sigma^2(\widetilde{f}_n^{(N)}) + \sigma^2(h_n)\right] = O\Big(\sum_{n=r}^{N} u_n^2(\mathbb{F})\Big) + O(1), \text{ by choice of } \widetilde{\mathsf{f}} \text{ and } \mathsf{h}.$$

Since $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$, $u_n^2(\mathbb{F}) = u_n^2(\mathsf{f} - \mathsf{h} - \mathsf{c}) \leq 2[u_n^2(\mathsf{f}) + u_n^2(\mathsf{h})]$, see Lemma 2.16(4). Thus by Theorem 3.7 and the assumption that h has summable variances,

$$\sum_{n=r}^{N} u_n^2(\mathbb{F}) \leq 2\sum_{n=r}^{N} \left[u_n^2(f) + u_n^2(h)\right] = O\big(\mathrm{Var}(S_N - S_r)\big) + O(1) = O\big(\mathrm{Var}(S_N - S_r)\big).$$

We now apply Lemma 6.7 to g, and deduce that for every $\overline{K} > 0$ and $m \in \mathbb{Z}$ there are $\overline{C}, \overline{N} > 0$ such that for all $N > \overline{N} + r$, $|s| \leq \overline{K}$, and v_{N+1} in the unit ball of L^∞

$$\begin{aligned}&\mathbb{E}\left(\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_N}})(S_N - S_{r-1})} v_{N+1}(X_{N+1}) \Big| X_r\right)\\ &= \mathrm{e}^{2\pi \mathrm{i} m c^{(N)}} \cdot \mathrm{e}^{-s^2/2} \mathbb{E}(v_{N+1}(X_{N+1})) + \eta_{N-r}(X_r),\end{aligned}$$

where $c^{(N)} := -\sum_{n=r}^{N} \mathbb{E}(\widetilde{f}_n^{(N)})$ and $\|\eta_{N-r}\|_1 \leq \overline{C}\sqrt{\epsilon}$. Since $\|v_{N+1}\|_\infty = 1$, we also have the trivial bound $\|\eta_{N-r}\|_\infty \leq 2$.

We are ready to prove the proposition:

$$\begin{aligned}&\frac{\mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_N}})S_N} v_{N+1}(X_{N+1})\right)}{\mathbb{E}(v_{N+1}(X_{N+1}))} = \mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_N}})S_{r-1}} \frac{\mathbb{E}\left(\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_n}})(S_N - S_{r-1})} v_{N+1}(X_{N+1}) | X_r\right)}{\mathbb{E}(v_{N+1}(X_{N+1}))}\right)\\ &= \mathbb{E}_x\left[\mathrm{e}^{\mathrm{i}(2\pi m + \frac{s}{\sqrt{V_N}})S_{r-1}} \left(\mathrm{e}^{2\pi \mathrm{i} m c^{(N)} - s^2/2} + \frac{\eta_{N-r}(X_r)}{\mathbb{E}(v_{N+1}(X_{N+1}))}\right)\right]\\ &= \underbrace{\mathrm{e}^{2\pi \mathrm{i} m c^{(N)} - s^2/2} \mathbb{E}_x(\mathrm{e}^{2\pi \mathrm{i} m S_{r-1} + o(1)})}_{A} + O(\overline{\delta}^{-1}) \underbrace{\mathbb{E}_x\left(|\eta_{N-r}(X_r)|\right)}_{B}, \text{ as } N \to \infty.\end{aligned}$$

Summand A: By assumption, $\mathsf{f} = \mathbb{F} + \mathsf{h} + \mathsf{c}$ with $\mathbb{F}$ integer valued. Necessarily,

$$\exp(2\pi \mathrm{i} m S_{r-1}) = \exp(2\pi \mathrm{i} m \mathfrak{H}_r + 2\pi \mathrm{i} m \widehat{c}^{(r-1)}), \tag{6.38}$$

where $\mathfrak{H}_r := \sum_{k=1}^{r-1} h_k(X_k, X_{k+1})$ and $\widehat{c}^{(r-1)} := \sum_{k=1}^{r-1} c_k = -\mathbb{E}(\sum_{k=1}^{r-1} \mathbb{F}_k(X_k, X_{k+1}))$. Substituting (6.38) in A, we obtain

$$A = \mathrm{e}^{2\pi \mathrm{i} m[c^{(N)}+\widetilde{c}^{(r-1)}]-s^2/2}\mathbb{E}_x(\mathrm{e}^{2\pi \mathrm{i} m \mathfrak{H}_r}).$$

We claim that $c^{(N)} + \widetilde{c}^{(r-1)} = c(N) \bmod \mathbb{Z}$:

$$\begin{aligned} c(N) &\equiv -\sum_{k=1}^{N} \mathbb{E}(\mathbb{F}_k) \stackrel{!}{=} -\sum_{k=1}^{r-1} \mathbb{E}(\mathbb{F}_k) - \sum_{k=r}^{N} \mathbb{E}(\widetilde{f}_k^{(N)}) - \sum_{k=r}^{N} \mathfrak{c}_k^{(N)}, \quad \text{by (6.36)} \\ &\equiv \widetilde{c}^{(r-1)} + c^{(N)} - \sum_{k=r}^{N} \mathfrak{c}_k^{(N)} \stackrel{!}{=} \widetilde{c}^{(r-1)} + c^{(N)} \mod \mathbb{Z}, \quad \text{because } \mathfrak{c}_k^{(N)} \in \mathbb{Z}. \end{aligned}$$

The following bound holds uniformly when ξ varies in a compact domain, by the choice of r, Lemma 3.4, and the Cauchy-Schwarz inequality:

$$\left|\mathbb{E}_x(\mathrm{e}^{\mathrm{i}\xi\mathfrak{H}}) - \mathbb{E}_x(\mathrm{e}^{\mathrm{i}\xi\mathfrak{H}_r})\right| \le |\xi|\mathbb{E}_x(|\mathfrak{H} - \mathfrak{H}_r|) \le |\xi| \mathrm{Var}\Big(\sum_{k=r}^{\infty} h_k(X_k, X_{k+1})\Big)^{\frac{1}{2}} = O(\sqrt{\epsilon}).$$

It follows that $A = [1 + o(1)]\mathrm{e}^{2\pi \mathrm{i} m c(N)-s^2/2}\mathbb{E}_x(\mathrm{e}^{2\pi \mathrm{i} m\mathfrak{H}}) + O\left(\sqrt{\varepsilon}\right)$.

Summand B: By the exponential mixing of X, for all N large enough,

$$B = \mathbb{E}_x(|\eta_{N-r}(X_r)|) = \mathbb{E}(|\eta_{N-r}(X_r)|) + o(1) = O(\sqrt{\epsilon}).$$

Thus the left-hand-side of (6.35) equals $\mathrm{e}^{2\pi \mathrm{i} m c(N)-s^2/2}\mathbb{E}_x(\mathrm{e}^{2\pi \mathrm{i} m\mathfrak{H}+o(1)}) + O(\sqrt{\epsilon})$. The lemma follows, because ϵ was arbitrary. □

6.2.2 Proof of the LLT in the Reducible Case

We prove Theorem 6.3.

It is sufficient to prove parts 2 and 3, on $\mathbb{E}_x(\phi(S_N - z_N - b_N)|X_{N+1} \in \mathfrak{A}_{N+1})$. Part 1, which deals with $\mathbb{E}(\phi(S_N - z_N - b_N))$, can be deduced as follows: The conditioning on $X_1 = x$ can be removed using Lemma 2.27; and the conditioning on $X_{N+1} \in \mathfrak{A}_{N+1}$ can be removed by taking $\mathfrak{A}_{N+1} := \mathfrak{S}_{N+1}$.

Suppose f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X, and assume $G_{ess}(\mathsf{X}, \mathsf{f}) = \delta(\mathsf{f})\mathbb{Z}$ with $\delta(\mathsf{f}) \neq 0$.

We begin with some reductions. By Theorem 4.5, f has an optimal reduction, and we can write $\mathsf{f} = \mathbb{F} + \mathsf{F}$ where $\mathbb{F}$ has algebraic range $\delta(\mathsf{f})\mathbb{Z}$ and F is a.s. uniformly bounded and center-tight. Trading constants in $\delta(\mathsf{f})\mathbb{Z}$ between $\mathbb{F}_n$ and F_n we can arrange for $\mathbb{F}_n$ to be a.s. $\delta(\mathsf{f})\mathbb{Z}$-valued.

Let $t = [t]_{\mathbb{Z}} + \{t\}_{\mathbb{Z}}$ denote the unique decomposition of $t \in \mathbb{R}$ into the ordered sum of an integer and a number in $[0, 1)$. Replacing F by $\{\mathsf{F}\}_{\delta(\mathsf{f})\mathbb{Z}}$ and $\mathbb{F}$ by $\mathbb{F} + [\mathsf{F}]_{\delta(\mathsf{f})\mathbb{Z}}$, we can also arrange ess sup $|\mathsf{F}| \le \delta(\mathsf{f})$.

By the gradient lemma (Lemma 3.14), we can decompose

$$\mathsf{F} = \nabla \mathsf{a} + \widetilde{\mathsf{f}} + \widetilde{\mathsf{c}}$$

where ess sup $|\mathsf{a}| \le 2$ess sup $|\mathsf{F}|$, $\widetilde{\mathsf{f}}$ has summable variances, and $\widetilde{c}_n$ are constants. Let $f_n^* := \dfrac{1}{\delta(\mathsf{f})}[f_n - \nabla a_n - \mathbb{E}(f_n - \nabla a_n)]$, then $G_{ess}(\mathsf{X}, \mathsf{f}^*) = \mathbb{Z}$, and

$$\mathsf{f}^* = \frac{1}{\delta(\mathsf{f})}\mathbb{F} + \mathsf{h} + \mathsf{c}, \tag{6.39}$$

where $h_n := \frac{1}{\delta(\mathsf{f})}[\widetilde{f}_n - \mathbb{E}(\widetilde{f}_n)]$ is a centered additive functional with summable variances, and $c_n := \frac{1}{\delta(\mathsf{f})}[\widetilde{c}_n + \mathbb{E}(\widetilde{f}_n) - \mathbb{E}(f_n - \nabla a_n)]$.

A Special Case: We begin with the special case when

$$\delta(\mathsf{f}) = 1, \quad \mathbb{E}(f_n) = 0 \quad \text{for all } n, \quad \text{and} \quad \mathsf{a} \equiv 0. \tag{6.40}$$

In this case $\mathsf{f} = \mathsf{f}^*$, and (6.39) places us in the setup of Proposition 6.8. Given this proposition, the proof of part (2) of the theorem is very similar to the proof of the mixing LLT in the irreducible non-lattice case, but we give it for completeness.

As in (6.33), let $c(N) := -\sum_{k=1}^{N} \mathbb{E}[\mathbb{F}_k(X_k, X_{k+1})]$, $\mathfrak{H} := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$, and

$$b_N := \{c(N)\}_{\mathbb{Z}}.$$

Fix $\phi \in L^1(\mathbb{R})$ such that $\mathrm{supp}(\widehat{\phi}) \subset [-L, L]$, and let ν_{N+1} denote the indicator function of $\mathfrak{A}_{N+1}$. By the Fourier inversion formula

$$\begin{aligned} &\mathbb{E}_x(\phi(S_N - b_N - z_N)|X_{N+1} \in \mathfrak{A}_{N+1}) \\ &= \frac{1}{2\pi} \int_{-L}^{L} \widehat{\phi}(\xi) \frac{\mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}\xi(S_N - b_N - z_N)} \nu_{N+1}(X_{N+1})\right)}{\mathbb{E}(\nu_{N+1}(X_{N+1}))} \mathrm{d}\xi. \end{aligned} \tag{6.41}$$

Our task is to find the asymptotic behavior of (6.41) in case $z_N \in \mathbb{Z}$, $\frac{z_N}{\sqrt{V_N}} \to z$.

Let $K :=$ ess sup $|\mathsf{f}|$ and recall the constant $\widetilde{\delta} = \widetilde{\delta}(K)$ from Lemma 5.8. Split $[-L, L]$ into a finite collection of subintervals I_j of length less than $\min\{\widetilde{\delta}, \pi\}$, in such a way that every I_j is either bounded away from $2\pi\mathbb{Z}$, or intersects it an unique point $2\pi m$ exactly at its center. Let $J_{j,N}$ denote the contribution of $I_{j,N}$ to (6.41).

If $I_j \cap 2\pi\mathbb{Z} \neq \varnothing$, then the center of I_j equals $2\pi m$ for some $m \in \mathbb{Z}$. Fix some large R. Let $J'_{j,N}$ be the contribution to (6.41) from $\{\xi \in I_j : |\xi - 2\pi m| \le R V_N^{-1/2}\}$, and let $J''_{j,N}$ be the contribution to (6.41) from $\{\xi \in I_j : |\xi - 2\pi m| > R V_N^{-1/2}\}$.

Working as in Claim 2 in §5.2.4, one can show that for every R,

$$|J''_{j,N}| \le C \int_{|u| > R V_N^{-1/2}} \mathrm{e}^{-c V_N u^2} \mathrm{d}u \le C \frac{\mathrm{e}^{-cR^2}}{R\sqrt{V_N}} = \frac{o_{R\to\infty}(1)}{\sqrt{V_N}}.$$

Thus the main contribution comes from $J'_{j,N}$. We make the change of variables $\xi = 2\pi m + \frac{s}{\sqrt{V_N}}$. Since $z_N \in \mathbb{Z}$ and $b_N = \{c(N)\}_{\mathbb{Z}}$, we have

$$\xi(S_N - b_N - z_N) = \xi S_N - 2\pi m c(N) - \frac{s}{\sqrt{V_N}}(z_N + \{c(N)\}_{\mathbb{Z}}) \bmod 2\pi.$$

So $J'_{j,N}$ is equal to

$$\frac{1}{2\pi\sqrt{V_N}} \int_{-R}^{R} \widehat{\phi}\left(2\pi m + \frac{s}{\sqrt{V_N}}\right) \frac{\mathrm{e}^{-2\pi \mathrm{i} m c(N)} \mathbb{E}_x\left(\mathrm{e}^{\mathrm{i}\xi S_N} \nu_{N+1}(X_{N+1})\right)}{\mathbb{E}(\nu_{N+1}(X_{N+1}))} \mathrm{e}^{-\mathrm{i}s \frac{z_N + O(1)}{\sqrt{V_N}}} \mathrm{d}s.$$

Fixing R and letting $N \to \infty$, we see by Proposition 6.8 that

$$\sqrt{V_N} J'_{j,N} = \frac{1}{2\pi} \widehat{\phi}(2\pi m) \mathbb{E}_x\left(\mathrm{e}^{2\pi \mathrm{i} m \mathfrak{H}}\right) \int_{|s|<R} \mathrm{e}^{-\mathrm{i}sz - s^2/2} \mathrm{d}s + o_{N\to\infty}(1)$$

$$= \frac{1}{\sqrt{2\pi}} \widehat{\phi}(2\pi m) \mathbb{E}_x\left(\mathrm{e}^{2\pi \mathrm{i} m \mathfrak{H}}\right) \mathrm{e}^{-z^2/2} + o_{R\to\infty}(1) + o_{N\to\infty}(1).$$

Combining the estimates for $J'_{j,N}$, $J''_{j,N}$, we obtain that if I_j intersects $2\pi\mathbb{Z}$, then

$$\lim_{N\to\infty} \sqrt{V_N}\, J_{j,N} = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \mathbb{E}_x\left(e^{2\pi i m \mathfrak{H}}\right)\widehat{\phi}(2\pi m).$$

If $I_j \cap 2\pi\mathbb{Z} = \varnothing$, then $\sum d_n^2(\xi) = \infty$ uniformly on I_j (Theorem 4.9). Thus by (5.19), $\Phi_N(x,\xi) \to 0$ uniformly on I_j. In this case we can argue as in the proof of (5.29) and show that the contribution of I_j to the integral (6.41) is $o(V_N^{-1/2})$. Hence,

$$\begin{aligned}
&\lim_{n\to\infty} \sqrt{V_N}\mathbb{E}_x(\phi(S_N - b_N - z_N)|X_{N+1} \in \mathfrak{A}_{N+1})\\
&= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}\cap[-L,L]} \mathbb{E}_x\left(e^{2\pi i m\mathfrak{H}}\right)\widehat{\phi}(2\pi m) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}_x\left(e^{2\pi i m\mathfrak{H}}\right)\widehat{\phi}(2\pi m)\\
&= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}_x\left(e^{2\pi i m\mathfrak{F}}\right)\widehat{\phi}(2\pi m), \text{ where } \mathfrak{F}\in[0,1),\ \mathfrak{F} := \mathfrak{H} \bmod \mathbb{Z}\\
&= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \widehat{(C_x\phi)}(2\pi m), \text{ where } (C_x\phi)(t) := \mathbb{E}_x[\phi(t+\mathfrak{F})],\\
&\overset{!}{=} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (C_x\phi)(m) \equiv \frac{e^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}_x[\phi(m+\mathfrak{F})],
\end{aligned}$$

by Poisson's summation formula.

This proves part (2) of the theorem, in the special case (6.40), and in particular for the additive functional

$$\mathsf{f}^* = \frac{1}{\delta(\mathsf{f})}[\mathsf{f} - \nabla\mathsf{a} - \mathbb{E}(\mathsf{f} - \nabla\mathsf{a})].$$

Proof of the Theorem in the General Case: $\mathsf{f} - \mathbb{E}(\mathsf{f}) = \delta(\mathsf{f})\mathsf{f}^* + \nabla\mathsf{a} - \mathbb{E}(\nabla\mathsf{a})$, so

$$S_N(\mathsf{f}) - \mathbb{E}[S_N(\mathsf{f})] \equiv \delta(\mathsf{f})S_N(\mathsf{f}^*) + a_{N+1}(X_{N+1}) - a_1(X_1) + \mathbb{E}[a_1(X_1) - a_{N+1}(X_{N+1})].$$

Since part (2) of Theorem 6.3 holds for f^* with $\mathfrak{F}=\{\sum h_n\}\in[0,1)$ and $b_N=\{c(N)\}_{\mathbb{Z}}$, it must also hold for f with $\delta(\mathsf{f})\mathfrak{F}$ and

$$b_N(X_1, X_{N+1}) := \delta(\mathsf{f})\{c(N)\}_{\mathbb{Z}} + a_{N+1}(X_{N+1}) - a_1(X_1) + \mathbb{E}[a_1(X_1) - a_{N+1}(X_{N+1})].$$

Clearly $|b_N| \le \delta(\mathsf{f}) + 4\text{ess sup}|\mathsf{a}|$. Recalling that $\text{ess sup}|\mathsf{a}| \le 2\text{ess sup}|\mathsf{F}| \le 2\delta(\mathsf{f})$, we find that $\text{ess sup}|b_N| \le 9\delta(\mathsf{f})$, proving part (3) as well.

As we explained in the beginning of the proof, part (1) of Theorem 6.3 is a special case of part (2). □

6.2.3 Necessity of the Irreducibility Assumption

Suppose f is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X. Recall that $\mathsf{f}_r = \{f_n\}_{n\ge r}$ and $\mathsf{X}_r = \{X_n\}_{n\ge r}$. In this section we prove Theorem 6.5, which asserts the equivalence of the following three conditions:

(a) f is irreducible with algebraic range $\mathbb{R}$.
(b) $(\mathsf{X}_r, \mathsf{f}_r)$ satisfies the mixing non-lattice LLT, for all r.
(c) $(\mathsf{X}_r, \mathsf{f}_r)$ satisfies the mixing mod t LLT for all r and $t > 0$.

(a)⇒(b): By Theorem 4.4, if (a) holds then $H(\mathsf{X},\mathsf{f}) = \{0\}$. Clearly, $H(\mathsf{X}_r,\mathsf{f}_r) = H(\mathsf{X},\mathsf{f})$ for all r, therefore $H(\mathsf{X}_r,\mathsf{f}_r) = \{0\}$ for all r. By Theorem 4.4, $(\mathsf{X}_r,\mathsf{f}_r)$ are all irreducible with essential range $\mathbb{R}$. Part (b) now follows from Theorem 5.4.

(b)⇒(a): Without loss of generality, $\mathbb{E}(f_n(X_n, X_{n+1})) = 0$ for all n.

If $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$ then f is irreducible with algebraic range $\mathbb{R}$, and we are done. Assume by way of contradiction that $G_{ess}(\mathsf{X},\mathsf{f}) \neq \mathbb{R}$, then $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$ for some t. If t were equal to zero, then (X,f) would have been center-tight, and V_N would have been bounded (Theorem 3.8). By the definition of the mixing non-lattice LLT, $V_N \to \infty$, so $t \neq 0$. There is no loss of generality in assuming that $t = 1$. So

$$G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{Z} \text{ and } H(\mathsf{X},\mathsf{f}) = 2\pi\mathbb{Z}.$$

Let $S_N^{(r)} := f_r(X_r, X_{r+1}) + \cdots + f_N(X_N, X_{N+1})$ and $V_N^{(r)} := \mathrm{Var}(S_N^{(r)})$. Clearly, $\mathbb{E}(S_N^{(r)}) = \mathbb{E}(S_N) = 0$. Next, by the exponential mixing of X (Proposition 2.13),

$$|V_N - V_N^{(r)}| = |V_{r-1} + 2\mathrm{Cov}(S_N^{(r)}, S_{r-1})| \leq V_r + 2\sum_{j=1}^{r-1}\sum_{k=r}^{\infty} |\mathrm{Cov}(f_j, f_k)| = O(1).$$

Therefore, for fixed r, $V_N/V_N^{(r)} \xrightarrow[N\to\infty]{} 1$.

By the reduction lemma,

$$\mathsf{f} = \mathbb{F} + \nabla \mathsf{a} + \mathsf{h} + \mathsf{c},$$

where $\mathbb{F}$ is irreducible with algebraic range $\mathbb{Z}$, $a_n(x)$ are uniformly bounded, h has summable variances, $\mathbb{E}(a_n) = 0$, $\mathbb{E}(h_n) = 0$, and c are constants.

Let $\mathfrak{F} := \sum_{n=1}^{\infty} h_n(X_n, X_{n+1})$, $\mathfrak{F}_r := \sum_{n=r}^{\infty} h_n(X_n, X_{n+1})$ (the sums converge a.s. and in L^2 by Theorem 3.12). Next, set $\beta_N^{(r)} := \left\{-\sum_{k=r}^{N} \mathbb{E}(\mathbb{F}_k(X_k, X_{k+1}))\right\}$ (where $\{\cdot\}$ denotes the fractional part), and define

$$b_N^{(r)}(X_r, X_{N+1}) := a_{N+1}(X_{N+1}) - a_r(X_r) + \beta_N^{(r)}.$$

If $\dfrac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\dfrac{z_N - \mathbb{E}(S_N^{(r)})}{\sqrt{V_N^{(r)}}} \to 0$.

By Theorem 6.3 and its proof, if $\liminf_{N\to\infty} \mathbb{P}(X_n \in \mathfrak{A}_n^{(r)}) > 0$, then the following holds for all $\phi \in C_c(\mathbb{R})$ and $x_r \in \mathfrak{S}_r$:

$$\lim_{N\to\infty} \sqrt{2\pi V_N^{(r)}}\,\mathbb{E}_{x_r}[\phi(S_N^{(r)} - b_N^{(r)} - z_N) | X_{N+1} \in \mathfrak{A}_{N+1}^{(r)}] = \sum_{m\in\mathbb{Z}} \mathbb{E}_{x_r}[\phi(m + \mathfrak{F}_r)] \tag{6.42}$$

Here and throughout, we abuse notation and write for $x_r \in \mathfrak{S}_r$, $\mathbb{E}_{x_r} := \mathbb{E}(\cdot|X_r = x_r)$ and $\mathbb{P}_{x_r} := \mathbb{P}(\cdot|X_r = x_r)$.

The plan is to choose $r, x_r, \mathfrak{A}_{N+1}^{(r)}, z_N$ and ϕ in such a way that (6.42) is inconsistent with the mixing LLT.

Choice of r: Let $K \in \mathbb{N}$ be a bound for $\operatorname{ess\,sup} |\mathsf{a}|$, and fix $\gamma := 10^{-6}(4K+2)^{-4}$. Since $\mathfrak{F}_r \xrightarrow[r\to\infty]{} 0$ almost surely, we can choose r such that

$$\mathbb{P}(|\mathfrak{F}_r| < 0.2) > 1 - \gamma^2.$$

Let μ_r denote the marginal distribution of X_r. We claim that

$$\mu_r\{x_r \in \mathfrak{S}_r : \mathbb{P}_{x_r}(|\mathfrak{F}_r| < 0.2) > 1 - \gamma\} > 1 - \gamma. \tag{6.43}$$

To see this, let $p_r(x_r) := \mathbb{P}_{x_r}(|\mathfrak{F}_r| < 0.2)$ and $\alpha := \mu_r[p_r > 1 - \gamma]$, then

$$1-\gamma^2 < \mathbb{P}(|\mathfrak{F}_r| < 0.2) = \int p_r \mathrm{d}\mu_r = \int_{[p_r > 1-\gamma]} p_r \mathrm{d}\mu_r + \int_{[p_r \le 1-\gamma]} p_r \mathrm{d}\mu_r$$
$$\le \alpha + (1-\alpha)(1-\gamma) = 1-\gamma(1-\alpha). \text{ So } \alpha > 1-\gamma.$$

Choice of x_r. Divide $[-2K-1, 2K+1]$ into a family $\mathcal{I}$ of $10^3(4K+2)$ pairwise disjoint intervals of equal length 10^{-3}. There must be some $J'_r \in \mathcal{I}$ such that $\mathbb{P}(a_r(X_r) \in J'_r) \ge |\mathcal{I}|^{-1} \gg \gamma$. Then

$$\mu_r\{x_r \in \mathfrak{S}_r : a_r(x_r) \in J'_r\} + \mu_r\{x_r \in \mathfrak{S}_r : \mathbb{P}_{x_r}(|\mathfrak{F}_r| < 0.2) > 1-\gamma\} > 1.$$

Choose x_r in the intersection of the last two events. Then

$$\mathbb{P}_{x_r}(|\mathfrak{F}_r| < 0.2) > 1-\gamma \text{ and } a_r(x_r) \in J'_r.$$

Choice of $\mathfrak{A}^{(r)}_{N+1}$: Choose $J'_N \in \mathcal{I}$ such that $\mathbb{P}[a_{N+1}(X_{N+1}) \in J'_N] \ge |\mathcal{I}|^{-1}$.

Let $J_N := J'_N - J'_r + \beta^{(r)}_N := \{a - b + \beta^{(r)}_N : a \in J'_N, b \in J'_r\}$. This is an interval of length $|J_N| = |J'_N| + |J'_r| < 0.01$, and $J_N \subset [-4K-3, 4K+3]$. Define

$$\mathfrak{A}^{(r)}_{N+1} := \{y \in \mathfrak{S}_{N+1} : b^{(r)}_N(x_r, y) \in J_N\}.$$

We claim that $\mathfrak{A}_N$ is regular, i.e. $\liminf_{N\to\infty} \mathbb{P}[X_{N+1} \in \mathfrak{A}^{(r)}_{N+1}] > 0$:

$$\begin{aligned}
&\mathbb{P}[X_{N+1} \in \mathfrak{A}^{(r)}_{N+1}] \ge \mathbb{P}_{x_r}[X_{N+1} \in \mathfrak{A}^{(r)}_{N+1}] - C_{mix}\theta^{N-r}, \text{ with } 0<\theta<1, \text{ see (2.11)}\\
&\equiv \mathbb{P}_{x_r}[b^{(r)}_N(x_r, X_{N+1}) \in J_N] - C_{mix}\theta^{N-r}\\
&\equiv \mathbb{P}_{x_r}\big[a_{N+1}(X_{N+1}) - a_r(x_r) + \beta^{(r)}_N \in J'_N - J'_r + \beta^{(r)}_N\big] - C_{mix}\theta^{N-r}\\
&\ge \mathbb{P}_{x_r}\big[a_{N+1}(X_{N+1}) \in J'_N\big] - C_{mix}\theta^{N-r}, \text{ because } a_r(x_r) \in J'_r\\
&\ge \mathbb{P}\big[a_{N+1}(X_{N+1}) \in J'_N\big] - 2C_{mix}\theta^{N-r} \ge |\mathcal{I}|^{-1} - o(1), \text{by the choice of } J'_N, \text{(2.11)}.
\end{aligned}$$

Choice of z_N: Let $\zeta_N :=$ −center of J_N, then $|\zeta_N| \le 2K+1$. Let $z_N := [\zeta_N]_{\mathbb{Z}}$ (the integer part of ζ_N). Then $z_N \in \mathbb{Z}$ and $\frac{z_N - \mathbb{E}(S^{(r)}_N)}{\sqrt{V^{(r)}_N}} \to 0$.

Choice of N_k and ϕ: Choose a sequence $N_k \to \infty$ such that $\zeta_{N_k} \to a$. Let $I := -a + [0.4, 0.6]$, and choose $\phi \in C_c(\mathbb{R})$ such that $1_{[0.3,0.7]} \le \phi \le 1_{[0.2,0.8]}$.

The Contradiction. With these choices, if (b) holds but $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{Z}$, then

$$\begin{aligned}
|I| &= \lim_{N\to\infty} \sqrt{2\pi V^{(r)}_N}\,\mathbb{P}_{x_r}(S^{(r)}_N - z_N \in I | X_{N+1} \in \mathfrak{A}^{(r)}_{N+1}), \text{ by (b)}\\
&= \lim_{k\to\infty} \sqrt{2\pi V^{(r)}_{N_k}}\,\mathbb{P}_{x_r}(S^{(r)}_{N_k} - z_{N_k} \in I | b^{(r)}_{N_k}(X_r, X_{N_k+1}) \in J_{N_k}), \text{by choice of } x_r, \mathfrak{A}^{(r)}_{N+1}\\
&\overset{!}{\le} \liminf_{k\to\infty} \sqrt{2\pi V^{(r)}_{N_k}}\,\mathbb{P}_{x_r}(S^{(r)}_{N_k} - b^{(r)}_{N_k} - z_{N_k} \in [0.3, 0.7] \,|\, b^{(r)}_{N_k} \in J_{N_k}),
\end{aligned}$$

because $I - b^{(r)}_{N_k} \subset I - J_{N_k} \subset I + \left(\zeta_{N_k} - \frac{|J_{N_k}|}{2}, \zeta_{N_k} + \frac{|J_{N_k}|}{2}\right)$, and $\zeta_{N_k} \to a$, so for $k \gg 1$, $I - b^{(r)}_{N_k} \subset I + (a - 0.1, a + 0.1) \subset [0.3, 0.7]$. Hence

$$\begin{aligned}
|I| &\le \lim_{k\to\infty} \sqrt{2\pi V^{(r)}_{N_k}}\,\mathbb{E}_{x_r}(\phi(S^{(r)}_{N_k} - b^{(r)}_{N_k} - z_{N_k}) \,|\mathfrak{A}^{(r)}_{N+1}), \text{ because } \phi \ge 1_{[0.3,0.7]}\\
&= \sum_{m\in\mathbb{Z}} \mathbb{E}_{x_r}[\phi(m + \mathfrak{F}_r)] \le \sum_{m\in\mathbb{Z}} \mathbb{P}_{x_r}(m + \mathfrak{F}_r \in [0.2, 0.8]) \le \mathbb{P}_{x_r}(|\mathfrak{F}_r| \ge 0.2) < \gamma
\end{aligned}$$

by (6.42) and the choice of r, x_r, and $\mathfrak{A}^{(r)}_{N+1}$. But $|I| = 0.2$ and $\gamma < 10^{-6}$.

(a)⇒(c): Fix r, $t > 0$, $x \in \mathfrak{S}_r$, and some sequence of measurable events $\mathfrak{A}_n \subset \mathfrak{S}_n$ such that $\mathbb{P}(X_n \in \mathfrak{A}_n)$ is bounded below. We need to show that if $0 < b - a < t$, then

$$\mathbb{P}_x(S_N^{(r)} \in (a,b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N\to\infty]{} \frac{|a-b|}{t}.$$

By standard approximation arguments, it suffices to show that for every continuous periodic function $\phi(x)$ with period t,

$$\mathbb{E}_x(\phi(S_N^{(r)}) | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N\to\infty]{} \frac{1}{t}\int_0^t \phi(x)\mathrm{d}x. \tag{6.44}$$

By the Stone-Weierstrass theorem, it is sufficient to do this for trigonometric polynomials $\phi(u) = \sum_{|n|<L} c_n \mathrm{e}^{2\pi \mathrm{i} n u/t}$. For such ϕ, we have the following:

$$\begin{aligned}\mathbb{E}_x(\phi(S_N^{(r)}) | X_{N+1} \in \mathfrak{A}_{N+1}) &= \sum_{|n|<L} c_n \mathbb{E}_x(\mathrm{e}^{2\pi \mathrm{i} n S_N^{(r)}/t} | X_{N+1} \in \mathfrak{A}_{N+1}) \\ &= c_0 + \sum_{0<|n|<L} \Phi_N\left(x, \tfrac{2\pi n}{t} | \mathfrak{A}_{N+1}\right), \text{ with } \Phi_N \text{ as in §5.2.2.}\end{aligned}$$

As we saw in the proof of (a)⇒(b), (a) implies that $(\mathsf{X}_r, \mathsf{f}_r)$ are non-lattice and irreducible for all r. Therefore $G_{ess}(\mathsf{X}_r, \mathsf{f}_r) = \mathbb{R}$, and $H(\mathsf{X}_r, \mathsf{f}_r) = \{0\}$. It follows that the structure constants $D_N^{(r)}(\xi)$ of $(\mathsf{X}_r, \mathsf{f}_r)$ tend to infinity for all $\xi \neq 0$, and in particular for $\xi = \frac{2\pi n}{t}$, $n \neq 0$. By (5.19), $\Phi_N(x, \frac{2\pi n}{t} | \mathfrak{A}_{N+1}) \to 0$. (6.44) follows.

(c)⇒(a): We need the following lemma.

Lemma 6.9 *Fix a regular sequence of sets $\mathfrak{A}_N$, x, and $t > 0$, and suppose that $\mathbb{P}_x(S_N^{(r)} \in (a,b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N\to\infty]{} |a-b|/t$ for all intervals (a,b) such that $0 < b - a < t$. Then the convergence is uniform in (a,b).*

Proof Without loss of generality, $(a,b) \subset [0,t)$. Given $\epsilon > 0$, we need to find an N_0 such that

$$\left|\mathbb{P}_x\left(S_N^{(r)} \in (a,b) + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}\right) - \tfrac{|a-b|}{t}\right| < \epsilon \text{ for all } N > N_0 \text{ and } a < b.$$

Choose $\delta > 0$ such that

$$\frac{4\delta}{t} + \delta < \epsilon,$$

and divide $[0,t]$ into finitely many equal disjoint intervals $\{I_j\}$ with length $|I_j| < \delta$. Choose N_0 so that for all $N > N_0$, for all I_j,

$$\left|\mathbb{P}_x(S_N^{(r)} \in I_j + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) - \frac{|I_j|}{t}\right| < \frac{\delta |I_j|}{t}. \tag{6.45}$$

$I := (a,b)$ can be approximated from within and from outside by finite (perhaps empty) unions of intervals I_j whose total length differs from $|a-b|$ by at most 2δ.

Summing (6.45) over these unions we see that for all $N > N_0$,

$$\begin{aligned}\mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) &\leq \frac{|a-b| + 2\delta}{t} + \frac{\delta(|a-b| + 2\delta)}{t} \\ \mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) &\geq \frac{|a-b| - 2\delta}{t} - \frac{\delta|a-b|}{t}.\end{aligned}$$

By choice of δ, $|\mathbb{P}_x(S_N^{(r)} \in I + t\mathbb{Z} | X_{N+1} \in \mathfrak{A}_{N+1}) - \frac{|a-b|}{t}| < \epsilon$. □

We can now prove that (c)⇒(a). Suppose $(\mathsf{X}_r, \mathsf{f}_r)$ satisfies the "mixing mod t LLT" for all r and t. This property is invariant under centering, because of Lemma 6.9. So we may assume without loss of generality that $\mathbb{E}[f_n(X_n, X_{n+1})] = 0$ for all n.

First we claim that (X, f) is not center-tight. Otherwise there are constants c_N and M such that $\mathbb{P}(|S_N - c_N| > M) < 0.1$ for all N. Take $t := 5M$ and $N_k \to \infty$ such that $c_{N_k} \xrightarrow[k\to\infty]{} c \bmod t\mathbb{Z}$, then by the bounded convergence theorem and (c),

$$0.9 \le \lim_{k\to\infty} \mathbb{P}(S_{N_k} \in [c-2M, c+2M]) \le \lim_{N\to\infty} \mathbb{P}(S_N \in [c-2M, c+2M] + t\mathbb{Z})$$
$$= \int_{\mathfrak{S}_1} \lim_{N\to\infty} \mathbb{P}_x(S_N \in [c-2M, c+2M] + t\mathbb{Z} | X_{N+1} \in \mathfrak{S}_{N+1})\mu_1(\mathrm{d}x) = \frac{4M}{t} = 0.8.$$

Thus (X, f) is not center-tight, and $V_N \to \infty$.

Assume by way of contradiction that $G_{ess}(\mathsf{X}, \mathsf{f}) \neq \mathbb{R}$, then $G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ for some t, and $t \neq 0$ because $V_N \to \infty$. Without loss of generality $t = 1$, otherwise rescale f. By the reduction lemma, we can write

$$f_n(x, y) + a_n(x) - a_{n+1}(y) = \mathbb{F}_n(x, y) + h_n(x, y) + c_n$$

where $a_k, \mathbb{F}_k, h_k, c_k$ are uniformly bounded, $\mathbb{F}_n$ are integer valued, c_k are constants, h_n have summable variances, and $\mathbb{E}(h_n) = 0$.

Then $\mathfrak{F} := \sum_{n\ge1} h_n(X_n, X_{n+1})$ converges a.s., and $\mathfrak{F}_r := \sum_{n\ge r} h_n(X_n, X_{n+1}) \xrightarrow[r\to\infty]{} 0$ almost surely. Working as in the proof of (b)⇒(a), we construct $r > 1$, $x_r \in \mathfrak{S}_r$, measurable sets $\mathfrak{A}_{N+1} \subset \mathfrak{S}_{N+1}$, and intervals J_N with the following properties:

- $|\mathbb{E}_{x_r}(\mathrm{e}^{2\pi\mathrm{i}\mathfrak{F}_r})| > 0.9$,
- J_N are intervals with lengths less than 10^{-4} and centers $\zeta_N = O(1)$,
- $y \in \mathfrak{A}_{N+1} \Rightarrow a_r(x_r) - a_{N+1}(y) \in J_N$,
- and $\liminf_{N\to\infty} \mathbb{P}[X_{N+1} \in \mathfrak{A}_{N+1}] > 0$.

If $X_r = x_r$ and $X_{N+1} \in \mathfrak{A}_{N+1}$, then $|a_r(X_r) - a_{N+1}(X_{N+1}) - \zeta_N| \le |J_N|$, so

$$\left|\mathbb{E}_{x_r}(\mathrm{e}^{2\pi\mathrm{i}S_N^{(r)}} | X_{N+1} \in \mathfrak{A}_{N+1})\right| \ge \left|\mathbb{E}_{x_r}(\mathrm{e}^{2\pi\mathrm{i}(S_N^{(r)} + a_r - a_{N+1} - \zeta_N)} | X_{N+1} \in \mathfrak{A}_{N+1})\right| - 0.1$$
$$\overset{!}{\ge} |\mathbb{E}_{x_r}(\mathrm{e}^{2\pi\mathrm{i}\mathfrak{F}_r})| - 0.2 \text{ for } N \gg 1, \text{ by (6.35) for } (\mathsf{X}_r, \mathsf{f}_r - \nabla\mathsf{a}) \text{ with } s = 0, m = 1$$
$$> 0.7, \text{ by the choice of } x_r.$$

But by (c), $\mathbb{E}_{x_r}(\mathrm{e}^{2\pi\mathrm{i}S_N^{(r)}} | X_{N+1} \in \mathfrak{A}_{N+1}) \xrightarrow[N\to\infty]{} \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{\mathrm{i}u}\mathrm{d}u = 0$, a contradiction. □

6.2.4 Universal Bounds for Markov Chains

The aim of this section is to prove Theorem 6.6. We begin with two simple lemmas.

Lemma 6.10 *Suppose $\mathfrak{F}$ is a real random variable. If $b - a = L > 0$, then*

$$\left(1 - \frac{\delta}{L}\right)|a - b| < \delta \sum_{m\in\mathbb{Z}} \mathbb{E}\left[1_{(a,b)}(m\delta + \mathfrak{F})\right] < \left(1 + \frac{\delta}{L}\right)|a - b|.$$

Proof By the monotone convergence theorem,

$$\delta\mathbb{E}\left(\sum_{m\in\mathbb{Z}} 1_{(a,b)}(m\delta + \mathfrak{F})\right) \equiv \delta \cdot \mathbb{E}(\#[(a, b) \cap (\mathfrak{F} + \delta\mathbb{Z})]).$$

For each realization of $\mathfrak{F}$, $(a, b) \cap (\mathfrak{F} + \delta\mathbb{Z})$ contains at least $(|a - b|/\delta) - 1$ points, and at most $(|a - b|/\delta) + 1$ points. The lemma follows. □

Proof of Theorem 6.6: By Lemma 2.27, it is sufficient to consider the case when $\mathbb{P}[X_1 = x] = 1$ for some x. In this case $\mathbb{P}_x = \mathbb{P}, \mathbb{E}_x = \mathbb{E}$.

If $\delta(\mathsf{f}) = \infty$ then there is nothing to prove, and if $\delta(\mathsf{f}) = 0$ then $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, and we can use Theorem 5.1. It remain to consider the case when $\delta := \delta(\mathsf{f}) \in (0, \infty)$.

Suppose $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$. Let $\mathfrak{F}$ and $b_N(X_1, X_N)$ be as in Theorem 6.3.

Upper Bound (6.7)**:** Suppose (a, b) is an interval of length $L > \delta$.

Suppose $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, and write $z_N = \overline{z}_N + \zeta_N$, where $\overline{z}_N \in \delta\mathbb{Z}$ and $|\zeta_N| \le \delta$.

Fix $\varepsilon > 0$ small and choose $\phi \in C_c(\mathbb{R})$ such that

$$1_{[a-10\delta, b+10\delta]} \le \phi \le 1_{(a-10\delta-\varepsilon, b+10\delta+\varepsilon)}.$$

By Theorem 6.3, $|b_N| \le 9\delta$, so $1_{(a,b)}(S_N - z_N) \le \phi(S_N - \overline{z}_N - b_N)$, and

$$\begin{aligned}
&\limsup_{N\to\infty} \sqrt{2\pi V_N}\mathbb{P}[S_N - z_N \in (a, b)] \le \limsup_{N\to\infty} \sqrt{2\pi V_N}\mathbb{E}[\phi(S_N - \overline{z}_N - b_N)] \\
&= e^{-z^2/2}\delta \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(m\delta + \mathfrak{F})], \quad \text{by Theorem 6.3} \\
&\le e^{-z^2/2}\delta \sum_{m\in\mathbb{Z}} \mathbb{E}[1_{(a-10\delta-\varepsilon, b+10\delta+\varepsilon)}(m\delta + \mathfrak{F})], \quad \text{since } \phi \le 1_{(a-10\delta-\varepsilon, b+10\delta+\varepsilon)} \\
&\le \left(1 + \frac{\delta}{|a-b| + 20\delta + 2\varepsilon}\right) e^{-z^2/2}(|a-b| + 20\delta + 2\varepsilon), \quad \text{by Lemma 6.10} \\
&\le (|a-b| + 21\delta + 2\varepsilon)\, e^{-z^2/2} \le \left(1 + \frac{21\delta + 2\varepsilon}{L}\right) e^{-z^2/2}|a-b|.
\end{aligned}$$

Since ε is arbitrary, the result follows.

Lower Bound (6.8)**:** Fix an interval (a, b) with length bigger than some $L > \delta(\mathsf{f})$. Recall that $|b_N|$ are uniformly bounded and that $b_N = b_N(X_1, X_{N+1})$. Choose some $\overline{K}$ so that $\mathbb{P}[|b_N| \le \overline{K}] = 1$. Since $\mathbb{P}[X_1 = x] = 1$, $\mathbb{P}_x[|b_N| \le \overline{K}] = 1$.

Next, divide $[-\overline{K}, \overline{K}]$ into k disjoint intervals $I_{j,N}$ of equal length $\frac{2\overline{K}}{k}$, with k large. For each N,

$$\sum_{\mathbb{P}_x[b_N \in I_{j,N}] \ge k^{-2}} \mathbb{P}_x[b_N \in I_{j,N}] \ge 1 - \frac{1}{k},$$

because to complete the left-hand-side to one we need to add the probabilities of $[b_N \in I_{j,N}]$ for the j such that $\mathbb{P}_x[b_N \in I_{j,N}] < k^{-2}$, and $1 \le j \le k$.

Therefore, we can divide $\{I_{j,N}\}$ into two groups of size at most k: The first contains the $I_{j,N}$ with $\mathbb{P}_x[b_N \in I_{j,N}] \ge k^{-2}$, and the second corresponds to events with total probability less than or equal to $\frac{1}{k}$.

Re-index the intervals in the first group (perhaps with repetitions) in such a way that it takes the form $I_{j,N}$ $(j = 1, \ldots, k)$ for all N. For each $1 \le j \le k$, let

$$\mathfrak{A}_{j,N} := \{y \in \mathfrak{S}_{N+1} : b_N(x, y) \in I_{j,N}\}.$$

These are regular sequences of events, because by the assumption $\mathbb{P}[X_1 = x] = 1$, $\mathbb{P}[X_{N+1} \in \mathfrak{A}_{j,N}] = \mathbb{P}_x[b_N(X_1, X_{N+1}) \in I_{j,N}] \ge k^{-2}$.

Let $\beta_{j,N}$:= center of $I_{j,N}$ and set $z_{j,N} := z_N - \beta_{j,N}$. Every sequence has a subsequence such that $z_{j,N}$ converges mod $\delta\mathbb{Z}$. We will henceforth assume that $z_{j,N} = \overline{z}_{j,N} + \zeta_0 + \zeta_{j,N}$ where $\overline{z}_{j,N} \in \delta\mathbb{Z}$ and $|\zeta_{j,N}| < \overline{K}/k$, and $|\zeta_0| < \delta$ is fixed.

Recall that $|I_{j,N}| = 2\overline{K}/k$. Conditioned on $\mathfrak{A}_{j,N}$, $b_N = \beta_{j,N} \pm \frac{\overline{K}}{k}$, therefore $\overline{z}_{j,N} + \zeta_0 + b_N = z_N \pm \frac{3\overline{K}}{k}$. It follows that if $S_N - \overline{z}_{j,N} - b_N \in \left(a + \zeta_0 + \frac{3\overline{K}}{k}, b + \zeta_0 - \frac{3\overline{K}}{k}\right)$, then $S_N - z_N \in (a, b)$.

There is no loss of generality in assuming that $a + \zeta_0 \pm \frac{3\overline{K}}{k}$ are not atoms of the distribution of $\mathfrak{F}$, otherwise perturb $\overline{K}$ a little. Since $\mathfrak{A}_{j,N}$ is a regular sequence, we have by Theorem 6.3(2) and Lemma 6.10 that

$$\begin{aligned}
&\liminf_{N\to\infty} \sqrt{2\pi V_N}\mathbb{P}_x(S_N - z_N \in (a,b)|X_{N+1} \in \mathfrak{A}_{j,N}) \\
&\geq \liminf_{N\to\infty} \sqrt{2\pi V_N}\mathbb{P}_x(S_N - \overline{z}_{j,N} - b_N \in (a+\zeta_0+\tfrac{3\overline{K}}{k}, b+\zeta_0-\tfrac{3\overline{K}}{k})|X_{N+1} \in \mathfrak{A}_{j,N}) \\
&= \delta e^{-z^2/2} \sum_{m\in\mathbb{Z}} \mathbb{E}_x[1_{(a+\zeta_0+\frac{3\overline{K}}{k}, b+\zeta_0-\frac{3\overline{K}}{k})}(m\delta + \mathfrak{F})] \geq \left(1 - \tfrac{\delta}{L}\right)(|a-b| - \tfrac{6\overline{K}}{k})e^{-z^2/2}.
\end{aligned}$$

We now multiply these bounds by $\mathbb{P}_x[X_{N+1} \in \mathfrak{A}_{j,N}]$ and sum over j. This gives

$$\begin{aligned}
&\liminf_{N\to\infty} \sqrt{2\pi V_N}\mathbb{P}_x([S_N - z_N \in (a,b)] \cap \bigcup_{j=1}^{k}[X_{N+1} \in \mathfrak{A}_{j,N}]) \\
&\geq (1 - \tfrac{\delta}{L})(|a-b| - \tfrac{6\overline{K}}{k})e^{-z^2/2}\left(1 - \tfrac{1}{k}\right).
\end{aligned}$$

Passing to the limit $k \to \infty$, we obtain

$$\liminf_{N\to\infty} \sqrt{2\pi V_N}\mathbb{P}_x\left([S_N - z_N \in (a,b)]\right) \geq \left(1 - \frac{\delta}{L}\right)e^{-z^2/2}|a-b|.$$

Proof of Equation (6.9)**.** Let $\mathcal{A}$ be the positive functional on $C_c(\mathbb{R})$ defined by (6.4), and let $\mu_{\mathcal{A}}$ be the Radon measure on $\mathbb{R}$ which represents $\mathcal{A}$.

Clearly, $\mu_{\mathcal{A}}$ is invariant under translation by $\delta = \delta(\mathsf{f})$. By Lemma 6.10, $\lim_{L\to\infty} \dfrac{\mu_{\mathcal{A}}[0,L]}{L} = 1$. Necessarily, for each a, $\mu_{\mathcal{A}}[a, a+\delta) = \delta$, and

$$\forall k \in \mathbb{N} \quad \mu_{\mathcal{A}}([a, a+k\delta)) = k\delta. \tag{6.46}$$

Given $k\delta < L < (k+1)\delta$ and an interval (a,b) of length L, take two intervals I^-, I^+ such that $I^- \subset (a,b) \subset I^+$, $\mu_{\mathcal{A}}(\partial I^-) = \mu_{\mathcal{A}}(\partial I^+) = 0$, $|I^-| = k\delta$, $|I^+| = (k+1)\delta$. Choose $\phi^-, \phi^+ \in C_c(\mathbb{R})$ such that $1_{I^-} < \phi^- < 1_{[a,b]} < \phi^+ < 1_{I^+}$.

By Theorem 6.3, for large N, $e^{z^2/2}\sqrt{2\pi V_N}\mathbb{P}(S_N - z_N \in (a,b))$ is sandwiched between $\mathcal{A}(\phi^-)$ and $\mathcal{A}(\phi^+)$ which in turn is sandwiched between $\mu_{\mathcal{A}}(I^-) = k\delta$ and $\mu_{\mathcal{A}}(I^+) = (k+1)\delta$ (see (6.46)). The proof of the theorem is complete. □

6.2.5 Universal Bounds for Markov Arrays

The proof in the last section uses Theorem 6.3, and is therefore restricted to Markov chains. We will now consider the more general case of arrays.

Theorem 6.11 *Let* X *be a uniformly elliptic Markov array, and* f *an a.s. uniformly bounded additive functional which is stably hereditary and not center-tight.*[1] *Suppose* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \xrightarrow[N\to\infty]{} z \in \mathbb{R}$. *For every* $L, \epsilon > 0$ *there is* $N_\epsilon(L) > 0$ *such that for every* $[a,b] \subset [-L, L]$ *such that* $|a-b| > 2\delta(\mathsf{f}) + \epsilon$, *and for all* $N > N_\epsilon(L)$,

$$\frac{1}{3}\left(\frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\right) \leq \mathbb{P}(S_N - z_N \in (a,b)) \leq 3\left(\frac{e^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}\right).$$

Remark. Recall that $\delta(\mathsf{f}) \leq 6\text{ess sup}|\mathsf{f}|$. Hence the theorem applies to every interval with length bigger than $13\text{ess sup}|\mathsf{f}|$.

[1] In particular, the theorem applies to all a.s. uniformly bounded additive functionals on uniformly elliptic Markov chains, assuming only that $V_N \to \infty$.

The upper bound in the Theorem 6.11 holds in much greater generality, for *all* sequences z_N, and without assuming that f is stably hereditary or non center-tight:

Theorem 6.12 (Anti-Concentration Inequality) *For each K, ϵ_0 and ℓ there is a constant $C^* = C^*(K, \epsilon_0, \ell)$ such that if f is an additive functional of a uniformly elliptic Markov array with ellipticity constant ϵ_0, and if $|\mathsf{f}| \le K$ a.s., then for every $N \ge 1$, $x \in \mathfrak{S}_1^{(N)}$, and for each interval J of length ℓ,*

$$\mathbb{P}_x\left(S_N \in J\right) \le \frac{C^*}{\sqrt{V_N}}.$$

Recall that by our conventions, the Fourier transform of an L^1 function $\gamma : \mathbb{R} \to \mathbb{R}$ is $\widehat{\gamma}(x) = \int_{-\infty}^{\infty} e^{-itx}\gamma(t)dt$. Fix some $b > 0$, and let

$$\psi_b(t) := \frac{\pi}{4b}1_{[-b,b]}(t), \quad \widehat{\psi}_b(x) = \frac{\pi}{2b}\left(\frac{\sin(bx)}{x}\right).$$

Lemma 6.13 $1 \le \widehat{\psi}_b(x) \le \dfrac{\pi}{2}$ *for* $|x| \le \dfrac{\pi}{2b}$, *and* $|\widehat{\psi}_b(x)| < 1$ *for* $|x| > \dfrac{\pi}{2b}$.

Proof The function $\widehat{\psi}_b(x)$ is even, with zeroes at $z_n = \pi n/b$, $n \in \mathbb{Z} \setminus \{0\}$. The critical points are $c_0 = 0$ and $\pm c_n$, where $n \ge 1$ and c_n is the unique solution of $\tan(bc_n) = bc_n$ in $\left(z_n, z_n + \frac{\pi}{2b}\right)$. It is easy to see that $c_n = z_n + \frac{\pi}{2b} - o(1)$ as $n \to \infty$, and that $\mathrm{sgn}[\widehat{\psi}_b(c_n)] = (-1)^n$, $|\widehat{\psi}_b(c_n)| \le \frac{1}{2n}$, $\widehat{\psi}_b(c_n) \sim \frac{(-1)^n}{2n}$ as $n \to \infty$.

So $\widehat{\psi}_b$ attains the global maximum $\widehat{\psi}_b(0) = \frac{\pi}{2}$ at c_0, and $|\widehat{\psi}_b(t)| \le \frac{1}{2n}$ everywhere on $[\pi n/b, \pi(n+1)/b]$. In particular, $|\widehat{\psi}_b(t)| < 1/2$ for $|t| \ge \pi/b$.

On $(0, \pi/b)$, $\widehat{\psi}_b$ decreases from $\widehat{\psi}_b(0) = \frac{\pi}{2}$ to $\widehat{\psi}_b(\frac{\pi}{b}) = 0$, passing through $\widehat{\psi}_b(\frac{\pi}{2b}) = 1$. It follows that $1 \le \widehat{\psi}_b(t) \le \frac{\pi}{2}$ on $(0, \frac{\pi}{2b})$ and $|\widehat{\psi}_b(t)| < 1$ for $t > \frac{\pi}{2b}$. The lemma follows, because $\widehat{\psi}_b(-t) = \widehat{\psi}_b(t)$. □

Lemma 6.14 *There exist two continuous functions $\gamma_1(x), \gamma_2(x)$ such that* $\mathrm{supp}(\gamma_i) \subset [-2, 2]$*;* $\gamma_1(0) > \frac{1}{3}$*;* $\gamma_2(0) < 3$*; and* $\widehat{\gamma}_1(x) \le 1_{[-\pi,\pi]}(x) \le \widehat{\gamma}_2(x) \quad (x \in \mathbb{R})$.

Proof Throughout this proof, $\psi^{*n} := \psi * \cdots * \psi$ (n times), and $*$ denotes the convolution. Let $\gamma_1(t) := \frac{1}{4}[\psi_{\frac{1}{2}}^{*4}(t) - \psi_{\frac{1}{2}}^{*2}(t)]$. Then $\widehat{\gamma}_1(x) = \frac{1}{4}[\widehat{\psi}_{\frac{1}{2}}(x)^4 - \widehat{\psi}_{\frac{1}{2}}(x)^2]$. By Lemma 6.13, $1 \le \widehat{\psi}_{\frac{1}{2}} \le \frac{\pi}{2}$ on $[-\pi, \pi]$ and $|\widehat{\psi}_{\frac{1}{2}}| < 1$ outside $[-\pi, \pi]$. So

$$\max_{|x|\le\pi} \widehat{\gamma}_1(x) \le \max_{1\le y\le\frac{\pi}{2}} \frac{1}{4}(y^4 - y^2) = \frac{1}{4}\left[\left(\frac{\pi}{2}\right)^4 - \left(\frac{\pi}{2}\right)^2\right] < 1,$$
$$\max_{|x|\ge\pi} \widehat{\gamma}_1(x) \le \max_{|y|\le 1} \frac{1}{4}(y^4 - y^2) = 0.$$

So $\widehat{\gamma}_1(x) \le 1_{[-\pi,\pi]}(x)$ for all $x \in \mathbb{R}$.

It is obvious from the definition of the convolution that $\mathrm{supp}(\gamma_1) \subset [-2, 2]$.

Here is the calculation showing that $\gamma_1(0) > \frac{1}{3}$:

$$(\psi_b^{*2})(t) = \frac{\pi^2}{16b^2}(1_{[-b,b]} * 1_{[-b,b]})(t) = \frac{\pi^2}{16b^2}1_{[-2b,2b]}(t)(2b - |t|);$$
$$(\psi_b^{*4})(0) = (\psi_b^{*2} * \psi_b^{*2})(0)$$
$$= \frac{\pi^4}{256b^4}\int_{-2b}^{2b}(2b - |t|)^2dt = \frac{\pi^4}{128b^4}\int_0^{2b}(2b - t)^2dt = \frac{\pi^4}{128b^4}\cdot\frac{(2b)^3}{3} = \frac{\pi^4}{48b}.$$

So $\psi_{\frac{1}{2}}^{*4}(0) = \frac{\pi^4}{24}$, $\psi_{\frac{1}{2}}^{*2}(0) = \frac{\pi^2}{4}$, and $\gamma_1(0) = \frac{1}{4}(\frac{\pi^4}{24} - \frac{\pi^2}{4}) > \frac{1}{3}$.

Next we set $\gamma_2(t) := (\psi_{\frac{1}{2}} * \psi_{\frac{1}{2}})(t) \equiv \frac{\pi^2}{4}1_{[-1,1]}(t)(1 - |t|)$. Then $\mathrm{supp}(\gamma_2) = [-1, 1]$ and $\gamma_2(0) = \frac{\pi^2}{4} < 3$. Finally, $\widehat{\gamma}_2(x) \ge 1_{[-\pi,\pi]}(x)$, because by Lemma 6.13,

- $\widehat{\gamma}_2(t) = (\widehat{\psi}_{\frac{1}{2}})^2(x) \geq 1$ for all $|x| \leq \frac{\pi}{2\cdot\frac{1}{2}} = \pi$, and (trivially)
- $\widehat{\gamma}_2(t) = (\widehat{\psi}_{\frac{1}{2}})^2(x) \geq 0$ for all $|x| \geq \pi$. □

Lemma 6.15 *For every $a > 1$, $z_N \in \mathbb{R}$ and $x_1^{(N)} \in \mathfrak{S}_1$, we have*

$$\mathbb{P}_{x_1^{(N)}}(|S_N - z_N| \leq a) \geq \frac{a}{\pi}\int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})\gamma_1\left(\tfrac{a\xi}{\pi}\right)\mathrm{d}\xi, \tag{6.47}$$

$$\mathbb{P}_{x_1^{(N)}}(|S_N - z_N| \leq a) \leq \frac{a}{\pi}\int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})\gamma_2\left(\tfrac{a\xi}{\pi}\right)\mathrm{d}\xi. \tag{6.48}$$

Proof Let $\gamma_i(t)$ be the functions from Lemma 6.14, and set $I := [-a, a]$, then $\widehat{\gamma}_1\left(\frac{\pi t}{a}\right) \leq 1_I(t) \leq \widehat{\gamma}_2\left(\frac{\pi t}{a}\right)$. Therefore, for every choice of $x_1^{(N)} \in \mathfrak{S}_1^{(N)}$ $(N \geq 1)$,

$$\begin{aligned}\mathbb{P}_{x_1^{(N)}}(S_N - z_N \in I) &= \mathbb{E}_{x_1^{(N)}}[1_I(S_N - z_N)] \geq \mathbb{E}_{x_1^{(N)}}\left[\widehat{\gamma}_1\left(\frac{\pi(S_N - z_N)}{a}\right)\right] \\ &= \mathbb{E}_{x_1^{(N)}}\left[\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\frac{\pi t}{a}(S_N - z_N)}\gamma_1(t)\mathrm{d}t\right] = \int_{-\infty}^{\infty}\mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\frac{\pi t}{a}(S_N - z_N)})\gamma_1(t)\mathrm{d}t.\end{aligned}$$

Recalling that $\mathrm{supp}(\gamma_1) \subset [-2, 2]$, and substituting $t = a\xi/\pi$, we obtain (6.47). The proof of (6.48) is similar. □

Lemma 6.16 *Under the assumptions of Theorem 6.11, if $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{Z}$ and $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}$ converges to a real number z, then for every $a > 1$*

$$\sqrt{V_N}\int_{-2\pi/a}^{2\pi/a} \mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})\gamma_i\left(\frac{a\xi}{\pi}\right)\mathrm{d}\xi \xrightarrow[N\to\infty]{} \sqrt{2\pi}\mathrm{e}^{-\frac{1}{2}z^2}\gamma_i(0).$$

Moreover, the convergence is uniform on compact subsets of $a \in (1, \infty)$.

Proof. In what follows we fix $i \in \{1, 2\}$, and let $\gamma(\xi) := \gamma_i\left(\frac{a\xi}{\pi}\right)$. Divide $[-\frac{2\pi}{a}, \frac{2\pi}{a}]$ into segments I_j of length at most $\widetilde{\delta}$, where $\widetilde{\delta}$ is the constant in Lemma 5.8 and Corollary 5.10, making sure that I_0 is centered at zero. Let

$$J_{j,N} := \int_{I_j} \mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})\gamma(\xi)\mathrm{d}\xi.$$

Claim 1. $\sqrt{V_N}J_{0,N} \xrightarrow[N\to\infty]{} \sqrt{2\pi}\mathrm{e}^{-z^2/2}\gamma(0)$.

Proof of the Claim. The proof is similar to the proof of (5.28), and we use the notation of that proof. Applying Corollary 5.10 to the interval I_0, and noting that $A_N(I_0) = 0$ and $\widetilde{\xi}_N = 0$, we find that $|\mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})| \leq \widetilde{C}\exp(-\widehat{\varepsilon}\xi^2 V_N)$.

So for $R > 1$,

$$\sqrt{V_N}\int_{[\xi\in I_0:|\xi|>\frac{R}{\sqrt{V_N}}]} \mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)})\gamma(\xi)\mathrm{d}\xi = O(\mathrm{e}^{-\widehat{\varepsilon}R^2}).$$

Similarly, for all N large enough,

$$\sqrt{V_N}\int_{[\xi\in I_0:|\xi|\leq\frac{R}{\sqrt{V_N}}]} \mathbb{E}_{x_1^{(N)}}\left(\mathrm{e}^{-\mathrm{i}\xi(S_N - z_N)}\right)\gamma(\xi)\mathrm{d}\xi = \int_{-R}^{R}\mathbb{E}_{x_1^{(N)}}\left(\mathrm{e}^{-\mathrm{i}\eta\frac{S_N - z_N}{\sqrt{V_N}}}\right)\gamma\left(\frac{\eta}{\sqrt{V_N}}\right)\mathrm{d}\eta =$$

$$\int_{-R}^{R}\mathbb{E}_{x_1^{(N)}}\left(\mathrm{e}^{-\mathrm{i}\eta\frac{S_N-\mathbb{E}(S_N)}{\sqrt{V_N}}}\right)\mathrm{e}^{\mathrm{i}\eta\frac{z_N-\mathbb{E}(S_N)}{\sqrt{V_N}}}\gamma\left(\frac{\eta}{\sqrt{V_N}}\right)\mathrm{d}\eta \overset{!}{=} \int_{-R}^{R}\mathrm{e}^{-\frac{1}{2}\eta^2+\mathrm{i}\eta z}\gamma(0)\mathrm{d}\eta+o_{N\to\infty}(1) \quad \text{uniformly on compacts}$$
$$=\sqrt{2\pi}\mathrm{e}^{-\frac{1}{2}z^2}\gamma(0)+o_{R\to\infty}(1)+o_{N\to\infty}(1). \tag{6.49}$$

Let us justify the equality (6.49). Arguing as in the proof of (6.20), one shows that $|\mathbb{E}(S_N)-\mathbb{E}_{x_1^{(N)}}(S_N)| = O(1)$ and $\mathrm{Var}(S_N|X_1^{(N)} = x_1^{(N)}) \sim V_N$, therefore

$$\frac{z_N-\mathbb{E}(S_N)}{\sqrt{V_N}} \to z \Leftrightarrow \frac{z_N-\mathbb{E}_{x_1^{(N)}}(S_N)}{\sqrt{\mathrm{Var}(S_N|X_1^{(N)}=x_1^{(N)})}} \to z.$$

(6.49) follows from Dobrushin's CLT for $\mathsf{Y}_x := \mathsf{X}$, conditioned on $X^{(N)} = x_1^{(N)}$.

In summary, $\sqrt{V_N}J_{0,N} = \sqrt{2\pi}\mathrm{e}^{-\frac{1}{2}z^2}\gamma(0)+o_{R\to\infty}(1)+o_{N\to\infty}(1)$. Fixing R, we see that $\limsup\sqrt{V_N}J_{0,N}$ and $\liminf\sqrt{V_N}J_{0,N}$ are both equal to $\sqrt{2\pi}\mathrm{e}^{-\frac{1}{2}z^2}\gamma(0)+o_{R\to\infty}(1)$.

Passing to the limit $R\to\infty$ gives us that the limit exists and is equal to $\sqrt{2\pi}\mathrm{e}^{-\frac{1}{2}z^2}\gamma(0)$. The convergence is uniform on compact subsets of a.

Claim 2. $\sqrt{V_N}J_{j,N} \xrightarrow[N\to\infty]{} 0$ *for every* $j\neq 0$.

Proof of the Claim. Since $G_{ess}(\mathsf{X},\mathsf{f})=\mathbb{Z}$, the co-range is $H(\mathsf{X},\mathsf{f})=2\pi\mathbb{Z}$. So

$$I_j \subset [-\tfrac{2\pi}{a},\tfrac{2\pi}{a}]\setminus\mathrm{int}(I_0) \subset \text{ a compact subset of } \mathbb{R}\setminus H(\mathsf{X},\mathsf{f}).$$

This implies by the stable hereditary property of f that $D_N(\xi)\xrightarrow[N\to\infty]{}\infty$ uniformly on I_j, whence by (5.18), $|\mathbb{E}_{x_1^{(N)}}(\mathrm{e}^{-\mathrm{i}\eta(S_N-z_N)})| \xrightarrow[N\to\infty]{} 0$ uniformly on I_j.

Let $A_{j,N} := -\log\{\sup|\mathbb{E}_x(\mathrm{e}^{-\mathrm{i}\xi(S_N-z_N)})| : (x,\xi)\in\mathfrak{S}_1^{(N)}\times I_j\}$, then $A_{j,N}\xrightarrow[N\to\infty]{}\infty$, and this divergence is uniform for a ranging over compact subsets of $(1,\infty)$.

From this point onward, the proof of the claim is identical to the proof of (5.29). We omit the details.

The lemma follows by summing over all subintervals I_j in $[-\frac{2\pi}{a},\frac{2\pi}{a}]$, and noting that the number of these intervals is uniformly bounded $\left(\text{by } 1+\frac{4\pi}{\delta}\right)$. □

Proof of Theorem 6.11. If $G_{ess}(\mathsf{X},\mathsf{f})=\mathbb{R}$ then the theorem follows from the LLT in the irreducible case. Otherwise (since f is not center-tight), $G_{ess}(\mathsf{X},\mathsf{f})=t\mathbb{Z}$ for some $t>0$, and there is no loss of generality in assuming that $G_{ess}(\mathsf{X},\mathsf{f})=\mathbb{Z}$.

In this case our interval $I:=[a,b]$ has length bigger than 2. Notice that we can always center I by modifying z_N by a constant. So we may take our interval to be of the form $I=[-a,a]$, with $a>1$.

Lemma 6.16, (6.47), (6.48), and the inequalities $\gamma_1(0)>\frac{1}{3}$ and $\gamma_2(0)<3$ imply that for every choice of $\{x_1^{(N)}\}_{N\geq 1}$, for all N sufficiently large,

$$\frac{1}{3}\cdot\frac{|I|}{\sqrt{2\pi V_N}}\mathrm{e}^{-z^2/2} \leq \mathbb{P}_{x_1^{(N)}}(S_N-z_N\in I) \leq 3\cdot\frac{|I|}{\sqrt{2\pi V_N}}\mathrm{e}^{-z^2/2}. \tag{6.50}$$

Thus we have proved the theorem for all Markov arrays with point mass initial distributions. By Lemma 2.27, the theorem follows for general arrays. □

Proof of Theorem 6.12. It is sufficient to prove the theorem for N such that $V_N\geq 1$. If $V_N<1$, the theorem holds (trivially) provided that $C^*\geq 1$.

It is also sufficient to prove the result for intervals J with length 4, since longer intervals can be covered by no more than $|J|/4+1$ such intervals. Thus $J=\zeta+[-2,2]$ and ζ :=center of J. By (6.48) (with $a=2$ and $z_N:=\zeta$),

$$\mathbb{P}_x(S_N \in J) \le \frac{2}{\pi}\|\gamma_2\|_\infty \int_{-\pi}^{\pi} |\Phi_N(x,\xi)|\,\mathrm{d}\xi.$$

To prove the theorem, we need to bound the integral from above.

The argument is similar to the previous proof, except that we cannot assume that $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{Z}$ or $V_N \to \infty$, and we must pay closer attention to the uniformity of the estimates in N (the statement in Theorem 6.12 is for *all* N, not just for $N \gg 1$).

Recall the notation

$$A_N(I) := -\log\sup\left\{|\Phi_N(x,\xi)| : (x,\xi) \in \mathfrak{S}_1^{(N)} \times I\right\},$$

and let $(\widetilde{x}_N(I), \widetilde{\xi}_N(I)) \in I \times \mathfrak{S}_1^{(N)}$ be a pair where

$$|A_N(I)| \le -\log|\Phi_N(\widetilde{x}_N(I), \widetilde{\xi}_N(I))| + \log 2.$$

By Lemma 5.8 and Corollary 5.10, there are constants $\widetilde{\delta}, \widetilde{C}, \widehat{\varepsilon}, \overline{c} > 0$ which depend only on ϵ_0 and K, so that if $|I| \le \widetilde{\delta}$, then for all $(x,\xi) \in \mathfrak{S}_1^{(N)} \times I$ and N,

$$|\Phi_N(x,\xi)| \le \widetilde{C}\exp\left(-\widehat{\varepsilon}\,V_N(\xi - \widetilde{\xi}_N(I))^2 + \overline{c}|\xi - \widetilde{\xi}_N(I)|\sqrt{V_N A_N(I)}\right). \tag{6.51}$$

We now divide $[-\pi,\pi]$ into no more than $4\pi/\widetilde{\delta} + 1$ intervals of length at most $\widetilde{\delta}/2$. We claim that for each interval I in our partition, $\int_I |\Phi(x,\xi)|\mathrm{d}\xi \le \text{const.}V_N^{-1/2}$.

To prove this we consider two cases.

(1) Suppose $A(I) \le 1$, then $I \subset [\widetilde{\xi}_N(J) - \widetilde{\delta}/2, \widetilde{\xi}_N(J) + \widetilde{\delta}/2]$, and by (6.51),

$$\int_I \Phi_N(x,\xi)|\mathrm{d}\xi \le \widetilde{C}\int_{-\widetilde{\delta}/2}^{\widetilde{\delta}/2} \mathrm{e}^{-\widehat{\varepsilon}V_N\xi^2 + \overline{c}|\xi|\sqrt{V_N}}\mathrm{d}\xi \le \text{const.}V_N^{-1/2},$$

with the constant only depending on $\widehat{\varepsilon}$ and $\overline{c}$, whence only on K and ϵ_0.

(2) If $A_N(I) \ge 1$, then (5.30) (with $R = 1$) gives $\int_I |\Phi_N(x,\xi)|\mathrm{d}\xi \le \text{const.}V_N^{-1/2}$, with the constant only depending on K and ϵ_0.

Summing over all intervals I, and recalling that there are at most $4\pi/\widetilde{\delta} + 1$ such intervals, we obtain $\int_{-\pi}^{\pi} |\Phi_N(x,\xi)|\mathrm{d}\xi \le \text{const.}V_N^{-1/2}$, with the constant only depending on ϵ_0 and K. As explained at the beginning of the proof, this implies the theorem. □

It is interesting to note that the anti-concentration inequality could be used to provide a different justification for Stone's trick of using $\phi \in L^1$ with Fourier transform with compact support in the proof of the LLT (see §5.2.1).

Namely let f be an irreducible additive functional with algebraic range $\mathbb{R}$, such that $\mathbb{E}(S_N) = 0$ for all N. Approximating $1_{[a,b]}$ from above and below by compactly supported functions we see that the LLT follows if one could show that

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}\left(\phi(S_N - z_N)\right) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\int_{\mathbb{R}} \phi(x)dx \tag{6.52}$$

for each C^2 compactly supported function ϕ and every sequence z_N such that $\lim_{N\to\infty} \frac{z_N}{\sqrt{V_N}} = z$. Fix a small $\varepsilon > 0$ and let $\overline{\phi}$ be a function such that the Fourier transform of $\overline{\phi}$ has compact support, $\int_{\mathbb{R}} \overline{\phi}(x)dx = \int_{\mathbb{R}} \phi(x)dx$ and [2]

[2] To find such a function take a large L and define $\overline{\phi}$ by the condition that the Fourier transforms of ϕ and $\overline{\phi}$ are related by $\widehat{\overline{\phi}}(\xi) = \widehat{\phi}(\xi)\chi_L(\xi)$ where χ_L is smooth, $\chi_L(\xi) = 1$ on $-[L,L]$, $0 \le \chi_L(\xi) \le 1$ for $|\xi| \in [L, L+1]$, and $\chi_L(\xi) = 0$ for $\xi \notin [-(L+1), L+1]$. To verify (6.53) we use the Fourier inversion formula, and the obvious inequality $|\widehat{\psi}(\xi)| \le \frac{1}{\xi^2}\int |\psi''(x)|dx$.

$$\left|\overline{\phi}(x) - \phi(x)\right| \le \frac{\varepsilon}{1+x^2}. \tag{6.53}$$

Combining (6.53) and Theorem 6.12 we get

$$\left|\mathbb{E}\left(\phi(S_N - z_N) - \overline{\phi}(S_N - z_N)\right)\right| \le \varepsilon \sum_{j=0}^{\infty} \frac{\mathbb{P}(|S_N - z_N| \in [j, j+1))}{1+j^2} \le 2C^* \varepsilon$$

where C^* is the constant obtain by applying Theorem 6.12 with $\ell = 1$.

Since ε is arbitrary, we see that to prove the LLT, it is sufficient to show (6.52) for all ϕ with *compactly supported* Fourier transform. This justifies "Stone's trick." We will meet this idea again, when we discuss large deviations, see §7.3.8.

6.3 Notes and References

Theorem 6.3 extends an earlier result for sums of independent random variables, due to Dolgopyat [56]. In this case, one can take b_N to be constants, see Theorem 8.3(2b) and the discussion in §8.2.

The connection between the LLT and mixing mod t LLT was considered for sums of independent random variables by Prokhorov [163], Rozanov [169], and Gamkrelidze [76].

As far as we know, the first paper devoted to the perturbative analysis of non stationary product of transfer operators is due to Bakhtin [12].

The study of the concentration function $\Lambda_N(h) = \sup_{x\in\mathbb{R}} \mathbb{P}(S_N \in [x, x+h])$ goes back to works of Paul Lévy [130] and Wolfgang Doeblin [53]. The idea to use the characteristic function to study this function is due to Esseen [71, 72]. We refer the reader to [156, Chapter III] for a detailed discussion, and the history of the subject. Our proof of the anti-concentration inequality (Theorem 6.12) follows [156, Section III.1] closely. [156] considers independent random variables, but given the results of §5.2.2, the argument in the Markov case is essentially the same.

For long intervals, the universal bounds in §6.2.4 and 6.2.5 can be obtained from the a **Berry-Esseen Estimate** for the rate of convergence in the CLT. Suppose we could show that $\exists L$ s.t.

$$\sup_{z\in\mathbb{R}} \left| \mathbb{P}\left(\frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \le z\right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-t^2/2} \mathrm{d}t \right| \le \frac{L}{\sqrt{V_N}}.$$

Then $\exists M$ such that for all $|a-b| > M$, if $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, then for all N large enough, $\mathbb{P}[S_N - z_N \in (a, b)]$ equals $\dfrac{\mathrm{e}^{-z^2/2}|a-b|}{\sqrt{2\pi V_N}}$ up to bounded multiplicative error.

For the additive functionals considered in this monograph the Berry-Esseen estimate has been obtained in [59] using the results from Chapters 3–5. The Berry-Esseen approach has the advantage of giving information on the time N when the universal estimates kick in, but it only applies to large intervals (the largeness depends on the bound on $\sup_k \|f_k\|_\infty$ but it does not take into account the graininess constant $\delta(\mathsf{f})$). By contrast, the results of §6.2.5 apply to intervals of length larger than $\delta(\mathsf{f})$, which is optimal, but do not say how large N should be for the estimates to work.

Chapter 7
Local Limit Theorems for Moderate Deviations and Large Deviations

Abstract We prove the local limit theorem in the regimes of moderate deviations and large deviations. In these cases the asymptotic behavior of $\mathbb{P}(S_N - z_N \in (a,b))$ is determined by the Legendre transforms of the log-moment generating functions.

7.1 Moderate Deviations and Large Deviations

Suppose f is an irreducible and a.s. uniformly bounded additive functional on a uniformly elliptic Markov *chain* X. Let

$$S_N = f_1(X_1, X_2) + \cdots + f_N(X_N, X_{N+1}), \ V_N := \mathrm{Var}(S_N),$$

and suppose $V_N \to \infty$.

In the previous chapters, we analyzed $\mathbb{P}(S_N - z_N \in (a,b))$ as $N \to \infty$, in the regime of **local deviations**, $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to \text{const}$. In this chapter we consider the following more general scenarios, which include cases when $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to \infty$:

(1) **Moderate Deviations:** $z_N - \mathbb{E}(S_N) = o(V_N)$,

(2) **Large Deviations:** $|z_N - \mathbb{E}(S_N)| \leq \epsilon V_N$ for some $\epsilon > 0$ small enough.

We should explain why we did not define the large deviations regime by the more natural condition that $|z_N - \mathbb{E}(S_N)| \geq \overline{\varepsilon} V_N$ for some $\overline{\varepsilon} > 0$. We should also explain the role of the upper bound on $|z_N - \mathbb{E}(S_N)|/V_N$.

The decision not to impose a lower bound on $|z_N - \mathbb{E}(S_N)|/V_N$ is mainly a matter of convenience; It allows us to view moderate deviations as a special case of large deviations, and handle the two regimes simultaneously. The decision to impose an upper bound on $|z_N - \mathbb{E}(S_N)|/V_N$ reflects a limitation of our methods: We do not know how to handle the degeneracies which may occur when $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ is "too large." Let us indicate briefly what could go wrong in this case.

The most extreme scenario is when $\frac{z_N - \mathbb{E}(S_N)}{V_N} > r_N$, where $r_N = \frac{\text{ess sup}\, S_N - \mathbb{E}(S_N)}{V_N}$. In this case, $\mathbb{P}[S_N - z_N \in (0,\infty)] = 0$ for all N. A more subtle degeneracy may happen when $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ falls near the boundary of the domain of the Legendre transform of $\mathcal{F}_N^c(t) := \frac{1}{V_N} \log \mathbb{E}(e^{t(S_N - \mathbb{E}(S_N))})$ (Legendre transforms are discussed in §7.2.2.). The following example shows that in this case, the probabilities $\mathbb{P}[S_N - z_N \in (a,b)]$ may be so sensitive to z_N, that they could have different asymptotic behaviors for $z_N^{(1)}, z_N^{(2)}$ with the same limit $\lim_{N\to\infty} \frac{z_N^{(i)} - \mathbb{E}(S_N)}{V_N}$:

Example 7.1 Let $S_N := X_1 + \cdots + X_N$, where X_i are identically distributed independent random variables, equal to $-1, 0, 1$ with equal probabilities.

Here $\mathbb{E}(S_N) = 0$, $V_N = 2N/3$, and the Legendre transforms of the log-moment generating functions have domains $(-\frac{3}{2}, \frac{3}{2})$. Clearly: if $z_N = N$, then $\mathbb{P}[S_N - z_N \in (0,2)] = 0$; if $z_N = N-1$, then $\mathbb{P}[S_N - z_N \in (0,2)] = 3^{-N}$;

if $z_N = N-2$, then $\mathbb{P}[S_N - z_N \in (0,2)] = 3^{-N}N$. In all cases, $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to \frac{3}{2}$, but the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (0,2)]$ is completely different.

The assumption that $|z_N - \mathbb{E}(S_N)| < \varepsilon V_N$ with ε "small enough" guarantees that $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ falls "well inside" the domain of the Legendre transform of $\mathcal{F}_N^c$, and this prevents these pathologies. A detailed discussion of the sequences $\{z_N\}$ to which our results apply appears in Section 7.4.

It is instructive to compare the regime of large deviations to the regime of local deviations from the point of view of universality.

In the regime of local deviations, the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (a,b)]$ does not depend on the details of the distributions of $f_n(X_n, X_{n+1})$. It depends only on rough features such as $\mathrm{Var}(S_N)$, the algebraic range, and (in case the algebraic range is $t\mathbb{Z}$) on the constants c_N s.t. $S_N \in c_N + t\mathbb{Z}$ almost surely.

By contrast, in the regime of large deviations the asymptotic behavior of $\mathbb{P}[S_N - z_N \in (a,b)]$ depends on the entire distribution of S_N. The dependence is through the Legendre transform of $\log \mathbb{E}(e^{tS_N})$, a function which encodes the entire distribution of S_N, not just its rough features.

We will consider two partial remedies to the lack of universality:

(a) *Conditioning:* The distribution of $S_N - z_n$ conditioned on $S_N - z_N > a$ has a universal scaling limit, see Corollary 7.10.

(b) *Moderate Deviations:* If $|z_N - \mathbb{E}(S_N)| = o(\mathrm{Var}(S_N))$, then $\mathbb{P}[S_N - z_N \in (a,b)]$ have universal lower and upper bounds (Theorems 7.5 and 7.6).

7.2 Local Limit Theorems for Large Deviations

7.2.1 The Log Moment Generating Functions

Suppose $|\mathsf{f}| < K$ almost surely. For every N such that $V_N \neq 0$, we define the **normalized log moment generating function** of S_N to be

$$\mathcal{F}_N(\xi) := \frac{1}{V_N} \log \mathbb{E}(e^{\xi S_N}) \quad (\xi \in \mathbb{R}).$$

The uniform boundedness of f implies the finiteness of the expectation, and the real-analyticity of $\mathcal{F}_N(\xi)$ on $\mathbb{R}$.

Example 7.2 (Sums of IID's) Let $S_N = \sum_{n=1}^{N} X_n$ where X_n are iid bounded random variables with non-zero variance. Let X denote the common law of X_n. Then $\mathcal{F}_N(\xi) = \mathcal{F}_X(\xi) := \frac{1}{\mathrm{Var}(X)} \log \mathbb{E}(e^{\xi X})$, for all N. Clearly,

(i) $\mathcal{F}_N(0) = 0$, $\mathcal{F}_N'(0) = \mathbb{E}(X)/\mathrm{Var}(x)$ and $\mathcal{F}_N''(0) = 1$. (ii) $\mathcal{F}_N(\xi)$ are uniformly strictly convex on compacts.

These properties play a key role in the study of large deviations for sums of i.i.d. random variables. A significant part of the effort in this chapter is to understand to which extent similar results hold in the setting of bounded additive functionals of uniformly elliptic Markov chains. We start with the following facts.

Theorem 7.3 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*, and assume* $V_N \neq 0$ *for all* $N \geq N_0$*. Then:*

(1) $\forall N \geq N_0$, $\mathcal{F}_N(0) = 0$, $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$ *and* $\mathcal{F}_N''(0) = 1$. *(2)* $\forall N \geq N_0$, $\mathcal{F}_N(\xi)$ *is strictly convex on* $\mathbb{R}$.

(3) The convexity is uniform on compacts: For every $R > 0$ *there is* $C = C(R)$ *positive such that for all* $N \geq N_0$,

$$C^{-1} \leq \mathcal{F}_N''(\xi) \leq C \text{ on } [-R, R].$$

(4) Suppose $V_N \to \infty$. $\forall \epsilon > 0\ \exists \delta, N_\epsilon > 0$ *such that for all* $|\xi| \leq \delta$ *and* $N > N_\epsilon$*, we have* $e^{-\epsilon} \leq \mathcal{F}_N''(\xi) \leq e^{\epsilon}$, *and* $e^{-\epsilon}\left(\frac{\xi^2}{2}\right) \leq \mathcal{F}_N(\xi) - \frac{\mathbb{E}(S_N)}{V_N}\xi \leq e^{\epsilon}\left(\frac{\xi^2}{2}\right)$.

This is very similar to what happens for iid's, but there is one important difference: in our setting, V_N may be much smaller than N. For the proof of this theorem, see §7.3.5.

7.2.2 The Rate Functions

Suppose $V_N \neq 0$. The **rate functions** $\mathcal{I}_N(\eta)$ are the **Legendre transforms** of $\mathcal{F}_N(\xi)$. Specifically, if $a_N := \inf_\xi \mathcal{F}_N'(\xi)$ and $b_N := \sup_\xi \mathcal{F}_N'(\xi)$, then $\mathcal{I}_N : (a_N, b_N) \to \mathbb{R}$ is

$$\mathcal{I}_N(\eta) := \xi\eta - \mathcal{F}_N(\xi), \text{ for the unique } \xi \text{ s.t. } \mathcal{F}_N'(\xi) = \eta.$$

The existence and uniqueness of ξ when $\eta \in (a_N, b_N)$ is because of the smoothness and strict convexity of $\mathcal{F}_N$ on $\mathbb{R}$. We call (a_N, b_N) the **domain** of $\mathcal{I}_N$, and write

$$\mathrm{dom}(\mathcal{I}_N) := (a_N, b_N).$$

Equivalently,

$$\mathrm{dom}(\mathcal{I}_N) = (\mathcal{F}'(-\infty), \mathcal{F}'(+\infty)), \text{ where } \mathcal{F}'(\pm\infty) := \lim_{t\to\pm\infty} \mathcal{F}'(t).$$

Later we will also need the sets $(a_N^R, b_N^R) \subset \mathrm{dom}(\mathcal{I}_N)$, where $R > 0$ and

$$a_N^R := \mathcal{F}_N'(-R), \quad b_N^R := \mathcal{F}_N'(R). \tag{7.1}$$

The functions $\mathcal{I}_N$ and their domains depend on N. The following theorem identifies certain uniformity and universality in their behavior.

Theorem 7.4 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*, and assume* $V_N \neq 0$ *for all* N *large enough, then:*

(1) $\exists c, N_1, R > 0$ *such that for all* $N > N_1$,

$$\mathrm{dom}(\mathcal{I}_N) \supset [a_N^R, b_N^R] \supset \left[\frac{\mathbb{E}(S_N)}{V_N} - c, \frac{\mathbb{E}(S_N)}{V_N} + c\right].$$

(2) For each R *there exists* $\rho = \rho(R)$ *such that for all* $N > N_1$,

$$\rho^{-1} \leq \mathcal{I}_N'' \leq \rho \text{ on } [a_N^R, b_N^R].$$

(3) Suppose $V_N \to \infty$*. For every* $\epsilon > 0$ *there exists* $\delta > 0$ *and* N_ϵ *such that for all* $\eta \in [\frac{\mathbb{E}(S_N)}{V_N} - \delta, \frac{\mathbb{E}(S_N)}{V_N} + \delta]$ *and* $N > N_\epsilon$,

$$e^{-\epsilon}\frac{1}{2}\left(\eta - \frac{\mathbb{E}(S_N)}{V_N}\right)^2 \leq \mathcal{I}_N(\eta) \leq e^{\epsilon}\frac{1}{2}\left(\eta - \frac{\mathbb{E}(S_N)}{V_N}\right)^2.$$

(4) Suppose $V_N \to \infty$ *and* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$*, then*

$$V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right) = \frac{1 + o(1)}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2 \quad \text{as} \quad N \to \infty.$$

The proof of the theorem is given in §7.3.6.

7.2.3 The LLT for Moderate Deviations

Theorem 7.5 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X. *Suppose* f *is irreducible with algebraic range* $\mathbb{R}$. *If* $z_N \in \mathbb{R}$ *satisfy* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, *then for every non-empty interval* (a, b), *when* $N \to \infty$,

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1+o(1)]\frac{|a-b|}{\sqrt{2\pi V_N}} \exp\left(-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)\right),$$
$$\mathbb{P}[S_N - z_N \in (a,b)] = [1+o(1)]\frac{|a-b|}{\sqrt{2\pi V_N}} \exp\left[-\frac{1+o(1)}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right].$$

Theorem 7.6 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X. *Assume* f *is irreducible with algebraic range* $\mathbb{Z}$, *and* $S_N \in c_N + \mathbb{Z}$ *a.s. If* $z_N \in c_N + \mathbb{Z}$ *and* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, *then*

$$\mathbb{P}[S_N = z_N] = \frac{[1+o(1)]}{\sqrt{2\pi V_N}} \exp\left(-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)\right),$$
$$\mathbb{P}[S_N = z_N] = \frac{[1+o(1)]}{\sqrt{2\pi V_N}} \exp\left[-\frac{1+o(1)}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right] \text{ as } N \to \infty.$$

We will deduce these results from the more general Theorem 7.8, below.

Remark. The first asymptotic relation in Theorems 7.5 and 7.6 is not universal, because of the dependence on $\mathcal{I}_N$. The second asymptotic relation is universal, but it is not a proper asymptotic equivalence because of the $o(1)$ in the exponent.

The following result provides less information than Theorems 7.5 and 7.6, but requires no irreducibility assumptions:

Theorem 7.7 *Suppose* f *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X. *If* $V_N \to \infty$, *then for all* $0 < \alpha < \frac{1}{2}$ *and* $\kappa > 0$, *if* $\dfrac{z_N - \mathbb{E}(S_N)}{V_N} \sim \kappa V_N^{-\alpha}$ *as* $N \to \infty$, *then*

$$\lim_{N\to\infty} \frac{1}{V_N^{1-2\alpha}} \log \mathbb{P}[S_N - z_N \geq 0] = -\frac{1}{2}\kappa^2.$$

Proof There is no loss of generality in assuming that $\mathbb{E}(S_N) = 0$ for all N. Let $a_n := V_n^{1-2\alpha}$, $b_n := V_n^{\alpha}$, $W_n := S_n/b_n$. Then $a_n \to \infty$, and $\lim\limits_{n\to\infty} \dfrac{1}{a_n} \log \mathbb{P}[S_n - z_n \geq 0] = \lim\limits_{n\to\infty} \dfrac{1}{a_n} \log \mathbb{P}[W_n/a_n \geq \kappa + o(1)]$. By Theorem 7.3(4), $\lim\limits_{n\to\infty} \frac{1}{a_n} \log \mathbb{E}(\mathrm{e}^{\xi W_n}) = \lim\limits_{n\to\infty} V_n^{2\alpha} \mathcal{F}_N(\frac{\xi}{V_n^{\alpha}}) = \frac{1}{2}\xi^2$. Thus, by the Gärtner-Ellis Theorem (see Appendix A), $\lim\limits_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[\frac{W_n}{a_n} \geq \kappa + o(1)] = -\frac{1}{2}\kappa^2$. □

7.2.4 The LLT for Large Deviations

Recall that $(a_N^R, b_N^R) := (\mathcal{F}_N'(-R), \mathcal{F}_N'(R)) \subset \mathrm{dom}(\mathcal{I}_N)$. It is convenient to define

$$[\widehat{a}_N^R, \widehat{b}_N^R] := \left[a_N^R - \frac{\mathbb{E}(S_N)}{V_N}, b_N^R - \frac{\mathbb{E}(S_N)}{V_N}\right].$$

Theorem 7.8 *Let* f *be an a.s. uniformly bounded, irreducible, additive functional on a uniformly elliptic Markov chain* X. *For every* R *large enough there are functions* $\rho_N : \left[\widehat{a}_N^R, \widehat{b}_N^R\right] \to \mathbb{R}^+$ *and* $\xi_N : [\widehat{a}_N^R, \widehat{b}_N^R] \to \mathbb{R}$ *as follows:*

(1) $\exists c > 0$ *such that* $[\widehat{a}_N^R, \widehat{b}_N^R] \supset [-c, c]$ *for all* N *large enough.*

(2) **Non Lattice Case:** *Suppose* $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$*. For every sequence of* $z_N \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$*, and for all finite non-empty intervals* (a, b)*, we have the following asymptotic result as* $N \to \infty$*:*

$$\mathbb{P}[S_N - z_N \in (a, b)] = [1 + o(1)] \cdot \frac{e^{-V_N \mathcal{I}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} |a - b| \rho_N \left(\tfrac{z_N - \mathbb{E}(S_N)}{V_N}\right) \times \frac{1}{|a - b|} \int_a^b e^{-t\xi_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)} dt.$$

(3) **Lattice Case:** *Suppose* $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{Z}$ *and* $S_N \in c_N + \mathbb{Z}$ *a.s., then for every sequence of* $z_N \in c_N + \mathbb{Z}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$*, the following asymptotic relation holds when* $N \to \infty$*:*

$$\mathbb{P}[S_N = z_N] = [1 + o(1)] \cdot \frac{e^{-V_N \mathcal{I}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} \times \rho_N \left(\tfrac{z_N - \mathbb{E}(S_N)}{V_N}\right).$$

(4) **Properties of the Error Terms:**

(a) $\rho_N(\eta)$ *are bounded away from* 0 *and* ∞ *on* $[\widehat{a}_N^R, \widehat{b}_N^R]$*, uniformly in* N*, and* $\rho_N(\eta) \xrightarrow[\eta \to 0]{} 1$ *uniformly in* N*.*

(b) For each $R > 0$ *there exists* $C = C_R > 0$ *such that for all* $\eta \in [\widehat{a}_N^R, \widehat{b}_N^R]$ *and* N*,* $C^{-1}|\eta| \le |\xi_N(\eta)| \le C|\eta|$ *and* $\mathrm{sgn}(\xi(\eta)) = \mathrm{sgn}(\eta)$*.*

Warning. ρ_N depends on the initial distribution.

Theorem 7.8 assumes irreducibility. The following coarser result, does not:

Theorem 7.9 *Suppose* f *is an additive functional on a uniformly elliptic Markov chain* X *such that* $K := \operatorname{ess\,sup}|\mathsf{f}| < \infty$*, and suppose* $V_N \to \infty$*. For each* $\varepsilon, R > 0$ *there is* $D(\varepsilon, R, K)$ *and* N_0 *such that for all* $N > N_0$*, if* $\frac{z_N}{V_N} \in [\mathcal{F}_N'(\varepsilon), \mathcal{F}_N'(R)]$*, then*

$$\frac{D^{-1}}{\sqrt{V_N}} e^{-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)} \le \mathbb{P}(S_N \ge z_N) \le \frac{D}{\sqrt{V_N}} e^{-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)}.$$

Theorem 7.8 is proved in §§7.3.1–7.3.7. Theorem 7.9 is proved in §7.3.8.

To assist the reader to digest the statement of Theorem 7.8, let us see how to use it to obtain Theorems 7.5 and 7.6 on moderate deviations.

Proof of Theorems 7.5 and 7.6: Fix $R > 0$, and suppose $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$. By Theorem 7.8(1), $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ for all N large enough. By Theorem 7.8(4), $\rho_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right) \xrightarrow[N \to \infty]{} 1$, $\xi_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right) \to 0$ and $\frac{1}{b-a}\int_a^b e^{-t\xi_N(\frac{z_N - \mathbb{E}(S_N)}{V_N})} dt \to 1$. Thus by Theorem 7.8(2), if $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, then

$$\mathbb{P}[S_N - z_N \in (a, b)] \sim \frac{|a - b|}{\sqrt{2\pi V_N}} \exp\left(-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)\right).$$

By Theorem 7.4(4), $V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right) \sim \frac{1}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2$. Hence $\mathbb{P}[S_N - z_N \in (a, b)] \sim \frac{|a - b|}{\sqrt{2\pi V_N}} \exp\left(-\frac{1 + o(1)}{2}\left(\frac{z_n - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2\right)$. This proves Theorem 7.5. The proof of Theorem 7.6 is similar, and we omit it. □

Corollary 7.10 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Suppose* f *is irreducible, with algebraic range* $\mathbb{R}$*.*

(1) If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$*, then for any finite non empty interval* (a, b)*, the distribution of* $S_N - z_N$ *conditioned on* $S_N - z_N \in (a, b)$ *is asymptotically uniform on* (a, b)*.*

(2) If $\liminf \frac{z_N - \mathbb{E}(S_N)}{V_N} > 0$ *and there exists* R *such that* $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ *for all sufficiently large* N*, then the distribution of* $\xi_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right) \cdot (S_N - z_N)$ *conditioned on* $S_N > z_N$ *is asymptotically exponential with parameter 1.*

Remark. Condition (2) holds when $\liminf \frac{z_N-\mathbb{E}(S_N)}{V_N}>0$, and $\limsup \frac{z_N-\mathbb{E}(S_N)}{V_N}>0$ is small, see Theorem 7.8(1).

Proof To see part (1), note using Theorem 7.8(4) that if $\frac{z_N-\mathbb{E}(S_N)}{V_N} \to 0$, then $\xi_N = \xi_N(\frac{z_N-\mathbb{E}(S_N)}{V_N}) \to 0$, whence $\frac{1}{\beta-\alpha}\int_\alpha^\beta e^{-t\xi_N}dt \xrightarrow[N\to\infty]{} 1$ for every non-empty interval (α,β). Thus by Theorem 7.8, for every interval $[c,d]\subset[a,b]$, $\lim\limits_{N\to\infty}\dfrac{\mathbb{P}[S_N-z_N\in(c,d)]}{\mathbb{P}[S_N-z_N\in(a,b)]}=\dfrac{|c-d|}{|a-b|}$. (the prefactors $\frac{e^{-V_N\mathcal{I}_N}}{\sqrt{2\pi V_N}}\rho_N$ are identical, and cancel out).

To see part (2), note first that our assumptions on z_N guarantee that $\xi_N=\xi_N\left(\frac{z_N-\mathbb{E}(S_N)}{V_N}\right)$ is bounded away from zero and infinity, and that all its limit points are strictly positive.

Suppose $\xi_{N_k}\to\xi$. Then arguing as in part (1) it is not difficult to see that for all $(a,b)\subset(0,\infty)$ and $r>0$,

$$\lim_{k\to\infty}\frac{\mathbb{P}[\xi_{N_k}(S_{N_k}-z_{N_k})\in(a+r,b+r)|S_{N_k}>z_{N_k}]}{\mathbb{P}[\xi_{N_k}(S_{N_k}-z_{N_k})\in(a,b)|S_{N_k}>z_{N_k}]}=e^{-r}.$$

Since this is true for all convergent $\{\xi_{N_k}\}$, and since any subsequence of $\{\xi_N\}$ has a convergent subsequence,

$$\liminf_{N\to\infty}\frac{\mathbb{P}[\xi_N(S_N-z_N)\in(a+r,b+r)|S_N>z_N]}{\mathbb{P}_x[\xi_N(S_N-z_N)\in(a,b)|S_N>z_N]}=e^{-r},$$
$$\limsup_{N\to\infty}\frac{\mathbb{P}[\xi_N(S_N-z_N)\in(a+r,b+r)|S_N>z_N]}{\mathbb{P}[\xi_N(S_N-z_N)\in(a,b)|S_N>z_N]}=e^{-r},$$

and so $\lim\limits_{N\to\infty}\dfrac{\mathbb{P}[\xi_N(S_N-z_N)\in(a+r,b+r)|S_N>z_N]}{\mathbb{P}[\xi_N(S_N-z_N)\in(a,b)|S_N>z_N]}=e^{-r}$. Thus, conditioned on $S_N>z_N$, $\xi_N(S_N-z_N)$ is asymptotically exponential with parameter 1. □

Corollary 7.11 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Suppose* f *is irreducible, with algebraic range* $\mathbb{Z}$, *and* c_N *are constants such that* $S_N\in c_N+\mathbb{Z}$ *a.s. Suppose* $z_N\in c_N+\mathbb{Z}$. *(1) If* $\frac{z_N-\mathbb{E}(S_N)}{V_N}\to 0$, *then for any* $a<b$ *in* $\mathbb{Z}$, *the distribution of* S_N-z_N *conditioned on* $S_N-z_N\in[a,b]$ *is asymptotically uniform on* $\{a,a+1,\dots,b\}$.

(2) If $\liminf\frac{z_N-\mathbb{E}(S_N)}{V_N}>0$, $\xi_N\left(\frac{z_N-\mathbb{E}(S_N)}{V_N}\right)\to\xi$, *and there exists R such that* $\frac{z_N-\mathbb{E}(S_N)}{V_N}\in[\widehat{a}_N^R,\widehat{b}_N^R]$ *for all sufficiently large N , then* (S_N-z_N) *conditioned on* $S_N\ge z_N$ *is asymptotically geometric with parameter* $e^{-\xi}$.

The proof is similar to the proof in the non-lattice case, so we omit it.

It is worth noting the following consequence of this result, which we state using the point of view of §5.2.3.

Corollary 7.12 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain. Let* z_N *be a sequence s.t. for some R,* $\frac{z_N-\mathbb{E}(S_N)}{V_N}\in[\widehat{a}_N^R,\widehat{b}_N^R]$ *for large N. Let* ζ_N *be the Radon measure* $\zeta_N(\phi)=\mathbb{E}(\phi(S_N-z_N))$ *Let* ζ *be a weak limit of* $\{q_N\zeta_N\}$ *for some sequence* $q_N>0$. *If* f *is irreducible then* ζ *has density* $c_1e^{c_2t}$ *with respect to the Haar measure on the algebraic range of* f *for some* $c_1\in\mathbb{R}_+$, $c_2\in\mathbb{R}$.

If the restriction $\frac{z_N-\mathbb{E}(S_N)}{V_N}\in[\widehat{a}_N^R,\widehat{b}_N^R]$ is dropped, then it is likely that ζ is either as above, or an atomic measure with one atom, but our methods are insufficient for proving this.

7.3 Proofs

In this section we prove Theorems 7.3, 7.4 on the behavior of $\mathcal{F}_n$ and $\mathcal{I}_N$ as $N\to\infty$, and Theorems 7.8 and 7.9 on the LLT for large deviations.

We assume throughout that $\{X_n\}$ is a uniformly elliptic Markov chain with state spaces $\mathfrak{S}_n$ and transition kernels $\pi_{n,n+1}(x,dy)$, and suppose μ_k are the measures $\mu_k(E):=\mathbb{P}(X_k\in E)$. Let $\mathsf{f}=\{f_n\}$ be an a.s. uniformly bounded additive functional on X. Let ϵ_0 denote the ellipticity constant of X, and let $K=\operatorname{ess\,sup}|\mathsf{f}|$.

7.3.1 Strategy of Proof

The proof is an implementation of **Cramér's "change of measure" method**.

We explain the idea. Let z_N be numbers as in Theorem 7.8. We will modify the transition kernels of $\mathsf{X} = \{X_n\}$ to generate a Markov array $\widetilde{\mathsf{X}} = \{\widetilde{X}_n^{(N)}\}$ whose row sums $\widetilde{S}_N = f_1(\widetilde{X}_1^{(N)}, \widetilde{X}_2^{(N)}) + \cdots + f_N(\widetilde{X}_N^{(N)}, \widetilde{X}_{N+1}^{(N)})$ satisfy

$$|z_N - \mathbb{E}(\widetilde{S}_N)| \leq \text{const.} \tag{7.2}$$

(7.2) places us in the regime of local deviations, which was analyzed in Chapter 5. The results of that chapter provide asymptotics for $\mathbb{P}(\widetilde{S}_N - z_N \in (a,b))$, and these can be translated into asymptotics for $\mathbb{P}(S_N - z_N \in (a,b))$.

The array $\widetilde{\mathsf{X}}$ will have state spaces $\mathfrak{S}_n^{(N)} := \mathfrak{S}_n$, row lengths $N+1$, initial distributions $\pi^{(N)}(E) := \mathbb{P}[X_1 \in E]$, and transition probabilities

$$\widetilde{\pi}_{n,n+1}^{(N)}(x, \mathrm{d}y) := \mathrm{e}^{\xi_N f_n(x,y)} \frac{h_{n+1}(y, \xi_N)}{\mathrm{e}^{p_n(\xi_N)} h_n(x, \xi_N)} \cdot \pi_{n,n+1}(x, \mathrm{d}y), \tag{7.3}$$

where the real parameters ξ_N are calibrated to get (7.2), and the positive functions h_n, h_{n+1} and the real numbers p_n are chosen to guarantee that $\int \widetilde{\pi}_{n,n+1}^{(N)}(x, \mathrm{d}y) = 1$.

The value of ξ_N will depend on $\frac{z_N - \mathbb{E}(S_N)}{V_N}$. To construct ξ_N and to control it, we must know that $\frac{z_N}{V_N}$ belong to a sets where $\mathcal{F}_N$ are strictly convex, uniformly in N. This is the reason why we need to assume that $\exists R$ s.t. $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ for all N, a condition we can check as soon as $|\frac{z_N - \mathbb{E}(S_N)}{V_N}| < c$ with c small enough.

We remark that the dependence of ξ_N on N means that $\{\widetilde{X}_n^{(N)}\}$ is an array, not a chain. The fact that the change of measure produces arrays from chains is the main reason we insisted on working with arrays in the first part of this work.

7.3.2 A Parameterized Family of Changes of Measure

Let ξ_N be *arbitrary* bounded real numbers. In this section we construct functions $h_k^{\xi_N}(\cdot) = h_k(\cdot, \xi_N)$ and $p_n(\xi_N) \in \mathbb{R}$ so that the measures $\widetilde{\pi}_{n,n+1}^{(N)}(x, \mathrm{d}y)$ in (7.3) are probability measures. In the next section we will choose specific $\{\xi_N\}$ to get (7.2).

Lemma 7.13 *Given $\xi \in \mathbb{R}$ and a sequence of real numbers $\{a_n\}_{n\in\mathbb{N}}$, there are unique numbers $p_n(\xi) \in \mathbb{R}$, and unique positive $h_n(\cdot, \xi) \in L^\infty(\mathfrak{S}_n, \mathscr{B}(\mathfrak{S}_n), \mu_n)$ s.t. $\int_{\mathfrak{S}_n} h_n(x, \xi)\mu_n(\mathrm{d}x) = \exp(a_n \xi)$ for all n, and for a.e. x*

$$\int_{\mathfrak{S}_{n+1}} \mathrm{e}^{\xi f_n(x,y)} \frac{h_{n+1}(y, \xi)}{\mathrm{e}^{p_n(\xi)} h_n(x, \xi)} \pi_{n,n+1}(x, \mathrm{d}y) = 1. \tag{7.4}$$

We will sometimes write $h_n^\xi(\cdot) := h_n(\cdot, \xi)$.

Remark. If $\{\overline{h}_n(\cdot, \xi)\}$, $\{\overline{p}_n(\xi)\}$ satisfy the Lemma with $a_n = 0$, then the unique solution with general $\{a_n\}$ is given by

$$h_n(\cdot, \xi) := \mathrm{e}^{a_n \xi} \overline{h}_n(\cdot, \xi)\ ,\ p_n(\xi) := \overline{p}_n(\xi) - a_n \xi + a_{n+1} \xi. \tag{7.5}$$

Evidently, (h_n, p_n) and $(\overline{h}_n, \overline{p}_n)$ give rise to the same transition probabilities (7.3). We call $\{\overline{h}_n\}$ and $\{\overline{p}_n\}$ the **fundamental solution**.

Proof It is enough to prove the existence and uniqueness of the fundamental solution, so henceforth we assume $a_n = 0$. We may also assume without loss of generality that $|\xi| \leq 1$; otherwise we scale f.

Set $V_n := L^\infty(\mathfrak{S}_n, \mathscr{B}(\mathfrak{S}_n), \mu_n)$, and define the operators $L_n^\xi : V_{n+1} \to V_n$ by

$$(L_n^\xi h)(x) = \int\limits_{\mathfrak{S}_{n+1}} \mathrm{e}^{\xi f_n(x,y)} h(y)\pi_{n,n+1}(x, \mathrm{d}y). \tag{7.6}$$

L_n^ξ are linear, bounded, and positive.

For (7.4) to hold, it is necessary and sufficient that $h_n^\xi(\cdot) := h_n(\cdot,\xi)$ be positive a.e., and $L_n^\xi h_{n+1}^\xi = \mathrm{e}^{p_n(\xi)} h_n^\xi$. Positivity may be replaced by the weaker property that $h_n^\xi \in L^\infty \setminus \{0\}$ and $h_n^\xi \geq 0$. For such functions, since $|\mathsf{f}| \leq K$ a.s. and X is uniformly elliptic, $h_n^\xi(x) = \mathrm{e}^{-p_n(\xi)-p_{n+1}(\xi)}(L_n^\xi L_{n+1}^\xi h_{n+2}^\xi)(x) \geq \mathrm{e}^{-p_n(\xi)-p_{n+1}(\xi)-2K}\epsilon_0 \|h_{n+2}^\xi\|_1$. So to prove the lemma it is enough to find $p_n(\xi) \in \mathbb{R}$ and non-negative $h_n^\xi \in L^\infty \setminus \{0\}$ such that $L_n^\xi h_{n+1}^\xi = \mathrm{e}^{p_n(\xi)} h_n^\xi$.

The existence and uniqueness of such "generalized eigenvectors" can be proved using **Hilbert's projective metrics**. We recall, briefly, what these are, and refer the reader to Appendix B for more details.

Let $C_n := \{h \in V_n : h \geq 0 \text{ a.e. }\}$. These are closed cones and $L_n^\xi(C_{n+1}) \subset C_n$. Given h, g in the interior of C_n, let $M = M(h|g)$ and $m = m(h|g)$ denote the best constants in the double a.e. inequality $mg \leq h \leq Mg$, and set

$$d_n(h, g) := \log\left(\frac{M(h|g)}{m(h|g)}\right) \in [0, \infty], \quad (h, g \in C_n).$$

This is a pseudo-metric on the interior of C_n, and $d(h, g) = 0 \Leftrightarrow h, g$ are proportional. Also, for all h, g in the interior of C_n,

$$\left\| \frac{h}{\int h} - \frac{g}{\int g} \right\|_1 \leq \mathrm{e}^{d_n(h,g)} - 1. \tag{7.7}$$

Denote $T_n^\xi := L_n^\xi L_{n+1}^\xi : C_{n+2} \to C_n$. By the uniform ellipticity assumption and the bounds ess sup $|\mathsf{f}| < K$ and $|\xi| \leq 1$,

$$\mathrm{e}^{-2K}\epsilon_0 \|h\|_1 \leq (T_n^\xi h)(x) \leq \mathrm{e}^{2K}\epsilon_0^{-2}\|h\|_1 \quad (h \in C_{n+2}). \tag{7.8}$$

So $d_n(T_n^\xi h, 1) \leq 4K + 3\log(1/\epsilon_0)$, and the diameter of $T_n^\xi(C_{n+2})$ in C_n is less than $\Delta := 8K + 6\log(1/\epsilon_0)$. By Birkhoff's theorem (Theorem B.6), every linear map $T : C_{n+2} \to C_n$ such that the d_n–diameter of $T(C_{n+2})$ in C_n is less than Δ, contracts d_n at least by a factor $\theta := \tanh(\Delta/4) = \tanh(2K + \frac{3}{2}\log(1/\epsilon_0)) \in (0, 1)$. Hence

$$d_n(T_{n+1}^\xi h, T_{n+1}^\xi g) \leq \theta d_{n+2}(h, g) \quad (h, g \in C_{n+2}). \tag{7.9}$$

Since $\theta \in (0, 1)$, $\{L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}\}_{k\geq 1}$ is a Cauchy sequence in C_n, with respect to d_n. By (7.7), $\dfrac{L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}}{\|L_n^\xi L_{n+1}^\xi \cdots L_{n+k-1}^\xi 1_{\mathfrak{S}_{n+k}}\|_1}$ $(k \geq 1)$ is a Cauchy sequence in $L^1(\mathfrak{S}_n)$. Call the limiting function h_n^ξ. Clearly h_n^ξ has integral one, and h_n^ξ is positive and bounded because of (7.8). It is also clear that $L_n^\xi h_{n+1}^\xi = \mathrm{e}^{p_n} h_n^\xi$ for some $p_n \in \mathbb{R}$. So $\{h_n^\xi\}, \{p_n\}$ exist.

The proof shows that the d_n-diameter of $\bigcap_{k\geq 1} L_n^\xi \cdots L_{n+k-1}^\xi(C_{n+k})$ is zero. So h_n^ξ is unique up to multiplicative constant. Since $\int h_n^\xi = 1$, it is unique. □

The proof has a useful consequence: For every $R > 0$, there exists $C_0 > 0$ and $\theta \in (0, 1)$ (depending on R) such that for every $|\xi| \leq R$

$$d_1\left(L_1^\xi \cdots L_N^\xi h_{N+1}^\xi,\ L_1^\xi \cdots L_N^\xi 1\right) \leq C_0 \theta^{N/2} d_{N+1}\left(h_{N+1}^\xi, 1\right). \tag{7.10}$$

The case when N is even follows from (7.9). The case when N is odd is obtained from the even case, by using the exponential contraction of $L_2^\xi \cdots L_N^\xi$, and Proposition B.5, that says that (7.9) holds even when $\Delta = \infty$ and $\theta = 1$.

Lemma 7.14 *Let $h(\cdot,\xi)$ be as in Lemma 7.13. If a_n is bounded, then for every $R > 0$ there is $C = C(R, \sup a_n)$ such that for all $n \geq 1$, a.e. $x \in \mathfrak{S}_n$ and $|\xi| \leq R$,*

$$h_n(x, 0) = 1\,,\ p_n(0) = 0\,,\ C^{-1} \leq h_n(x, \xi) \leq C \quad \textit{and} \quad C^{-1} < \mathrm{e}^{p_n(\xi)} < C.$$

Proof It is enough to consider the fundamental solution ($a_n = 0$); the general case follows from (7.5). So henceforth assume that $\int h_n \mathrm{d}\mu_n = 1$. It is also sufficient to consider the case $|\xi| \le 1$; otherwise, we scale f.

The first two statements ($h_n(\cdot, 0) \equiv 1$ and $p_n(0) = 0$) are because $h_n \equiv 1$ solves (7.4) when $\xi = 0$, and the solution is unique.

Let $\{h_n^\xi\}$ be the fundamental solution, then in the notation of the previous proof, $T_n^\xi h_{n+2}^\xi = \mathrm{e}^{p_n(\xi)+p_{n+1}(\xi)} h_n^\xi$. By (7.8), $\mathrm{e}^{-2K}\epsilon_0 \le \mathrm{e}^{p_n(\xi)+p_{n+1}(\xi)} h_n^\xi \le \mathrm{e}^{2K}\epsilon_0^{-2}$. Integrating, we obtain $\mathrm{e}^{-2K}\epsilon_0 \le \mathrm{e}^{p_n(\xi)+p_{n+1}(\xi)} \le \mathrm{e}^{2K}\epsilon_0^{-2}$. So,

$$\mathrm{e}^{-4K}\epsilon_0^3 \le h_n^\xi(\cdot) \le \mathrm{e}^{4K}\epsilon_0^{-3}. \tag{7.11}$$

Next, $\mathrm{e}^{p_n(\xi)} = \int \mathrm{e}^{p_n} h_n^\xi \mathrm{d}\mu_n = \int L_n^\xi h_{n+1}^\xi \mathrm{d}\mu_n = \mathrm{e}^{\pm K} \iint h_{n+1}^\xi \pi_{n,n+1}(x, \mathrm{d}y)\mu_n(\mathrm{d}y)$. By (7.11), $\mathrm{e}^{p_n(\xi)} = \mathrm{e}^{\pm(5K+3|\log\epsilon_0|)}$. □

In the next section we will choose ξ_N to guarantee (7.2). As it turns out, the choice involves a condition on $p_n'(\xi)$. Later, we will also need to use $p_n''(\xi)$. In preparation for this, we will now analyze the differentiability of $\xi \mapsto h_n^\xi$ and $\xi \mapsto p_n(\xi)$.

The map $\xi \mapsto h_n^\xi$ takes values in the Banach space L^∞, and we will need the machinery of **real-analytic maps into Banach spaces** [49]. Here is a brief review. Suppose $\mathfrak{X}, \mathfrak{Y}$ are Banach spaces. Let $a_n : \mathfrak{X}^n \to \mathfrak{Y}$ be a multilinear map. We set $\|a_n\| := \sup\{\|a_n(x_1, \dots, x_n)\| : x_i \in \mathfrak{X},\ \|x_i\| \le 1 \text{ for all } i\}$. A multilinear map is called **symmetric** if it is invariant under the permutation of its coordinates. Given $x \in \mathfrak{X}$, we denote $a_n x^n := a_n(x, \dots, x)$. A **power series** is a formal expression $\sum_{n\ge1} a_n x^n$ where $a_n : \mathfrak{X}^n \to \mathfrak{Y}$ are multilinear and symmetric. A function $\phi : \mathfrak{X} \to \mathfrak{Y}$ is called **real-analytic** at x_0, if there is some $r > 0$ and a power series $\sum a_n x^n$, called the **Taylor series at** x_0, such that $\sum \|a_n\| r^n < \infty$ and $\phi(x) = \phi(x_0) + \sum_{n\ge1} a_n(x - x_0)^n$ whenever $\|x - x_0\| < r$. One can check that if this happens, then:

$$a_n(x_1, \dots, x_n) = \frac{1}{n!}\frac{\mathrm{d}}{\mathrm{d}t_1}\Big|_{t_1=0} \cdots \frac{\mathrm{d}}{\mathrm{d}t_n}\Big|_{t_n=0} \phi\Big(x_0 + \sum_{i=1}^n t_i x_i\Big). \tag{7.12}$$

Conversely, one can show that if $\sum a_n(x - x_0)^n$ has a positive radius of convergence with a_n as in (7.12), then ϕ is real-analytic, and equal to its Taylor series at x_0, on a neighborhood of x_0.

Example 7.15 Let $\phi : \mathfrak{X}\times\mathfrak{X}\times\mathbb{R} \to \mathfrak{X}$ be the map $\phi(x, y, z) := x - y/z$. Then ϕ is real-analytic at every (x_0, y_0, z_0) such that $z_0 \ne 0$, and the coefficients of its Taylor series $\phi(x, y, z) = \phi(x_0, y_0, z_0) + \sum_{n=1}^\infty a_n(x - x_0, y - y_0, z - z_0)^n$ satisfy $\|a_n\| = O(\|y_0\|/|z_0|^{n+1}) + O(n/|z_0|^{n+1})$.

Proof. If $|z - z_0| < |z_0|$, then $x - yz^{-1} = x - yz_0^{-1}\sum_{k\ge0}(-1)^k z_0^{-k}(z - z_0)^k$. Let $\underline{x}_i := (x_i, y_i, z_i)$ and $(t_1, \dots, t_n) \in \mathbb{R}^n$. The series

$$\phi\Big(\underline{x}_0 + \sum_{i=1}^n t_i \underline{x}_i\Big) = x_0 + \sum_{i=1}^n t_i x_i + \sum_{k=0}^\infty \frac{(-1)^{k+1}}{z_0^{k+1}}\Big(y_0 + \sum_{i=1}^n t_i y_i\Big)\Big(\sum_{i=1}^n t_i z_i\Big)^k \tag{7.13}$$

converges in norm whenever $(t_1, \dots, t_n) \in A_n := [|\sum_{i=1}^n t_i z_i| < |z_0|]$.

In particular, on A_n, this series is real-analytic separately in each t_i, and can be differentiated term-by-term infinitely many times.

To find $a_n(\underline{x}_1, \dots, \underline{x}_n)$ we observe that the differential (7.12) is equal to the coefficient of $t_1 \cdots t_n$ on the right-hand-side of (7.13). So for $n > 2$,

$$a_n(\underline{x}_1, \dots, \underline{x}_n) = (-1)^{n+1} y_0 z_0^{-(n+1)} \cdot z_1 \cdots z_n + (-1)^n z_0^{-n} \sum_{i=1}^n y_i z_1 \cdots \widehat{z_i} \cdots z_n$$

where the hat above z_i indicates that the i-th term should be omitted. It follows that $\|a_n\| = O(\|y_0\|/|z_0|^{n+1}) + O(n/|z_0|^n)$. □

Lemma 7.16 *The functions $\xi \mapsto h_n^\xi, p_n(\xi)$ are real-analytic. If a_n is bounded, then for every $R > 0$ there is $C(R, \sup a_n) > 0$ such that for every $|\xi| \le R$ and $n \ge 1$,*

$$\left\| \frac{\partial}{\partial \xi} h_n(\cdot, \xi) \right\|_\infty \le C(R), \quad \left\| \frac{\partial^2}{\partial \xi^2} h_n(\cdot, \xi) \right\|_\infty \le C(R).$$

Proof Without loss of generality, $R = 1$ and h_n^ξ are the fundamental solutions ($a_n = 0$), i.e. $\int h_n^\xi d\mu_n = 1$. Fix $|\xi| \le 1$ and let $T_n := T_n^\xi$, $h_n(\cdot) = h_n(\cdot, \xi)$ be as in the proof of Lemma 7.13. Define two Banach spaces:

$$X := \left\{ (S_n)_{n\in\mathbb{N}} : \begin{array}{l} S_n : L^\infty(\mathfrak{S}_{n+2}) \to L^\infty(\mathfrak{S}_n) \text{ are bounded linear operators,} \\ \text{and } \|S\| := \sup_n \|S_n\| < \infty \end{array} \right\}$$
$$Y := \{(\varphi_n)_{n\in\mathbb{N}} : \varphi_n \in L^\infty(\mathfrak{S}_n),\ \|\varphi\| := \sup \|\varphi_n\|_\infty < \infty\}.$$

By (7.8), $T := (T_n)$ belongs to X. By Lemma 7.14, $h := (h_n)_{n\in\mathbb{N}}$ belongs to Y.

Step 1. *There exists some $0 < \delta < 1$ such that for every $(S, \varphi) \in X \times Y$, for all $|\xi| \le 1$, if $\|S - T\| < \delta$ and $\|\varphi - h\| < \delta$, then* $\inf |\int (S_n\varphi_{n+2})| > \delta$.

Proof of the Step. Let $C := \sup_{|\xi|\le 1} \sup_n \|h_n^\xi\|_\infty$. By (7.8), $\|T_n\| \le M$ where $M := e^{2K}\epsilon_0^{-2}$, and by Lemma 7.14, there is a constant $\varepsilon_1 > 0$ so that for all n and $|\xi| \le 1$, $\varepsilon_1 \le (T_n h_{n+2})(x) \le \varepsilon_1^{-1}$.

So if $\|S - T\| < \delta$ and $\|\varphi - h\| < \delta$, then for a.e. x,

$$\begin{aligned} S_n\varphi_{n+2}(x) &= (T_n h_{n+2})(x) - (T_n - S_n)h_{n+2}(x) - S_n(h_{n+2} - \varphi_{n+2})(x) \\ &\ge \varepsilon_1 - \|T - S\|\|h\| - (\|S - T\| + \|T\|)\|h - \varphi\| \ge \varepsilon_1 - C\delta - (\delta + M)\delta. \end{aligned}$$

If δ is small enough, then $S_n\varphi_{n+2} > \delta$ a.e. for all n, and the step follows.

Henceforth we fix δ as in Step 1. Let $B_\delta(T) := \{S \in X : \|S - T\| < \delta\}$ and $B_\delta(h) := \{\varphi \in Y : \|\varphi - h\| < \delta\}$. Define

$$\Upsilon : B_\delta(T) \times B_\delta(h) \to Y, \quad \Upsilon(S, \varphi) := \left(\varphi_n - \frac{S_n\varphi_{n+2}}{\int (S_n\varphi_{n+2}) d\mu_n} \right)_{n\in\mathbb{N}}.$$

This is well-defined by the choice of δ, and $\Upsilon(T, h) = 0$.

Step 2. *Υ is real-analytic on $B_\delta(T) \times B_\delta(h)$.*

Proof of the Step. $\Upsilon = \Phi(\Upsilon^{(1)}, \Upsilon^{(2)}, \Upsilon^{(3)})$, with

- $\Phi((\varphi, \psi, \xi)_{i\ge 1}) = (\varphi_i - \xi_i^{-1}\psi_i)_{i\ge 1}$;
- $\Upsilon^{(1)} : X \times Y \to Y$, $\Upsilon^{(1)}(S, \varphi) = \varphi$;
- $\Upsilon^{(2)} : X \times Y \to Y$, $\Upsilon^{(2)}(S, \varphi) = (S_n\varphi_{n+2})_{n\in\mathbb{N}}$;
- $\Upsilon^{(3)} : X \times Y \to \ell^\infty$, $\Upsilon^{(3)}(S, \varphi) = (\int (S_n\varphi_{n+2}) d\mu_n)_{n\in\mathbb{N}}$.

We claim that for each i, some high-order derivative of $\Upsilon^{(i)}$ is identically zero. Let D be the total derivative, and let D_i be the partial derivative with respect to the i-th variable, then:

- $D^2\Upsilon^{(1)} = 0$, because $\Upsilon^{(1)}$ is linear, so its first total derivative is constant.
- $D^3\Upsilon^{(2)} = 0$: Starting with the identity $\Upsilon^{(2)}(S, \varphi) = (S_n\varphi_{n+2})_{n\in\mathbb{N}}$, we find by repeated differentiation that

$$\begin{array}{ll} (D_1\Upsilon^{(2)})(S, \varphi)(S') = (S'_n\varphi_{n+2})_{n\in\mathbb{Z}} \quad , & (D_1^2\Upsilon^{(2)})(S, \varphi) = 0 \\ (D_2\Upsilon^{(2)})(S, \varphi)(\varphi') = (S_n\varphi'_{n+2})_{n\in\mathbb{Z}} \quad , & (D_2^2\Upsilon^{(2)})(S, \varphi) = 0 \\ (D_1D_2\Upsilon^{(2)})(S, \varphi)(S', \varphi') = (S'_n\varphi'_{n+2})_{n\in\mathbb{Z}}. & \end{array}$$

  We see that all second-order partial derivatives of $\Upsilon^{(2)}$ at (S, φ) do not depend on (S, φ). Therefore $D^2\Upsilon^{(2)}$ is constant, and $D^3\Upsilon^{(2)} = 0$.

- $D^3\Upsilon^{(3)} = 0$, because $\Upsilon^{(3)} = L \circ \Upsilon^{(2)}$ where L is a bounded linear map, so $D^3(L \circ \Upsilon^{(2)}) = LD^3\Upsilon^{(2)} = 0$.

Consequently, $\Upsilon^{(i)}$ are real-analytic on their domains (with finite Taylor series).

By Step 1, $\overrightarrow{\Upsilon} := (\Upsilon^{(1)}, \Upsilon^{(2)}, \Upsilon^{(3)})$ maps $B_\delta(T) \times B_\delta(h)$ into

$$U := \{(\varphi, \psi, \xi) \in Y \times Y \times \ell^\infty : \|\varphi\| < C + \delta, \|\psi\| < (C+\delta)(M+\delta),\ \inf |\xi_i| > \delta/2\}.$$

We will show that Φ is real-analytic on U.

By Example 7.15, $x - \frac{y}{z} = \sum_{n=0}^{\infty} a_n(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0)^n$, where $a_n(x_0, y_0, z_0) : (\mathbb{R}^3)^n \to \mathbb{R}$ are symmetric multilinear functions depending on (x_0, y_0, z_0), and satisfying $\|a_n(x_0, y_0, z_0)\| = O(|y_0|/|z_0|^{n+1}) + O(n/|z_0|^{n+1})$. So

$$\Phi(\varphi, \psi, \xi) = \Phi(\varphi^{(0)}, \psi^{(0)}, \xi^{(0)}) + \sum_{n=1}^{\infty} A_n(\varphi - \varphi^{(0)}, \psi - \psi^{(0)}, \xi - \xi^{(0)})^n, \tag{7.14}$$

where $A_n((\varphi^{(1)}, \psi^{(1)}, \xi^{(1)}), \ldots, (\varphi^{(n)}, \psi^{(n)}, \xi^{(n)})) \in Y$ has i-th entry

$$a_n(\varphi_i^{(0)}(x), \psi_i^{(0)}(x), \xi_i^{(0)})\big((\varphi_i^{(1)}(x), \psi_i^{(1)}(x), \xi_i^{(1)}), \ldots, (\varphi_i^{(n)}(x), \psi_i^{(n)}(x), \xi_i^{(n)})\big).$$

A_n inherits multilinearity and symmetry from a_n, and by construction,

$$\|A_n\| \le \sup\left\{\|a_n(x_0, y_0, z_0)\| : |x_0|, |y_0| \le (C+\delta)(M+1+\delta), |z_0| > \tfrac{\delta}{2}\right\} = O(\tfrac{2^n n}{\delta^n}).$$

So the right-hand-side of (7.14) has positive radius of convergence, proving the analyticity of $\Phi : U \to Y$. The step follows from the well-known result that the composition of real-analytic functions is real-analytic, see [49].

STEP 3. $(D_2\Upsilon)(T, h) : Y \to Y$ *has a bounded inverse.*

Proof. A direct calculation gives $(D_2\Upsilon)(T, h)(\varphi) = \varphi - \Lambda\varphi$, where $(\Lambda\varphi)_n = \frac{T_n\varphi_{n+2}}{\int (T_n h_{n+2}) \mathrm{d}\mu_n} - \left(\frac{\int (T_n\varphi_{n+2})\mathrm{d}\mu_n}{\int (T_n h_{n+2})\mathrm{d}\mu_n}\right) h_n$. It is sufficient to show that Λ has spectral radius strictly smaller than one.

Let $T_n^{(k)} := T_n T_{n+2} \cdots T_{n+2(k-1)}$, then we claim that

$$(\Lambda^k\varphi)_n = \frac{T_n^{(k)}\varphi_{n+2k}}{\int (T_n^{(k)} h_{n+2k})\mathrm{d}\mu_n} - \left(\frac{\int (T_n^{(k)}\varphi_{n+2k})\mathrm{d}\mu_n}{\int (T_n^{(k)} h_{n+2k})\mathrm{d}\mu_n}\right) h_n. \tag{7.15}$$

To see this, we use $T_m h_{m+2} \propto h_m$ and $\int h_m \mathrm{d}\mu_m = 1$ to note that

$$\int (T_n^{(k+1)} h_{n+2(k+1)})\mathrm{d}\mu_n = \int (T_n h_{n+2})\mathrm{d}\mu_n \int (T_{n+2}^{(k)} h_{n+2(k+1)})\mathrm{d}\mu_{n+2}.$$

With this identity in mind, the formula for Λ^k follows by induction.

We now explain why (7.15) implies that the spectral radius of Λ is less than one.

Fix $\varphi \in Y$. Recall that $C^{-1} \le h_n \le C$ for all n, and let $\psi := \varphi + 2C\|\varphi\| h$. Then $\psi \in Y$, $\Lambda^k\psi = \Lambda^k\varphi$ for all k (because $\Lambda h = 0$), and for all n

$$C\|\varphi\| h_n \le \psi_n \le 3C\|\varphi\| h_n. \tag{7.16}$$

In particular, if C_n is the cone from the proof of Lemma 7.13, and d_n is its projective Hilbert metric, then $\psi_n \in C_n$ and $d_n(\psi_n, h_n) \le \log 3$. Since T_n contracts the Hilbert projective norm by a factor $\theta \in (0, 1)$, $d_n(T_n^{(k)}\psi_{n+2k}, T_n^{(k)} h_{n+2k}) \le \theta^k \log 3$. This implies by the definition of d_n that for a.e. $x \in \mathfrak{S}_n$,

$$\left|\frac{(T_n^{(k)}\psi_{n+2k})(x)/\int (T_n^{(k)}\psi_{n+2k})}{(T_n^{(k)} h_{n+2k})(x)/\int (T_n^{(k)} h_{n+2k})} - 1\right| \le \max\{3^{\theta^k} - 1, 1 - 3^{-\theta^k}\} =: \varepsilon_k.$$

The denominator simplifies to h_n. So

$$\left\| \frac{(T_n^{(k)}\psi_{n+2k})}{\int (T_n^{(k)}\psi_{n+2k})} - h_n \right\|_\infty \le \varepsilon_k \|h\|. \tag{7.17}$$

By (7.16), $C\|\varphi\| T_n^{(k)} h_{n+2k} \le T_n^{(k)}\psi_{n+2k} \le 3C\|\varphi\| T_n^{(k)} h_{n+2k}$; therefore

$$C\|\varphi\| \le \frac{\int (T_n^{(k)}\psi_{n+2k})}{\int (T_n^{(k)} h_{n+2k})} \le 3C\|\varphi\|. \tag{7.18}$$

By (7.15), (7.17) and (7.18),

$$\|\Lambda^k\psi\| = \sup_n \left\| \frac{T_n^{(k)}\psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} - \frac{\int T_n^{(k)}\psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} \cdot h_n \right\|_\infty \le \sup_n \left\| \frac{T_n^{(k)}\psi_{n+2k}}{\int T_n^{(k)}\psi_{n+2k}} - h_n \right\|_\infty \sup_n \left\| \frac{\int T_n^{(k)}\psi_{n+2k}}{\int T_n^{(k)} h_{n+2k}} \right\|_\infty \le 3C\varepsilon_k \|h\|\, \|\varphi\|.$$

In summary, $\|\Lambda^k\varphi\| = O(\varepsilon_k\|\varphi\|)$. Since $\lim \sqrt[k]{\varepsilon_k} = \theta < 1$, the spectral radius of Λ is less than 1. Therefore $D_2\Upsilon \equiv \mathrm{Id} - \Lambda$ has a bounded inverse.

We can now complete the proof of the lemma. We constructed a real-analytic function $\Upsilon : X \times Y \to Y$ such that $\Upsilon(T, h) = 0$ and $(D_2\Upsilon)(T, h) : Y \to Y$ has a bounded inverse. By the implicit function theorem for real-analytic functions on Banach spaces [199], T has a neighborhood $W \subset X$ where one can define a real-analytic function $h : W \to Y$ so that $\Upsilon(S, h(S)) = 0$.

Recall that $T = T^\xi := \{T_n^\xi\}_{n\in\mathbb{N}}$ and $h = \{h_n(\cdot,\xi)\}_{n\ge1}$. It is easy to see using ess sup $|\mathrm{f}| < \infty$ that $\eta \mapsto T^\eta$ is real-analytic (even holomorphic). Therefore the composition $\eta \mapsto h(T^\eta)$ is real-analytic.

We will now show that $h(T^\eta) = h^\eta = \{h_n^\eta\}_{n\ge1}$, for all $|\eta|$ small enough, and deduce real-analyticity with uniform derivative bounds for $\eta \mapsto h_n^\eta(x)$. Since $h_n(\cdot,\xi)$ are uniformly bounded below by a positive constant, we can choose W so that the functions $h(S)_n$ are uniformly bounded from below for $S \in W$ and $n \in \mathbb{N}$. In particular, for all η close enough to ξ, $T^\eta \in W$, and $h(T^\eta)_n$ are all positive. Next, by construction, $\Upsilon(T^\eta, h(T^\eta)) = 0$, and this implies that $T_n^\eta h(T^\eta)_{n+2} \propto h(T^\eta)_n$ and $\int h(T^\eta)_n \mathrm{d}\mu_n = 1$ for all n.

Lemma 7.13 and its proof show that there can be at most one sequence of functions like that. So $h(T^\eta) = \{h_n(\cdot,\eta)\}_{n\ge1}$ for all η sufficiently close to ξ.

Thus $\eta \mapsto \{h_n^\eta\}_{n\ge1}$ is real-analytic, as a function taking values in Y. It follows that $\eta \mapsto h_n(\cdot,\eta)$ is real-analytic for all n, and $\left\{\frac{\partial^j}{\partial\eta^j} h_n(\cdot,\eta)\right\}_{n\ge1} = \frac{\partial^j}{\partial\eta^j} h(T^\eta) \in Y$ for all j. By the definition of Y,
$\sup_{n\ge1} \left\| \frac{\partial^j}{\partial\eta^j}\Big|_{\eta=\xi} h_n(\cdot,\eta) \right\|_\infty = \left\| \frac{\partial^j}{\partial\eta^j}\Big|_{\eta=\xi} h(T^\eta) \right\| < \infty$ for all $j \in \mathbb{N}$. Since $\eta \mapsto h(T^\eta)$ is real-analytic on a neighborhood of ξ, this norm is uniformly bounded on compact sets of ξ. The lemma is proved. □

7.3.3 Choosing the Parameters

Fix ξ and $\{a_n\}$, and construct $p_n(\xi)$ and $h_k(\cdot,\xi)$ as in Lemma 7.13. Let $\widetilde{\mathsf{X}}^\xi$ denote the Markov chain with the initial distribution and state spaces of X, but with transition probabilities $\widetilde{\pi}^\xi_{n,n+1}(x, \mathrm{d}y) = \mathrm{e}^{\xi f_n(x,y)} \frac{h_{n+1}(y,\xi)}{\mathrm{e}^{p_n(\xi)} h_n(x,\xi)} \cdot \pi_{n,n+1}(x, \mathrm{d}y)$.

Denote the expectation and variance operators of this chain by $\widetilde{\mathbb{E}}^\xi, \widetilde{V}^\xi$. We now show that if $V_N := \mathrm{Var}(S_N) \to \infty$ and $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ is sufficiently small, then it is possible to choose ξ_N and a_n bounded such that

$$\widetilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)\,, \quad \frac{z_N - \widetilde{\mathbb{E}}^{\xi_N}(S_N)}{\sqrt{\widetilde{V}^{\xi_N}(S_N)}} \xrightarrow[N\to\infty]{} 0 \text{ and } \sum_{n=1}^N p_n'(0) = \mathbb{E}(S_N).$$

The construction will show that if $\dfrac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\xi_N \to 0$.

Recall that while the choice of $\{a_n\}$ affects $h_n^\xi := h_n(\cdot,\xi)$, it does not affect $\widetilde{\pi}^\xi_{n,n+1}$. Specifically, if $\overline{h}_n^\xi := \overline{h}_n(\cdot,\xi)$ and $\overline{p}_n(\xi) \in \mathbb{R}$ is the fundamental solution, then $\overline{h}_n^\xi = \mathrm{e}^{-a_n\xi} h_n(\cdot,\xi)$ and $\overline{p}_n(\xi) = p_n(\xi) + a_n\xi - a_{n+1}\xi$, so $\widetilde{\pi}^\xi_{n,n+1}(x,\mathrm{d}y) = \mathrm{e}^{\xi f_n(x,y)} \dfrac{\overline{h}_{n+1}(x,\xi)}{\mathrm{e}^{\overline{p}_n(\xi)}\overline{h}_n(y,\xi)} \pi_{n,n+1}(x,\mathrm{d}y)$. Let $\overline{P}_N(\xi) := \overline{p}_1(\xi) + \cdots + \overline{p}_N(\xi)$.

Lemma 7.17 *$\xi \mapsto \overline{P}_N(\xi)$ is real-analytic, and for every $R > 0$ there is a constant $C(R)$ such that for all $|\xi| \le R$ and $N \in \mathbb{N}$,*

(1) $|\overline{P}'_N(\xi) - \widetilde{\mathbb{E}}^\xi(S_N)| \le C(R)$;
(2) Suppose $V_N \to \infty$. Then $C(R)^{-1} \le \widetilde{V}^\xi(S_N)/V_N \le C(R)$ for all N and $|\xi| \le R$.
Moreover $\overline{P}''_N(\xi)/\widetilde{V}^\xi(S_N) \xrightarrow[N\to\infty]{} 1$ uniformly in $|\xi| \le R$.

Proof We have the identity $\mathrm{e}^{\overline{P}_N(\xi)} = \int (L_1^\xi \cdots L_N^\xi \overline{h}^\xi_{N+1})(x)\mu_1(\mathrm{d}x)$. Since $\xi \mapsto \overline{h}^\xi$ and $\xi \mapsto L_n^\xi$ are real-analytic, $\xi \mapsto \overline{P}_N(\xi)$ is real-analytic.

Given $x \in \mathfrak{S}_1$ (the state space of X_1), define two measures on $\prod_{i=2}^{N+1} \mathfrak{S}_i$ so that for every $E_i \in \mathscr{B}(\mathfrak{S}_i)$

$$\pi_x(E_2 \times \cdots \times E_{N+1}) := \mathbb{P}(X_2 \in E_2, \ldots, X_{N+1} \in E_{N+1} | X_1 = x),$$
$$\widetilde{\pi}_x^\xi(E_2 \times \cdots \times E_{N+1}) := \widetilde{\mathbb{P}}^\xi(\widetilde{X}_2^\xi \in E_2, \ldots, \widetilde{X}_{N+1}^\xi \in E_{N+1} | \widetilde{X}_1^\xi = x).$$

Let $S_N(x, y_2, \ldots, y_{N+1}) := f_1(x,y_1) + \sum_{i=2}^{N} f_i(y_i, y_{i+1})$, then $\dfrac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}(y_2,\ldots,y_{N+1}) = \mathrm{e}^{\xi S_N(x,\underline{y})}\mathrm{e}^{-\overline{P}_N(\xi)}\left(\dfrac{\overline{h}_{N+1}(y_{N+1},\xi)}{\overline{h}_1(x,\xi)}\right)$.

By Lemma 7.16, $\xi \mapsto \frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}(y_2,\ldots,y_{N+1})$ is real-analytic. Differentiating, gives

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left[\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right] = \left[S_N(x,\underline{y}) - \overline{P}'_N(\xi) + \epsilon_N(x, y_{N+1}, \xi)\right]\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}, \tag{7.19}$$

where $\epsilon_N(x,y_{N+1},\xi) := \dfrac{\overline{h}_1(x,\xi)}{\overline{h}_{N+1}(y_{N+1},\xi)}\dfrac{\mathrm{d}}{\mathrm{d}\xi}\left(\dfrac{\overline{h}_{N+1}(y_{N+1},\xi)}{\overline{h}_1(x,\xi)}\right)$. By Lemmas 7.14 and 7.16, $\epsilon_N(x,y_{N+1},\xi)$ is uniformly bounded in N, x, $\underline{y}$, and $|\xi| \le R$.

Fix N. By the intermediate value theorem and the uniform boundedness of $\frac{\mathrm{d}}{\mathrm{d}\xi}\left[\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right]$ on compact subsets of $\xi \in \mathbb{R}$, $\frac{1}{\delta}\left[\frac{\mathrm{d}\widetilde{\pi}_x^{\xi+\delta}}{\mathrm{d}\pi_x} - \frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right]$ is uniformly bounded for $0 < \delta < 1$. By the bounded convergence theorem,

$$\int \lim_{\delta\to 0}\frac{1}{\delta}\left[\frac{\mathrm{d}\widetilde{\pi}_x^{\xi+\delta}}{\mathrm{d}\pi_x} - \frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right]\mathrm{d}\pi_x = \lim_{\delta\to 0}\int \frac{1}{\delta}\left[\frac{\mathrm{d}\widetilde{\pi}_x^{\xi+\delta}}{\mathrm{d}\pi_x} - \frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right]\mathrm{d}\pi_x = 0.$$

So $\int \frac{\mathrm{d}}{\mathrm{d}\xi}\left[\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right]\mathrm{d}\pi_x = 0$, whence by (7.19), $0 = \widetilde{\mathbb{E}}_x^\xi(S_N) - \overline{P}'_N(\xi) + O(1)$ (here $\widetilde{\mathbb{E}}_x^\xi = \widetilde{\mathbb{E}}^\xi(\cdot|\widetilde{X}_1^\xi = x)$). Integrating with respect to x we obtain that $\overline{P}'_N(\xi) = \widetilde{\mathbb{E}}^\xi(S_N) + O(1)$ uniformly in $|\xi| \le R$, as $N \to \infty$.

Differentiating (7.19) again gives

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\left[\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\right] &= \frac{\mathrm{d}}{\mathrm{d}\xi}\left[\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\left(S_N(x,\underline{y}) - \overline{P}'_N(\xi) + \epsilon_N(x,y_{N+1},\xi)\right)\right]\\ &= \frac{\mathrm{d}\widetilde{\pi}_x}{\mathrm{d}\pi_x}\left[\left(S_N(x,\underline{y}) - \overline{P}'_N(\xi) + \epsilon_N(x,y_{N+1},\xi)\right)^2 - \overline{P}''_N(\xi) + \frac{\mathrm{d}\epsilon_N}{\mathrm{d}\xi}\right].\end{aligned}$$

By Lemmas 7.14 and 7.16, $\frac{\mathrm{d}\epsilon_N}{\mathrm{d}\xi}$ is uniformly bounded in x, y_{N+1}, N and $|\xi| \le R$. As before, $\int \frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\mathrm{d}\pi_x = \frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\int \frac{\mathrm{d}\widetilde{\pi}_x^\xi}{\mathrm{d}\pi_x}\mathrm{d}\pi_x = 0$. As $\overline{P}'_N(\xi) = \widetilde{\mathbb{E}}^\xi(S_N) + O(1)$, we get:

$$0 = \widetilde{\mathbb{E}}^{\xi}\left[\left(S_N - \widetilde{\mathbb{E}}^{\xi}(S_N) + O(1)\right)^2\right] - \overline{P}''_N(\xi) + O(1), \tag{7.20}$$

$$= \widetilde{V}^{\xi}(S_N) - \overline{P}''_N(\xi) + O(1)\sqrt{\widetilde{V}^{\xi}(S_N)} + O(1). \tag{7.21}$$

The proof shows that the big-Oh's are uniformly bounded for all N and $|\xi| \le R$.

If $|\xi| \le R$, then $\widetilde{\pi}^{\xi}_{n,n+1}(x, dy)$ are uniformly elliptic with ϵ_0 replaced by $\epsilon_0/(C^4 e^{2KR})$ with the C in Lemma 7.14. Thus by Theorem 3.7, $\widetilde{V}^{\xi}(S_N) \asymp \sum_{n=3}^{N} u_n^2(\xi)$ where $u_n(\xi)$ are the structure constants of $\{\widetilde{X}^{\xi}_n\}$.

Using Corollary 2.10, it is not difficult to show that the hexagon measures associated to $\widetilde{\mathsf{X}}^{\xi}$ and X on Hex (N, n) with $5 \le n \le N < \infty$ are equivalent, and their Radon-Nikodym derivatives are bounded away from zero and infinity uniformly in n, N. Therefore $u_n(\xi) \asymp u_n(\mathsf{X}, \mathsf{f})$. By Theorem 3.6, $\widetilde{V}^{\xi}(S_N) \asymp V_N \to \infty$, uniformly in N and $|\xi| \le R$. By (7.21), $\overline{P}''_N(\xi)/\widetilde{V}^{\xi}(S_N) \xrightarrow[N\to\infty]{} 1$ uniformly on compacts. □

The Choice of a_N: Lemma 7.17(1) with $\xi = 0$ says that $\overline{P}'_N(0) = \mathbb{E}(S_N) + O(1)$. The error term is a nuisance, and we will choose a_n to get rid of it. Given N, let

$$a_n := \mathbb{E}(S_{n-1}) - \overline{P}'_{n-1}(0)\,,\ a_1 := 0. \tag{7.22}$$

This is bounded, because of Lemma 7.17(1). Let $h^{\xi}_n(x) := e^{a_n\xi}\overline{h}_n(x, \xi)$ and

$$p_n(\xi) := \overline{p}_n(\xi) + (a_{n+1} - a_n)\xi. \tag{7.23}$$

The transition kernel $\widetilde{\pi}^{\xi}_{n,n+1}$ is left unchanged, because the differences between $\overline{h}_n$ and h_n and between $\overline{p}_n$ and p_n cancel out. But now, $P'_N(0) = \mathbb{E}(S_N)$, where

$$P_N(\xi) := p_1(\xi) + \cdots + p_N(\xi) \equiv \overline{P}_N(\xi) + (\mathbb{E}(S_N) - \overline{P}'_N(0))\xi. \tag{7.24}$$

Properties of $P_N(\xi)$: Recall that $\mathcal{F}_N(\xi) := \frac{1}{V_N}\log\mathbb{E}(e^{\xi S_N})$, and that $\widetilde{V}^{\xi}$ is the variance with respect to the change of measure $\widetilde{\mathsf{X}}^{\xi}$.

Lemma 7.18 *Suppose $V_N \to \infty$, then $\xi \mapsto P_N(\xi)$ is real-analytic, and:*

(1) $P'_N(0) = \mathbb{E}(S_N)$. (2) $\forall R > 0$, $\exists C(R) > 0$ s.t. $|P'_N(\xi) - \widetilde{\mathbb{E}}^{\xi}(S_N)| \le C(R)$ for all $|\xi| \le R, N \in \mathbb{N}$.

(3) $P''_N(\xi)/\widetilde{V}^{\xi}(S_N) \xrightarrow[N\to\infty]{} 1$ uniformly on compact subsets of ξ.

(4) $P_N(\xi)/V_N = \mathcal{F}_N(\xi) + O(V_N^{-1})$ uniformly on compact subsets of ξ: $\forall R > 0$, $\Delta_N(R) := \sup\limits_{|\xi|\le R}|V_N\mathcal{F}_N(\xi) - P_N(\xi)| = O(1)$, and $\sup\limits_N \Delta_N(R) \xrightarrow[R\to 0^+]{} 0$.

(5) $P'_N(\xi)/V_N = \mathcal{F}'_N(\xi) + O(V_N^{-1})$ uniformly on compact subsets of ξ, as $N \to \infty$: If $\overline{\Delta}_N(R) := \sup\limits_{|\xi|\le R}\left|V_N\mathcal{F}'_N(\xi) - P'_N(\xi)\right|$, then $\sup\limits_N \overline{\Delta}_N(R) < \infty$.

(6) $P_N(\cdot)$ are uniformly strictly convex on compacts: $\forall R > 0\, \exists N(R)$ such that $\inf\limits_{\xi\in[-R,R]}\inf\limits_{N\ge N(R)} P''_N(\xi) > 0$.

Proof The real-analyticity of $P_N(\xi)$ and parts (1)–(3) and (6) follow from Lemma 7.17, the identity $P_N(\xi) = \overline{P}_N(\xi) + (a_{N+1} - a_1)\xi$, and the boundedness of a_n.

The proof of part (4) uses the operators $L^{\xi}_n : L^{\infty}(\mathfrak{S}_{n+1}) \to L^{\infty}(\mathfrak{S}_n)$ from (7.6), $(L^{\xi}_n h)(x) := \int_{\mathfrak{S}_{n+1}} e^{\xi f_n(x,y)} h(y)\pi_{n,n+1}(x, dy) \equiv \mathbb{E}[e^{\xi f_n(X_n, X_{n+1})}h(X_{n+1})|X_n = x]$. Let $h^{\xi}_n := h_n(\cdot, \xi) \in L^{\infty}(\mathfrak{S}_n)$ be the unique positive functions s.t. $\int h^{\xi}_n = e^{a_n\xi}$ and $L^{\xi}_n h^{\xi}_{n+1} = e^{p_n(\xi)}h^{\xi}_n$. Then $h^0_n \equiv 1$, $p_1(\xi) + \cdots + p_N(\xi) = P_N(\xi)$ and

$$\begin{aligned}
&\mathbb{E}_x[e^{\xi S_N}h^{\xi}_{N+1}(X_{N+1})] = \mathbb{E}[\mathbb{E}(e^{\xi S_N}h^{\xi}_{N+1}(X_{N+1})|X_N, \ldots, X_1)|X_1 = x] = \mathbb{E}[e^{\xi S_{N-1}}\mathbb{E}(e^{\xi f_N(X_N, X_{N+1})}h^{\xi}_{N+1}(X_{N+1})|X_N)|X_1 = x] \\
&= \mathbb{E}[e^{\xi S_{N-1}}(L^{\xi}_N h^{\xi}_{N+1})(X_N)|X_1 = x] = e^{p_N(\xi)}\mathbb{E}_x[e^{\xi S_{N-1}}h^{\xi}_N(X_N)] = e^{p_N(\xi)+p_{N-1}(\xi)}\mathbb{E}_x[e^{\xi S_{N-2}}h^{\xi}_{N-1}(X_{N-1})] \\
&= \cdots = e^{p_N(\xi)+\cdots+p_1(\xi)}\mathbb{E}_x[h^{\xi}_1(X_1)] = e^{P_N(\xi)}h^{\xi}_1(x).
\end{aligned} \tag{7.25}$$

By Lemma 7.14, there exists $C_1 = C_1(R) > 1$ such that $C_1^{-1} \le h_j^\xi \le C_1$ for $j = 1, N+1$ and every $|\xi| \le R$ and $N \ge 1$. Thus by (7.25),

$$C_1(R)^{-2}\mathrm{e}^{P_N(\xi)} \le \mathbb{E}\left(\mathrm{e}^{\xi S_N}\right) \le C_1(R)^2\mathrm{e}^{P_N(\xi)}. \tag{7.26}$$

Taking logarithms, we deduce that $|\mathcal{F}_N(\xi) - P_N(\xi)/V_N| \le 2\log C_1(R)/V_N$ for all $N \ge 1$ and $|\xi| \le R$. Equivalently, $\sup_N \Delta_N(R) \le 2\log C_1 < \infty$.

By Lemma 7.16 and the identity $h_n^0 \equiv 1$, $\|h_N^\xi - 1\|_\infty \xrightarrow[\xi\to 0]{} 0$ uniformly in N. Returning to the definition of $C_1(R)$ we find that we may choose $C_1(R) \xrightarrow[R\to 0^+]{} 1$. As before, this implies that $\sup_N \Delta_N(R) \xrightarrow[R\to 0]{} 0$. This proves part (4).

Here is the proof of part (5). Fix $R > 0$, then

$$V_N\mathcal{F}_N'(\xi) = \frac{\mathbb{E}(S_N\mathrm{e}^{\xi S_N})}{\mathbb{E}(\mathrm{e}^{\xi S_N})} = \frac{\widetilde{\mathbb{E}}^\xi(S_N(h_1^\xi/h_{N+1}^\xi))}{\widetilde{\mathbb{E}}^\xi(h_1^\xi/h_{N+1}^\xi)}. \tag{7.27}$$

We have already remarked that X^ξ are uniformly elliptic, and that their uniform ellipticity constants are bounded away from zero for ξ ranging on a compact set. This gives us the mixing bounds in Proposition 2.13 with the same $C_{mix} > 0$, $0 < \theta < 1$ for all $|\xi| \le R$. So $\widetilde{\mathbb{E}}^\xi\left(h_1^\xi/h_{N+1}^\xi\right) = \widetilde{\mathbb{E}}^\xi(h_1^\xi)\widetilde{\mathbb{E}}^\xi\left(1/h_{N+1}^\xi\right) + O(\theta^N)$, as $N \to \infty$ Similarly, $\widetilde{\mathbb{E}}^\xi\left(S_N(h_1^\xi/h_{N+1}^\xi)\right) = \widetilde{\mathbb{E}}^\xi(h_1^\xi)\widetilde{\mathbb{E}}^\xi\left(1/h_{N+1}^\xi\right)\widetilde{\mathbb{E}}^\xi(S_N) + O(1)$ as $N \to \infty$.
The big oh's are uniform for $|\xi| \le R$. By (7.27), $V_N\mathcal{F}_N'(\xi) = \widetilde{\mathbb{E}}^\xi(S_N) + O(1)$ as $N \to \infty$, uniformly for $|\xi| \le R$. Part 5 now follows from part 2. □

The Choice of ξ_N: Recall that to reduce the regime of large deviations to the regime of local deviations, we need a change of measure for which $\widetilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)$. By Lemma 7.18(2) this will be the case for ξ_N such that $P_N'(\xi_N) = z_N$.

The following lemma gives sufficient conditions for the existence of such ξ_N.

Lemma 7.19 *Suppose $V_N \to \infty$, $R > 0$, and*

$$[\widehat{a}_N^R, \widehat{b}_N^R] := \left[\mathcal{F}_N'(-R) - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}_N'(R) - \frac{\mathbb{E}(S_N)}{V_N}\right].$$

(1) For each R there are $C(R)$ and $N(R)$ such that if $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ and $N > N(R)$, then

(a) $\exists!\xi_N \in [-(R+1), (R+1)]$ such that $P_N'(\xi_N) = z_N$; (b) $C(R)^{-1}\left|\frac{z_N - \mathbb{E}(S_N)}{V_N}\right| \le |\xi_N| \le C(R)\left|\frac{z_N - \mathbb{E}(S_N)}{V_N}\right|$;
(c) $\mathrm{sgn}(\xi_N) = \mathrm{sgn}(\frac{z_N - \mathbb{E}(S_N)}{V_N})$; (d) $\left|\widetilde{\mathbb{E}}^{\xi_N}(S_N) - z_N\right| \le C(R)$.

(2) For every $R > 2$ there exists $c(R) > 0$ such that for all N large enough,

$$\text{if} \left|\frac{z_N - \mathbb{E}(S_N)}{V_N}\right| \le c(R), \text{ then } \frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]. \tag{7.28}$$

Consequently, if $\left|\frac{z_N - \mathbb{E}(S_N)}{V_N}\right| < c(R)$, then there exists a unique ξ_N with (a)–(d).

Proof Let $[\widetilde{a}_N^R, \widetilde{b}_N^R] := \left[\frac{P_N'(-R) - \mathbb{E}(S_N)}{V_N}, \frac{P_N'(R) - \mathbb{E}(S_N)}{V_N}\right]$.

Claim: *For all $R > 0$, for all N large enough,*

$$[\widehat{a}_N^R, \widehat{b}_N^R] \subset [\widetilde{a}_N^{R+1}, \widetilde{b}_N^{R+1}] \subset [\widehat{a}_N^{R+2}, \widehat{b}_N^{R+2}]. \tag{7.29}$$

Proof of the Claim: By Lemmas 7.17 and 7.18, there exists a $\delta > 0$ such that for all sufficiently large N, $P_N''(\xi)/V_N \ge \delta$ on $[-(R+2), (R+2)]$. By the mean value theorem, $\widetilde{b}_N^{R+2} \ge \widetilde{b}_N^{R+1} + \delta$, $\widetilde{b}_N^{R+1} \ge \widetilde{b}_N^R + \delta$, $\widetilde{a}_N^{R+2} \le \widetilde{a}_N^{R+1} - \delta$, $\widetilde{a}_N^{R+1} \le \widetilde{a}_N^R - \delta$.

Next, by Lemma 7.18(5), $|\widehat{b}_N^{R'} - \widetilde{b}_N^{R'}| = O(V_N^{-1})$ and $|\widehat{a}_N^{R'} - \widetilde{a}_N^{R'}| = O(V_N^{-1})$ for all $R' \le R+2$. For all N large enough, $|O(V_N^{-1})| < \delta$, and $a_N^{R+2} < \widetilde{a}_N^{R+1} < \widehat{a}_N^{R} < \widehat{b}_N^{R} < \widetilde{b}_N^{R+1} < \widehat{b}_N^{R+2}$.

We can now prove part (1) of the lemma. Let $\varphi_N(\xi) := \dfrac{P_N(\xi) - \xi P_N'(0)}{V_N}$. By Lemma 7.18, for all N large enough, $\varphi_N(\xi)$ is strictly convex, smooth, and $P_N'(\xi_N) = z_N$ iff $\varphi_N'(\xi_N) = \dfrac{z_N - P_N'(0)}{V_N}$.

Fix $R > 0$. By the claim, for all N large enough, if $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$, then $\frac{z_N - P_N'(0)}{V_N} \equiv \frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widetilde{a}_N^{R+1}, \widetilde{b}_N^{R+1}] \equiv \varphi_N'[-(R+1),(R+1)]$. Since φ_N' is continuous and strictly increasing, there is a unique $\xi_N \in [-(R+1),(R+1)]$ s.t. $\varphi_N'(\xi_N) = \frac{z_N - P_N'(0)}{V_N}$. So there is a unique $|\xi_N| \le R+1$ s.t. $P_N'(\xi_N) = z_N$.

This argument shows that for every N sufficiently large, for every $\eta \in [\widehat{a}_N^R, \widehat{b}_N^R]$ there exists a unique $\xi = \xi(\eta) \in [-(R+1),(R+1)]$ such that $\varphi_N'(\xi(\eta)) = \eta$.

By Lemma 7.18(6), $\exists \delta(R) > 0$ so that $\delta(R) \le \varphi_N'' \le \delta(R)^{-1}$ on $[-(R+1),(R+1)]$. So $\eta \mapsto \xi(\eta)$ is $\frac{1}{\delta(R)}$-bi-Lipschitz on $[\widehat{a}_N^R, \widehat{b}_N^R]$. By construction, $\varphi_N'(0) = 0$. So $\xi(0) = 0$, whence by the bi-Lipschitz property

$$\delta(R)|\eta| \le |\xi(\eta)| \le \delta(R)^{-1}|\eta| \text{ on } [\widehat{a}_N^R, \widehat{b}_N^R].$$

Since φ_N is real-analytic and strictly convex, φ_N' is smooth and strictly increasing. By the inverse mapping theorem, $\eta \mapsto \xi(\eta)$ is smooth and strictly increasing. So $\mathrm{sgn}(\xi(\eta)) = \mathrm{sgn}(\eta)$ on $[\widehat{a}_N^R, \widehat{b}_N^R]$. Specializing to the case $\eta = \frac{z_N - \mathbb{E}(S_N)}{V_N}$, gives properties (a)–(c) of ξ_N.

Property (d) is because of Lemma 7.18, which says that $z_N = P_N'(\xi_N) = \widetilde{\mathbb{E}}^{\xi_N}(S_N) + O(1)$. The big oh is uniform because $|\xi_N| \le R+1$.

This completes the proof of part (1). To prove part (2), fix $R > 2$. By (7.29), for all N large enough, $[\widehat{a}_N^R, \widehat{b}_N^R] \supset [\widetilde{a}_N^{R-1}, \widetilde{b}_N^{R-1}] \equiv \varphi_N'[-(R-1),(R-1)]$. Since $\varphi_N'(0) = 0$ and $\varphi_N'' \ge \delta(R)$ on $[-R, R]$, we have that $\varphi'(\pm R) = \pm(R)\varphi''(\eta^\pm)$ for some $\eta^\pm \in [-R, R]$. Therefore $[\widehat{a}_N^R, \widehat{b}_N^R] \supset [-c, c]$, where $c := R\delta(R)$. □

Corollary 7.20 *Suppose $V_N \to \infty$ and $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then for all N large enough, there exists a unique ξ_N such that $P_N'(\xi_N) = z_N$. Furthermore, $\xi_N \to 0$.*

7.3.4 The Asymptotic Behavior of $\widetilde{V}^{\xi_N}(S_N)$

Recall that $\widetilde{V}_N^{\xi}$ denotes the variance of S_N with respect to the change of measure $\widetilde{\mathsf{X}}^{\xi}$ from (7.3).

Lemma 7.21 *Suppose $V_N \xrightarrow[N\to\infty]{} \infty$, and define ξ_N as in Lemma 7.19.*

(1) Suppose $R > 0$ and $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ for all N, then $\widetilde{V}_N^{\xi_N} \asymp V_N$ as $N \to \infty$.
(2) If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\widetilde{V}_N^{\xi_N} \sim V_N$ as $N \to \infty$.
(3) Uniformity: $\forall \epsilon > 0\ \exists \xi^* > 0$ and $\exists N_0 > 1$, so that $\forall N > N_0$, if $|\xi| < \xi^*$, then $\widetilde{V}_N^{\xi}/V_N \in [\mathrm{e}^{-\epsilon}, \mathrm{e}^{\epsilon}]$.

Proof We assume without loss of generality that $\mathbb{E}(f_k(X_k, X_{k+1})) = 0$ for all k; otherwise we subtract suitable constants from f_k, and note that this has no effect on $\widehat{a}_N, \widehat{b}_N, V_N$ or $\widetilde{V}_N^{\xi}$. In particular, $\mathbb{E}(S_N) = 0$ for all N.
If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$, then $|\xi_N| \le R+1$ (Lemma 7.19), and part (1) follows from Lemma 7.17.
If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\xi_N \to 0$ (Corollary 7.20), and part (2) follows from part (3). It remains to prove part (3).

To do this we decompose S_N into weakly correlated large blocks with roughly the same X-variances, and we check that the $\widetilde{\mathsf{X}}^{\xi}$-variance of the i-th block converges *uniformly* in i to its X-variance, as $\xi \to 0$.

We denote the entries of $\widetilde{\mathsf{X}}^{\xi}$ by $\{\widetilde{X}_n^{\xi}\}$, and define for $m \ge n+1$, $S_{n,m} := \sum_{k=n}^{m-1} f_k(X_k, X_{k+1})$,

$$\widetilde{S}^{\xi}_{n,m} := \sum_{k=n}^{m-1} f_k(\widetilde{X}^{\xi}_k, \widetilde{X}^{\xi}_{k+1}), \quad V_{n,m} := \mathrm{Var}(S_{n,m}), \quad \widetilde{V}^{\xi}_{n,m} := \mathrm{Var}(\widetilde{S}^{\xi}_{n,m}), \quad p_{n,m}(\xi) := \sum_{k=n}^{m-1} p_k(\xi).$$

We claim that for all $R > 0$, $n < m$, and $|\xi| \le R$,

$$p_{n,m}(0) = 0, \quad p'_{n,m}(0) = \mathbb{E}(S_{n,m}) = 0, \quad |p'_{n,m}(\xi) - \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n,m})| \le C(R). \tag{7.30}$$

The first identity is because $\overline{h}_n(\cdot, 0) \equiv 1$ and $\overline{p}_n(0) = 1$, by the uniqueness of the fundamental solution. The second identity is because $p'_{n,m}(0) = P'_{m-1}(0) - P'_{n-1}(0) = \mathbb{E}(S_{m-1}) - \mathbb{E}(S_{n-1}) = 0 - 0 = 0$. The third part of (7.30) can be shown by applying the proof of Lemma 7.18(2) to the truncated Markov chains $\{X_k\}_{k \ge n}$.

Similarly, applying the proof of Lemma 7.17 to the truncated chain $\{X_k\}_{k\ge n}$ gives a constant M_0 such that for all $n < m$ and $|\xi| \le R$,

$$V_{n,m} \ge M_0 \Rightarrow \begin{cases} C(R)^{-1} \le \widetilde{V}^{\xi}_{n,m}/V_{n,m} \le C(R) \\ 2^{-1} \le p''_{n,m}(\xi)/\widetilde{V}^{\xi}_{n,m} \le 2. \end{cases} \tag{7.31}$$

Step 1 (Mixing Estimates). *Recall that* $K = \operatorname{ess\,sup} |\mathsf{f}|$. *There are* $C^*_{mix} = C^*_{mix}(K, R) > 1$ *and* $\eta = \eta(K, R) \in (0, 1)$ *such that for every* $|\xi| \le R$ *and* $k \le n < m$*,*

(1) $\left|\mathrm{Cov}(f_m(\widetilde{X}^{\xi}_m, \widetilde{X}^{\xi}_{m+1}), f_n(\widetilde{X}^{\xi}_n, \widetilde{X}^{\xi}_{n+1}))\right| \le C^*_{mix}\eta^{m-n}$; (2) $\|\widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n,m}|\widetilde{X}^{\xi}_k) - \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n,m})\|_\infty \le C^*_{mix}$;

(3) $\|\widetilde{\mathbb{E}}^{\xi}((\widetilde{S}^{\xi}_{n,m})^2|\widetilde{X}_k) - \widetilde{\mathbb{E}}^{\xi}((\widetilde{S}^{\xi}_{n,m})^2)\|_\infty \le C^*_{mix}(1 + |p'_{n,m}(\xi)|)$.

Proof of the Step: Let $\widetilde{f}^{\xi}_i := f_i(\widetilde{X}^{\xi}_i, \widetilde{X}^{\xi}_{i+1})$ and $\widehat{f}^{\xi}_i := \widetilde{f}^{\xi}_i - \widetilde{\mathbb{E}}^{\xi}(\widetilde{f}^{\xi}_i)$. Then $|\widehat{f}^{\xi}_i| \le 2K$.

The Markov chains $\widetilde{X}^{\xi}_n$ are uniformly elliptic with the same ellipticity constant $\epsilon_0(R) > 0$ for all $|\xi| \le R$. By Proposition 2.13, there are constants $C'_{mix} = C'_{mix}(K, R) > 0$ and $\eta = \eta(K, R) \in (0, 1)$ such that for all $|\xi| \le R$ and for every $i > k$,

$$\|\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_i|\widetilde{X}^{\xi}_k)\|_\infty \le C'_{mix}\eta^{i-k}. \tag{7.32}$$

Thus $\|\widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n,m}|\widetilde{X}^{\xi}_k) - \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n,m})\|_\infty \le \sum \|\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_i|\widetilde{X}^{\xi}_k)\|_\infty < \frac{C'_{mix}}{1-\eta}$, a constant independent of ξ. This proves part 2.

Next, $\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_m\widehat{f}^{\xi}_n) = \widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_n\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_m|\widetilde{X}^{\xi}_n, \widetilde{X}^{\xi}_{n+1})) = \widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_n\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_m|\widetilde{X}^{\xi}_{n+1}))$; therefore $|\widetilde{\mathbb{E}}^{\xi}(\widehat{f}^{\xi}_m\widehat{f}^{\xi}_n)| \le 2K \cdot C'_{mix}\eta^{m-n-1}$. Part 1 follows.

Henceforth, we fix ξ and set $\widetilde{\mathbb{E}} = \widetilde{\mathbb{E}}^{\xi}$, $\widehat{f}_i = \widehat{f}^{\xi}_i$, $\widetilde{X}^{\xi}_k = \widetilde{X}_k$. We claim that:

$$\|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)\|_\infty \le \mathrm{const.}\eta^{\frac{i-k}{2}}\eta^{\frac{j-i}{2}} \quad (|\xi| \le R, k \le i < j). \tag{7.33}$$

Fix $k \le i < j$. To prove (7.33), we will estimate the LHS in two ways:

- $\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j) = \widetilde{\mathbb{E}}(g|\widetilde{X}_k)$, with $g := \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}_i, \widetilde{X}_{i+1}, \widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)$. A calculation shows that g depends only on $\widetilde{X}_i$ and $\widetilde{X}_{i+1}$. By (7.32),

$$\|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)\|_\infty = \|\widetilde{\mathbb{E}}(g|\widetilde{X}_k)\|_\infty \le \mathrm{const.}\eta^{i-k}. \tag{7.34}$$

- Consider the Markov chain $\mathsf{Y} = (\widetilde{X}^{\xi}_k, \widetilde{X}^{\xi}_{k+1}, \dots)$ with initial distribution $\widetilde{X}^{\xi}_k = x_k$ for some fixed $x_k \in \mathfrak{S}_k$. Y has the same uniform ellipticity constant as $\widetilde{\mathsf{X}}^{\xi}$, and therefore Y and $\widetilde{\mathsf{X}}^{\xi}$ have the same mixing rates. Thus for every $x_k \in \mathfrak{S}_k$, $|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}^{\xi}_k = x_k) - \widetilde{\mathbb{E}}(\widehat{f}_i|\widetilde{X}_k = x_k)\widetilde{\mathbb{E}}(\widehat{f}_j|\widetilde{X}_k = x_k)| \le \mathrm{const.}\eta^{j-i}$ uniformly in x_k and k, i, j. This, (7.32), and $i \ge k$ lead to $\|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}^{\xi}_k)\|_\infty \le \|\widehat{f}_i\|_\infty\|\widetilde{\mathbb{E}}(\widehat{f}_j|\widetilde{X}_k)\|_\infty + \mathrm{const.}\eta^{j-i} \le \mathrm{const.}\eta^{j-i}$.
  Obviously, this implies that $|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)| \le \mathrm{const.}\eta^{j-i}$ and

$$\|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}^{\xi}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)\|_\infty \le \mathrm{const.}\eta^{j-i}. \tag{7.35}$$

By (7.34) and (7.35), the RHS of (7.33) is $O(\min\{\eta^{i-k}, \eta^{j-i}\})$. Since $\min\{|a|, |b|\} \le \sqrt{|ab|}$, (7.33) follows.

We can now prove the Step 1(3). Let $\widetilde{S}_{n,m} := \widetilde{S}^{\xi}_{n,m}$ and $\widehat{S}_{n,m} := \widetilde{S}_{n,m} - \widetilde{\mathbb{E}}(\widetilde{S}_{n,m})$.

$$\|\widetilde{\mathbb{E}}(\widehat{S}^2_{n,m}|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{S}^2_{n,m})\|_\infty \le 2 \sum_{n\le i\le j\le m} \|\widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{f}_i\widehat{f}_j)\|_\infty \le C \sum_{n\le i\le j\le m} \eta^{\frac{i-k}{2}}\eta^{\frac{j-i}{2}} \le C(1-\eta^{1/2})^{-2} =: C_1.$$

$$\|\widetilde{\mathbb{E}}(\widetilde{S}^2_{n,m}|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widetilde{S}^2_{n,m})\|_\infty = \|\widetilde{\mathbb{E}}[(\widehat{S}_{n,m}+\widetilde{\mathbb{E}}(\widetilde{S}_{n,m}))^2|\widetilde{X}_k] - \widetilde{\mathbb{E}}[(\widehat{S}_{n,m}+\widetilde{\mathbb{E}}(\widetilde{S}_{n,m}))^2]\|_\infty$$
$$= \|\widetilde{\mathbb{E}}(\widehat{S}^2_{n,m}|\widetilde{X}_k) + 2\widetilde{\mathbb{E}}(\widehat{S}_{n,m}|\widetilde{X}_k)\widetilde{\mathbb{E}}(\widetilde{S}_{n,m}) + \widetilde{\mathbb{E}}(\widetilde{S}_{n,m})^2 - \widetilde{\mathbb{E}}(\widehat{S}^2_{n,m}) - \widetilde{\mathbb{E}}(\widetilde{S}_{n,m})^2\|_\infty$$
$$\le \|\widetilde{\mathbb{E}}(\widehat{S}^2_{n,m}|\widetilde{X}_k) - \widetilde{\mathbb{E}}(\widehat{S}^2_{n,m})\|_\infty + 2\|\widetilde{\mathbb{E}}(\widehat{S}_{n,m}|\widetilde{X}_k)\|_\infty|\widetilde{\mathbb{E}}(\widetilde{S}_{n,m})| \le C_1 + \sum_{n\le i\le m} \|\widetilde{\mathbb{E}}(\widehat{f}_i|\widetilde{X}_k)\|_\infty \left(|p'_{n,m}(\xi)| + |\widetilde{\mathbb{E}}(\widetilde{S}_{n,m}) - p'_{n,m}(\xi)|\right)$$
$$\le \text{const.}(1 + |p'_{n,m}(\xi)|), \text{ by (7.30) and (7.32).}$$

The constant is uniform for $|\xi| \le R$, k, n and m. Increasing C^*_{mix}, we obtain (3).

Step 2 (Block Decomposition). *For every $\epsilon > 0$ small enough, for every $R > 1$, there exists $M > 1$ and a sequence of integers $n_i \uparrow \infty$ such that:*

(1) $M > 1000(C^*_{mix})^2/\epsilon$; (2) $M \le V_{n_i,n_{i+1}} \le 2M$;

(3) $|\mathrm{Cov}(\widetilde{S}^\xi_{n_i,n_{i+1}}, \widetilde{S}^\xi_{n_j,n_{j+1}})| \le C^\#_{mix}\eta^{n_j-n_{i+1}}$ $\forall |\xi| \le R$ and $i < j$, where $C^\#_{mix}$ is independent of M, i, j and ξ;

(4) For all $|\xi| \le R$, for all $i > 3$, for all $n \in [n_i, n_{i+1}]$,

$$e^{-\epsilon} \le \frac{\widetilde{V}^\xi_n}{\sum_{k=1}^{i-1} \widetilde{V}^\xi_{n_k,n_{k+1}} + \widetilde{V}^\xi_{n_i,n}} \le e^\epsilon; \tag{7.36}$$

(5) $M^* := \sup_i \sup_{n\in[n_i,n_{i+1}]} \sup_{|\xi|\le R} |p''_{n_i,n}(\xi)| < \infty$; (6) $M^\# := \sup_i \sup_{n\in[n_i,n_{i+1}]} \sup_{|\xi|\le R} \sup_{x_{n_i}\in\mathfrak{S}_{n_i}} \mathbb{E}(e^{\xi S_{n_i,n}}|X_{n_i} = x_{n_i}) < \infty$.

Proof of the Step. Fix $M > \max\left\{M_0, \dfrac{1000(C^*_{mix})^2}{\epsilon}, 2\left(K^2 + \dfrac{C^*_{mix}}{1-\eta}\right), \dfrac{8C^*_{mix}C(K,R)}{\epsilon(1-\eta(K,R))^3}\right\}$. Set $n_1 := 1$, and define n_i inductively by $n_{i+1} := \min\{n > n_i : V_{n_i,n_{i+1}} > M\}$. Such $n > n_i$ exist, because $V_{n_i,n} \xrightarrow[n\to\infty]{} \infty$:

$$\infty \xleftarrow[\infty\leftarrow n]{} V_{n-1} \equiv V_{1,n} = V_{1,n_i} + V_{n_i,n} + 2\mathrm{Cov}(S_{1,n_i}, S_{n_i,n}) = V_{n_i,n} + V_{1,n_i} + O\left(\sum_{m=1}^{n_i-1}\sum_{k=0}^{\infty} |\mathrm{Cov}(f_m, f_{n_i+k})|\right) \overset{!}{=} V_{n_i,n} + V_{1,n_i} + O(1),$$

by Step 1 with $\xi = 0$.

By construction, $V_{n_i,n_{i+1}} > M$, and
$V_{n_i,n_{i+1}} \le V_{n_i,n_{i+1}-1} + |V_{n_i,n_{i+1}} - V_{n_i,n_{i+1}-1}| \le M + |V_{n_i,n_{i+1}} - V_{n_i,n_{i+1}-1}|$, (by the minimality of n_{i+1})

$$\le M + \mathrm{Var}(f_{n_{i+1}-1}(X_{n_{i+1}-1}, X_{n_{i+1}})) + 2|\mathrm{Cov}(f_{n_{i+1}-1}(X_{n_{i+1}-1}, X_{n_{i+1}}), S_{n_i,n_{i+1}-1})| \le M + 2\left(K^2 + \frac{C^*_{mix}}{1-\eta}\right) \le 2M$$

by the choice of M and η. So $M < V_{n_i,n_{i+1}} \le 2M$, and $\{n_i\}$ satisfies part 2.

If $i < j$, then $|\mathrm{Cov}(\widetilde{S}^\xi_{n_i,n_{i+1}}, \widetilde{S}^\xi_{n_j,n_{j+1}})| \le \sum_{k=n_i}^{n_{i+1}-1}\sum_{\ell=n_j}^{n_{j+1}-1} C^*_{mix}\eta^{\ell-k}$. The last sum is equal to

$$C^*_{mix}\sum_{k=n_i}^{n_{i+1}-1}\frac{\eta^{n_j-k}}{1-\eta} = \frac{C^*_{mix}\eta^{n_j-n_{i+1}}}{1-\eta}\sum_{k=n_i}^{n_{i+1}-1}\eta^{n_{i+1}-k} \le \frac{C^*_{mix}\eta^{n_j-n_{i+1}}}{(1-\eta)^2}.$$ Part 3 follows with $C^\#_{mix} := C^*_{mix}/(1-\eta)^2$.

Fix $n \in [n_i, n_{i+1}]$, then

$$\left|\widetilde{V}_n^{\xi} - \sum_{k=1}^{i-1} \widetilde{V}_{n_k,n_{k+1}}^{\xi} - \widetilde{V}_{n_i,n}^{\xi}\right| \le 2 \sum_{1\le k<\ell\le i-1} |\mathrm{Cov}(\widetilde{S}_{n_k,n_{k+1}}^{\xi}, \widetilde{S}_{n_\ell,n_{\ell+1}}^{\xi})| + 2\sum_{1\le k\le i-1} |\mathrm{Cov}(\widetilde{S}_{n_k,n_{k+1}}^{\xi}, \widetilde{S}_{n_i,n}^{\xi}) \le$$

$$2\sum_{1\le k<\ell\le i-1} C_{mix}^{\#}\eta^{n_\ell - n_{k+1}} + 2\sum_{1\le k\le i-1} C_{mix}^{\#}\eta^{n_i-n_{k+1}} \le 4\sum_{1\le k<\ell\le i} C_{mix}^{\#}\eta^{\ell-k-1} = 4C_{mix}^{\#}\sum_{k=1}^{i}\sum_{\ell=k+1}^{i}\eta^{\ell-k-1} \le \frac{4C_{mix}^{*} i}{(1-\eta)^3}.$$

By (7.31) and part 2, $\sum_{k=1}^{i-1} \widetilde{V}_{n_k,n_{k+1}}^{\xi} \ge \dfrac{M(i-1)}{C(R)}$. So

$$\left|\frac{\widetilde{V}_n^{\xi}}{\sum_{k=1}^{i-1}\widetilde{V}_{n_k n_{k+1}}^{\xi} + \widetilde{V}_{n_i,n}^{\xi}} - 1\right| \le \frac{\left(\frac{4C_{mix}^{*}}{(1-\eta)^3}\right) i}{C(R)^{-1}M(i-1)} = \frac{1}{M}\cdot\frac{4C_{mix}^{*}C(R)}{(1-\eta)^3}\cdot\frac{i}{i-1} \le \frac{\epsilon}{2}\cdot\frac{i}{i-1}.$$

If $i > 3$, then the last bound is less than $\frac{3}{4}\epsilon$, and part 4 follows for all ϵ small.

Part 5 is a uniform bound on $|p''_{n_i,n}(\xi)|$ for $i\in\mathbb{N}$, $n\in[n_i,n]$, $|\xi|\le R$. By construction, $V_{n_i,n}\le 2M$. By Theorem 3.7, this implies a uniform upper bound on $\sum_{k=n_i}^{n-1} u_k^2$. We have already seen that the structure constants of $\{X_n\}$ and $\{\widetilde{X}_n^{\xi}\}$ are equal up to a bounded multiplicative error. So the same theorem, applied to the Markov chain $\{\widetilde{X}_k^{\xi}\}_{k\ge n_i}$, gives a uniform upper bound for $\widetilde{V}_{n_i,n}^{\xi}$, whence $\sup_i \sup_{n\in[n_i,n_{i+1}]}\sup_{|\xi|\le R} \widetilde{V}_{n_i,n}^{\xi} < \infty$.

A routine modification of the argument used to show (7.20) gives $\left|p''_{n_i,n}(\xi) - \widetilde{\mathbb{E}}^{\xi}\left[\left(\widetilde{S}_{n_i,n}^{\xi} - \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}_{n_i,n}^{\xi}) + O(1)\right)^2\right]\right| \le C$.

$\widetilde{\mathbb{E}}^{\xi}\left[\left(\widetilde{S}_{n_i,n}^{\xi} - \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}_{n_i,n}^{\xi}) + O(1)\right)^2\right]$ is uniformly bounded because of the bound on $\widetilde{V}_{n_i,n}^{\xi}$ and the Minkowski inequality, so part 5 follows.

Given $x_{n_i}\in\mathfrak{S}_{n_i}$, let $\mathbb{E}_{x_{n_i}}(\cdot) := \mathbb{E}(\cdot|X_{n_i} = x_{n_i})$. We have the following variant of (7.25) (with the same proof): $\mathbb{E}_{x_{n_i}}(e^{\xi S_{n_i,n}} h_n^{\xi}(X_n)) = e^{p_{n_i n}(\xi)} h_{n_i}^{\xi}(x_{n_i})$. By Lemma 7.14, there is a constant $\widehat{C} := \widehat{C}(K,R)$ independent of x_{n_i} such that for all $n<m$ and $|\xi|<R$, $h_n^{\xi}, h_{n_i}^{\xi}\in[\widehat{C}^{-1},\widehat{C}]$. It follows that $\mathbb{E}_{x_{n_i}}(e^{\xi S_{n_i,n}}) \le \widehat{C}^2 e^{p_{n_i,n}(\xi)}$.

By (7.30) and part 5, $|p_{n_i,n}(\xi)| \le \frac{1}{2}M^*R^2$ on $[-R,R]$. So $\mathbb{E}_{x_{n_i}}(e^{\xi S_{n_i,n}}) \le M^{\#}$, where $M^{\#} := \widehat{C}^2\exp(\frac{1}{2}M^*R^2)$.

Step 3 (Block Variance). *Fix M and n_i as in Step 2. There is $\xi^*>0$ such that for all $|\xi|<\xi^*$, $i\in\mathbb{N}$, and $n_i\le n\le n_{i+1}$, $|\widetilde{V}_{n_i,n}^{\xi} - V_{n_i,n}| < \epsilon M/10$.*

Proof of the Step. Fix $0<\delta<(3M^{\#})^{-1}\wedge R$, and choose $L>0$ so big that $t^2 < \delta e^{\delta|t|}$, whenever $|t|>L$. Given $x_n\in\mathfrak{S}_n$, let $\mathbb{E}_{x_n}(\cdot) := \mathbb{E}(\cdot|X_n = x_n)$ and $\widetilde{\mathbb{E}}_{x_n}^{\xi}(\cdot) := \widetilde{\mathbb{E}}^{\xi}(\cdot|\widetilde{X}_n^{\xi} = x_n)$. By the definition of $\widetilde{\mathsf{X}}^{\xi}$, for every $i\in n$ and $n\in[n_i,n_{i+1}]$,

$$\widetilde{\mathbb{E}}_{x_{n_i}}^{\xi}[(\widetilde{S}_{n_i,n}^{\xi})^2] - \mathbb{E}_{x_{n_i}}(S_{n_i,n}^2) = \mathbb{E}_{x_{n_i}}\left[\left(e^{\xi S_{n_i,n} - p_{n_i,n}(\xi)}\frac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})} - 1\right)S_{n_i,n}^2\right]. \tag{7.37}$$

We will choose $\xi^*>0$ so that some of the terms inside the brackets are close to constants, whenever $|\xi|<\xi^*$. Firstly, by (7.30) and Step 2(5), if $\xi^*<R$, then $|p_{n_i,n}(\xi)| \le \dfrac{1}{2}M^*\xi^{*2}$ for all $|\xi|<\xi^*$. Next, by Lemma 7.16 and the identity $h_n^0\equiv 1$, $\|h_n^{\xi}-1\|_\infty \le c(R)|\xi|$ for some $c(R)$ and all n. Therefore by decreasing ξ^*, we can guarantee that for all $|\xi|<\xi^*$, $e^{-3c(R)\xi^*} \le \dfrac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})} \le e^{3c(R)\xi^*}$ a.s. Finally (for reasons which will soon become apparent), we make ξ^* so small that

$$\xi^*+\delta<R,\quad \left|e^{\pm\xi^*(L+\frac{1}{2}M^*\xi^{*2}+3c(R))}-1\right|L^2<1,\quad e^{\frac{1}{2}M^*\xi^{*2}+3c(R)\xi^*}<2,\quad M^*\xi^*<1.$$

We now return to (7.37), and decompose the RHS into the sum I+II+II, where I,II and III are the contributions to the expectation from $[|S_{n_i,n}|\le L]$, $[S_{n_i,n}>L]$, and $[S_{n_i,n}<-L]$ respectively. Suppose $|\xi|<\xi^*$, then $|p_{n_i,n}(\xi)| \le \frac{1}{2}M^*\xi^{*2}$, so

$$|\mathrm{I}| \le \mathbb{E}_{x_{n_i}}\left[\left|e^{\xi S_{n_i,n}-p_{n_i,n}(\xi)}\frac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})}-1\right|S_{n_i,n}^2 1_{[|S_{n_i,n}|\le L]}\right] = |e^{\pm\xi^*(L+\frac{1}{2}M^*\xi^{*2}+3c(R))}-1|L^2 < 1, \text{ if } \xi^* \text{ is small enough.}$$

$$\begin{aligned}|\mathrm{II}| &\le \mathbb{E}_{x_{n_i}}\left[\left|e^{\xi S_{n_i,n}-p_{n_i,n}(\xi)}\frac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})}-1\right|S_{n_i,n}^2 1_{[S_{n_i,n}>L]}\right] \le \mathbb{E}_{x_{n_i}}\left[\left|e^{\xi S_{n_i,n}-p_{n_i,n}(\xi)}\frac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})}-1\right|\delta e^{\delta S_{n_i,n}}\right],\\ &\le \mathbb{E}_{x_{n_i}}\left[e^{(\xi+\delta)S_{n_i,n}}\right]e^{\frac{1}{2}M^*\xi^{*2}+3c(R)\xi^*}\delta + \mathbb{E}_{x_{n_i}}\left[e^{\delta S_{n_i,n}}\right]\delta \le \delta M^{\#}\left(e^{\frac{1}{2}M^*\xi^{*2}+3c(R)\xi^*}+1\right) < 3\delta M^{\#} < 1,\end{aligned}$$

There the second inequality follows by the choice of L and the last one is by the choice of δ, $M^{\#}$ and ξ^*.

$$|\mathrm{III}|\le\mathbb{E}_{x_{n_i}}\left[\left|e^{\xi S_{n_i,n}-p_{n_i,n}(\xi)}\frac{h_n^{\xi}(X_n)}{h_{n_i}^{\xi}(x_{n_i})}-1\right|S_{n_i,n}^2 1_{[S_{n_i,n}<-L]}\right]\le\mathbb{E}_{x_{n_i}}\left[e^{(\xi-\delta)S_{n_i,n}}\right]e^{\frac{1}{2}M^*\xi^{*2}+3c(R)\xi^*}\delta + \mathbb{E}_{x_{n_i}}\left[e^{-\delta S_{n_i,n}}\right]\delta<1.$$

Thus, $|\widetilde{\mathbb{E}}^{\xi}_{x_{n_i}}[(\widetilde{S}^{\xi}_{n_i,n})^2] - \mathbb{E}_{x_{n_i}}(S^2_{n_i,n})| \le |\mathrm{I}| + |\mathrm{II}| + |\mathrm{III}| \le 3$. The same argument also shows that $|\widetilde{\mathbb{E}}^{\xi}_{x_{n_i}}[\widetilde{S}^{\xi}_{n_i,n}] - \mathbb{E}_{x_{n_i}}(S_{n_i,n})| \le 3 < 3C^*_{mix}$.

Applying Step 1(3) twice (once for $|\xi| < \xi^*$, and once for $\xi = 0$), we obtain

$$|\widetilde{\mathbb{E}}^{\xi}[(\widetilde{S}^{\xi}_{n_i,n})^2] - \mathbb{E}(S^2_{n_i,n})| \le 3C^*_{mix} + 2C^*_{mix}(1 + |p'_{n_i,n}(\xi)|) \le 3C^*_{mix}(1 + M^*\xi^*) < 6C^*_{mix}.$$

Similarly, $|\widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n_i,n}) - \mathbb{E}(S_{n_i,n})| \le 5C^*_{mix}$. Since $\mathbb{E}(S_{n_i,n}) = 0$, the last estimate gives $|\widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n_i,n})| \le 5C^*_{mix}$. Thus,

$$|\widetilde{V}^{\xi}_{n_i,n} - V_{n_i,n}| \le |\widetilde{\mathbb{E}}^{\xi}[(\widetilde{S}^{\xi}_{n_i,n})^2] - \mathbb{E}(S^2_{n_i,n})| + \widetilde{\mathbb{E}}^{\xi}(\widetilde{S}^{\xi}_{n_i,n})^2 < 50(C^*_{mix})^2 < \frac{\epsilon M}{10}. \tag{7.38}$$

Proof of Part (3) of the Lemma. Fix $\epsilon > 0$ very small. $V_{n_k,n_{k+1}} \ge M$, and by (7.38), for every $|\xi| < \xi^*$, $|\widetilde{V}^{\xi}_{n_k,n_{k+1}} - V_{n_k,n_{k+1}}| < (\epsilon/10)V_{n_k,n_{k+1}}$. So for all k, $e^{-\epsilon} < \widetilde{V}^{\xi}_{n_k,n_{k+1}}/V_{n_k,n_{k+1}} < e^{\epsilon}$. Also by (7.38), for all $n \in [n_k, n_{k+1}]$, $|\widetilde{V}^{\xi}_{n_k,n} - V_{n_k,n}| < (\epsilon/10)V_{n_1,n_2}$.

Fix n so large that $V_n > 4M$. Then $n \in [n_i, n_{i+1}]$ with $i \ge 2$, and by Step 2(4)

$$\frac{\widetilde{V}^{\xi}_n}{V_n} = \frac{\widetilde{V}^{\xi}_n}{\widetilde{V}^0_n} = \frac{e^{\pm\epsilon}(\sum_{k=1}^{i-1}\widetilde{V}^{\xi}_{n_k,n_{k+1}} + \widetilde{V}^{\xi}_{n_i,n})}{e^{\pm\epsilon}(\sum_{k=1}^{i-1}\widetilde{V}^{0}_{n_k,n_{k+1}} + \widetilde{V}^{0}_{n_i,n})} = \frac{e^{\pm2\epsilon}(\sum_{k=1}^{i-1}V_{n_k,n_{k+1}} + V_{n_i,n} \pm \frac{\epsilon}{10}V_{n_1,n_2})}{e^{\pm\epsilon}(\sum_{k=1}^{i-1}V_{n_k,n_{k+1}} + V_{n_i,n})}.$$

The last fraction is $e^{\pm3\epsilon}(1 \pm \frac{\epsilon}{10})$. If ϵ is small enough, this is inside $[e^{-4\epsilon}, e^{4\epsilon}]$.

We proved the lemma, with 4ϵ instead of ϵ. □

7.3.5 Asymptotics of the Log Moment Generating Functions

Let X, Y be two random variables on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Suppose X has finite variance, and Y is positive and bounded. Let $\mathrm{Var}^Y(X)$ be the variance of X with respect to the change of measure $\frac{Y}{\mathbb{E}(Y)}d\mathbb{P}$. Equivalently, $\mathrm{Var}^Y(X) := \dfrac{\mathbb{E}(X^2Y)}{\mathbb{E}(Y)} - \left(\dfrac{\mathbb{E}(XY)}{\mathbb{E}(Y)}\right)^2$.

Lemma 7.22 *Suppose* $0 < \mathrm{Var}(X) < \infty$ *and* $C^{-1} \le Y \le C$ *with* C *a positive constant, then* $C^{-4}\mathrm{Var}(X) \le \mathrm{Var}^Y(X) \le C^4\mathrm{Var}(X)$.

Proof Let (X_1, Y_1), (X_2, Y_2) be two independent copies of (X, Y), then $\mathrm{Var}^Y(X) = \dfrac{1}{2}\dfrac{\mathbb{E}[(X_1 - X_2)^2 Y_1 Y_2]}{\mathbb{E}(Y_1 Y_2)} = C^{\pm 4}\dfrac{1}{2}\mathbb{E}[(X_1 - X_2)^2] = C^{\pm 4}\mathrm{Var}(X)$. □

Proof of Theorem 7.3 Let f be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X, such that $V_N := \mathrm{Var}(S_N) \neq 0$ for $N \geq N_0$, and let $\mathcal{F}_N(\xi) := \frac{1}{V_N}\log\mathbb{E}(e^{\xi S_N})$.

Since $\|S_N\|_\infty < \infty$, we may differentiate under the expectation and obtain that for all k, $\frac{d^k}{d\xi^k}\mathbb{E}(e^{\xi S_N}) = \mathbb{E}(S_N^k e^{\xi S_N})$. A direct calculation shows that denoting $Y_N^\xi := e^{\xi S_N}$, we have

$$\mathcal{F}_N'(\xi) = \frac{1}{V_N}\frac{\mathbb{E}(S_N e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} = \frac{1}{V_N}\mathbb{E}^{Y_N^\xi}(S_N), \quad \mathcal{F}_N''(\xi) = \frac{1}{V_N}\left[\frac{\mathbb{E}(S_N^2 e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})} - \left(\frac{\mathbb{E}(S_N e^{\xi S_N})}{\mathbb{E}(e^{\xi S_N})}\right)^2\right] = \frac{\mathrm{Var}^{Y_N^\xi}(S_N)}{V_N}.$$

Part 1: Substituting $\xi = 0$ gives $\mathcal{F}_N(0) = 0$, $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$ and $\mathcal{F}_N''(0) = 1$.

Part 2: $\mathcal{F}_N''(\xi) = 0 \Leftrightarrow \mathrm{Var}^{Y_N^\xi}(S_N) = 0 \Leftrightarrow S_N = \text{const}\ \frac{Y_N^\xi}{\mathbb{E}(Y_N^\xi)}d\mathbb{P}$–a.s. $\Leftrightarrow S_N = \text{const}\ \mathbb{P}$–a.s. $\Leftrightarrow \mathrm{Var}(S_N) = 0$. So $\mathcal{F}_N$ is strictly convex on $\mathbb{R}$ for all $N > N_0$.

Part 3: $\widetilde{V}^\xi(S_N) \equiv \mathrm{Var}^{Z_N}(S_N)$, where $Z_N^\xi := e^{\xi S_N}\frac{h_{N+1}^\xi}{h_1^\xi}$ (the normalization constant does not matter). Next, $Z_N^\xi \equiv Y_N^\xi W_N^\xi$, where $W_N^\xi := h_{N+1}^\xi / h_1^\xi$. Lemma 7.14 says that for every $R > 0$ there is a constant $C = C(R)$ s.t. $C^{-1} \leq W_N^\xi \leq C$ for all N and $|\xi| \leq R$. Lemma 7.16 and the obvious identity $h_n^0 \equiv 1$ imply that $W_N^\xi \xrightarrow[\xi \to 0]{} 1$ uniformly in N. So there is no loss of generality in assuming that $C(R) \xrightarrow[R \to 0]{} 1$.

By Lemma 7.22 with the probability measure $\frac{e^{\xi S_N}}{\mathbb{E}(e^{\xi S_N})}d\mathbb{P}$ and $Y = W_N^\xi$,

$$\frac{\widetilde{V}^\xi(S_N)}{V_N \mathcal{F}_N''(\xi)} = \frac{\mathrm{Var}^{Y_N^\xi W_N^\xi}(S_N)}{\mathrm{Var}^{Y_N^\xi}(S_N)} \in \left[C(R)^{-4}, C(R)^4\right],\ \forall |\xi| \leq R,\quad N \geq 1. \tag{7.39}$$

By Lemma 7.17, $\widetilde{V}^\xi(S_N) \asymp V_N$ uniformly on compact sets of ξ, and by Lemma 7.21 for every ϵ there exists $\delta, N_\epsilon > 0$ s.t. $e^{-\epsilon} < \widetilde{V}^\xi(S_N)/V_N < e^\epsilon$ for all $N > N_\epsilon$ and $|\xi| \leq \delta$. It follows that for every R there exists $C_2(R) > 1$ such that $C_2(R) \xrightarrow[R \to 0]{} 1$ and $C_2(R)^{-1} \leq \mathcal{F}_N''(\xi) \leq C_2(R)$ for all $|\xi| \leq R$.

Part 4: Fix $\epsilon > 0$. Since $C_2(R) \xrightarrow[R \to 0]{} 1$, there exist $\delta > 0$ and $N_\epsilon \in \mathbb{N}$ such that $e^{-\epsilon} \leq \mathcal{F}_N''(\xi) \leq e^\epsilon$ for all $|\xi| \leq \delta, N \geq N_\epsilon$. So $\mathcal{F}_N(\xi) = \mathcal{F}_N(0) + \int_0^\xi \left(\mathcal{F}_N'(0) + \int_0^\eta \mathcal{F}_N''(\alpha)d\alpha\right)d\eta = \frac{\mathbb{E}(S_N)}{V_N}\xi + \frac{1}{2}e^{\pm\epsilon}\xi^2$. □

7.3.6 Asymptotics of the Rate Functions

The **rate functions** $\mathcal{I}_N(\eta)$ are the Legendre transforms of $\mathcal{F}_N(\xi) = \frac{1}{V_N}\log\mathbb{E}(e^{\xi S_N})$. Recall that the Legendre transform of a continuously differentiable and strictly convex function $\varphi : \mathbb{R} \to \mathbb{R}$ is the function $\varphi^* : (\inf\varphi', \sup\varphi') \to \mathbb{R}$ given by

$$\varphi^*(\eta) = \xi\eta - \varphi(\xi), \quad \text{for the unique } \xi \text{ such that } \varphi'(\xi) = \eta. \tag{7.40}$$

Lemma 7.23 *Suppose $\varphi(\xi)$ is strictly convex and twice differentiable on $\mathbb{R}$, and let $\varphi'(\pm\infty) := \lim_{\xi \to \pm\infty}\varphi'(\xi)$. Then the Legendre transform φ^* is strictly convex and twice differentiable on $(\varphi'(-\infty), \varphi'(+\infty))$. In addition, for every $t \in \mathbb{R}$,*

$$\varphi^*(\varphi'(t)) = t\varphi'(t) - \varphi(t),\ (\varphi^*)'(\varphi'(t)) = t\text{ , and } (\varphi^*)''(\varphi'(t)) = \frac{1}{\varphi''(t)}. \tag{7.41}$$

Proof Under the assumptions of the lemma, φ' is strictly increasing and differentiable. So $(\varphi')^{-1} : (\varphi'(-\infty), \varphi'(\infty)) \to \mathbb{R}$ is well-defined, strictly increasing, differentiable, and $\varphi^*(\eta) = \eta(\varphi')^{-1}(\eta) - \varphi[(\varphi')^{-1}(\eta)]$

The lemma follows by differentiation of right-hand-side. □

Proof of Theorem 7.4. Let $\mathcal{I}_N := \mathcal{F}_N^*$.

Part 1: Since $\mathcal{F}_N$ is strictly convex and smooth, $\mathcal{F}_N'$ is strictly increasing and continuous. So $\mathcal{F}_N'[-1,1] = [\mathcal{F}_N'(-1), \mathcal{F}_N'(1)] \equiv [a_N^1, b_N^1]$, and for every $\eta \in [a_N^1, b_N^1]$, there is a unique $\xi \in [-1,1]$ s.t. $\mathcal{F}_N'(\xi) = \eta$. So $\mathrm{dom}(\mathcal{I}_N) \supset [a_N^1, b_N^1]$.

By Theorem 7.3 there is $C > 0$ such that $C^{-1} \le \mathcal{F}_N'' \le C$ on $[-1,1]$ for all $N \ge N_0$. Since $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$ and $\mathcal{F}_N'(\rho) = \mathcal{F}_N'(0) + \int_0^\rho \mathcal{F}_N''(\xi)\mathrm{d}\xi$, we have $b_N^1 \equiv \mathcal{F}_N'(1) \ge \dfrac{\mathbb{E}(S_N)}{V_N} + C^{-1}$, $a_N^1 \equiv \mathcal{F}_N'(-1) \le \dfrac{\mathbb{E}(S_N)}{V_N} - C^{-1}$. So $\mathrm{dom}(\mathcal{I}_N) \supseteq [a_N^1, b_N^1] \supseteq \left[\frac{\mathbb{E}(S_N)}{V_N} - C^{-1}, \frac{\mathbb{E}(S_N)}{V_N} + C^{-1}\right]$ for all $N \ge N_0$.

Part 2 follows from Lemma 7.23 and the strict convexity of $\mathcal{F}_N$ on $[-R, R]$.

Part 3: Let $J_N := \left[\frac{\mathbb{E}(S_N)}{V_N} - C^{-1}, \frac{\mathbb{E}(S_N)}{V_N} + C^{-1}\right]$. In part 1 we constructed the functions $\xi_N : J_N \to [-1,1]$ such that $\mathcal{F}_N'(\xi_N(\eta)) = \eta$.

Since $C^{-1} \le \mathcal{F}_N'' \le C$ on $[-1,1]$, $\xi_N'(\eta) = \frac{1}{\mathcal{F}_N''(\xi_N(\eta))} \in [C^{-1}, C]$ on J_N. This and the identity $\xi_N\left(\frac{\mathbb{E}(S_N)}{V_N}\right) = 0$ lead to

$$|\xi_N(\eta)| = |\xi_N(\eta) - \xi_N(\tfrac{\mathbb{E}(S_N)}{V_N})| \le C|\eta - \tfrac{\mathbb{E}(S_N)}{V_N}| \text{ for all } \eta \in J_N, N \ge N_0.$$

Fix $0 < \epsilon < 1$. By Theorem 7.3(4) there are $\delta, N_\epsilon > 0$ such that $\mathrm{e}^{-\epsilon} \le \mathcal{F}_N'' \le \mathrm{e}^{\epsilon}$ on $[-\delta, \delta]$ for all $N > N_\epsilon$. If $|\eta - \frac{\mathbb{E}(S_N)}{V_N}| < \delta/C$, then $|\xi_N(\eta)| < \delta$, and $\mathcal{F}_N''(\xi_N(\eta)) \in [\mathrm{e}^{-\epsilon}, \mathrm{e}^{\epsilon}]$.

Since $\mathcal{F}_N(0) = 0$ and $\mathcal{F}_N'(0) = \frac{\mathbb{E}(S_N)}{V_N}$, we have by (7.41) that $\mathcal{I}_N(\frac{\mathbb{E}(S_N)}{V_N}) = \mathcal{I}_N'(\frac{\mathbb{E}(S_N)}{V_N}) = 0$ and $\mathcal{I}_N''(\eta) = 1/\mathcal{F}_N''(\xi_N(\eta)) \in [\mathrm{e}^{-\epsilon}, \mathrm{e}^{\epsilon}]$. Writing

$$\mathcal{I}_N(\eta) = \mathcal{I}_N(\tfrac{\mathbb{E}(S_N)}{V_N}) + \int_{\frac{\mathbb{E}(S_N)}{V_N}}^{\eta} \left(\mathcal{I}_N'(\tfrac{\mathbb{E}(S_N)}{V_N}) + \int_{\frac{\mathbb{E}(S_N)}{V_N}}^{\alpha} \mathcal{I}_N''(\beta)\mathrm{d}\beta\right)\mathrm{d}\alpha,$$

we find that $\mathcal{I}_N(\eta) = \mathrm{e}^{\pm\epsilon}\frac{1}{2}(\eta - \frac{\mathbb{E}(S_N)}{V_N})^2$ for all η such that $|\eta - \frac{\mathbb{E}(S_N)}{V_N}| \le \delta/C$.

Part 4: If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\frac{z_N}{V_N} \in \left[\frac{\mathbb{E}(S_N)}{V_N} - \delta_N, \frac{\mathbb{E}(S_N)}{V_N} + \delta_N\right]$ with $\delta_N \to 0$. By part 3, $\mathcal{I}_N(\frac{z_N}{V_N}) \sim \frac{1}{2}\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)^2$, whence $V_N \mathcal{I}_N(\frac{z_N}{V_N}) \sim \frac{1}{2}\left(\frac{z_n - \mathbb{E}(S_N)}{\sqrt{V_N}}\right)^2$. □

Let $H_N(\eta)$ denote the Legendre transform of $P_N(\xi)/V_N$. We will compare $H_N(\eta)$ to $\mathcal{I}_N(\eta)$. This is needed to link the change of measure we performed in section §7.3.3 to the functions $\mathcal{I}_N$ which appear in the statement of the LLT for large deviations.

Lemma 7.24 *Suppose $R > 0$ and $V_N \neq 0$ for all $N \ge N_0$. Then for all $N \ge N_0$,*

(1) H_N is well-defined and real-analytic on $[a_N^R, b_N^R] = \left[\mathcal{F}_N'(-R), \mathcal{F}_N'(R)\right]$.
(2) For all R large enough, there exists $c > 0$ such that $H_N(\cdot)$ is well-defined and real-analytic on $\left(\frac{\mathbb{E}(S_N)}{V_N} - c, \frac{\mathbb{E}(S_N)}{V_N} + c\right)$, for all N large enough.

Proof Let $[\overline{a}_N^R, \overline{b}_N^R] := [\widetilde{a}_N^R, \widetilde{b}_N^R] + \frac{\mathbb{E}(S_N)}{V_N} \equiv [\frac{P_N'(-R)}{V_N}, \frac{P_N'(R)}{V_N}]$. Lemma 7.19 and its proof provide real-analytic maps $\xi_N : [\overline{a}_N^R, \overline{b}_N^R] \to \mathbb{R}$ such that $\frac{P_N'(\xi_N(\eta))}{V_N} = \eta$. So for all $R > 0$, H_N is real-analytic and well-defined on $[\overline{a}_N^R, \overline{b}_N^R]$. By (7.29), $[a_N^R, b_N^R] \equiv [\widehat{a}_N^R, \widehat{b}_N^R] + \dfrac{\mathbb{E}(S_N)}{V_N} \subset [\widetilde{a}_N^{R+1}, \widetilde{b}_N^{R+1}] + \dfrac{\mathbb{E}(S_N)}{V_N} = [\overline{a}_N^{R+1}, \overline{b}_N^{R+1}]$ and part 1 follows. Part 2 follows from Lemma 7.19(2) and (7.29). □

Lemma 7.25 *Suppose $R > 0$ and $V_N \neq 0$ for all $N \geq N_0$. Then for all $N > N_0$,*

(1) $\mathrm{dom}(\mathcal{I}_N) \cap \mathrm{dom}(H_N) \supset [a_N^R, b_N^R]$.

(2) There exists $C(R) > 0$ such that if $\frac{z}{V_N} \in [a_N^R, b_N^R]$ and $N \geq N_0$, then

(a) $\left|V_N \mathcal{I}_N\left(\frac{z}{V_N}\right) - V_N H_N\left(\frac{z}{V_N}\right)\right| \leq C(R)$; *(b)* $C(R)^{-1} \leq H_N''(\frac{z}{V_N}) \leq C(R)$.

(3) For every $\epsilon > 0$, $\exists \delta, N_\epsilon > 0$ such that if $N \geq N_\epsilon$ and $\left|\frac{z-\mathbb{E}(S_N)}{V_N}\right| < \delta$, then

(a) $\left|V_N \mathcal{I}_N\left(\frac{z}{V_N}\right) - V_N H_N\left(\frac{z}{V_N}\right)\right| \leq \epsilon$; *(b)* $\mathrm{e}^{-\epsilon} \leq H_N''\left(\frac{z}{V_N}\right) \leq \mathrm{e}^{\epsilon}$.

Proof Part 1 is a direct consequence of Lemma 7.24 and Theorem 7.4(1).

To prove the other parts of the lemma, we use the following consequence of (7.29) and the continuity of P_N' and $\mathcal{F}_N'$: For every $R > 0$, for all N large enough, for every $\eta \in [a_N^R, b_N^R]$, there exist $\xi_N^{(1)}, \xi_N^{(2)} \in [-(R+1), (R+1)]$ such that $\dfrac{P_N'(\xi_N^{(1)})}{V_N} = \eta$ and $\mathcal{F}_N'(\xi_N^{(2)}) = \eta$. Arguing as in the proof of Theorem 7.4(3), we can also find a constant $C(R)$ such that $|\xi_N^{(i)}| \leq C(R)|\eta - \frac{\mathbb{E}(S_N)}{V_N}|$.

It is a general fact that the Legendre transform of a C^1 convex function φ is equal on its domain to $\varphi^*(\eta) = \sup_\xi\{\xi\eta - \varphi(\xi)\}$. Thus for every $z \in [a_N^R V_N, b_N^R V_N]$,

$$V_N \mathcal{I}_N\left(\frac{z}{V_N}\right) = V_N \sup_\xi\left\{\xi\frac{z}{V_N} - \mathcal{F}_N(\xi)\right\} = V_N\left(\xi_N^{(2)}\frac{z}{V_N} - \mathcal{F}_N(\xi_N^{(2)})\right) \overset{Lm.7.18(4)}{\leq} V_N\left(\xi_N^{(2)}\frac{z}{V_N} - \frac{P_N(\xi_N^{(2)})}{V_N}\right) + \Delta_N(R+1)$$

$$\leq V_N \sup_\xi\left\{\xi\frac{z}{V_N} - \frac{P_N(\xi)}{V_N}\right\} + \Delta_N(R+1) \equiv V_N H_N\left(\frac{z}{V_N}\right) + \Delta_N(R+1).$$

So $V_N \mathcal{I}_N(\frac{z}{V_N}) - V_N H_N(\frac{z}{V_N}) \leq \Delta_N(R+1)$. Similarly, one can show that $V_N H_N(\frac{z}{V_N}) - V_N \mathcal{I}_N(\frac{z}{V_N}) \leq \Delta_N(R+1)$. Part (2a) now follows from Lemma 7.18(4).

If $z/V_N \in \left(\frac{\mathbb{E}(S_N)}{V_N} - \delta, \frac{\mathbb{E}(S_N)}{V_N} + \delta\right)$, then $|\xi_N^{(i)}| < C\delta$, and the same argument gives

$$\sup_{N \geq N_0} \sup_{\left|\frac{z-\mathbb{E}(S_N)}{V_N}\right| \leq \delta} |V_N \mathcal{I}_N(z/V_N) - V_N H_N(z/V_N)| \leq \sup_{N \geq N_0} \Delta_N(C\delta).$$

Part (3a) follows from Lemma 7.18(4).

By (7.41), $H_N''(\frac{z}{V_N}) = \frac{V_N}{P_N''(\xi)} = \frac{V_N}{\widetilde{V}_N^\xi} \cdot \frac{\widetilde{V}_N^\xi}{P_N''(\xi)}$, for the unique ξ s.t. $\frac{P_N'(\xi)}{V_N} = \frac{z}{V_N}$. Part (2b) now follows from Lemmas 7.18(3) and 7.21(1).

Part (3b) is because if $|\frac{z-\mathbb{E}(S_N)}{V_N}|$ is small, then $|\xi|$ is small, and $\frac{V_N}{\widetilde{V}_N^\xi}$ and $\frac{\widetilde{V}_N^\xi}{P_N''(\xi)}$ are close to one, by Lemmas 7.18(3) and 7.21(3). □

7.3.7 Proof of the Local Limit Theorem for Large Deviations

Proof of Theorem 7.8. We consider the non-lattice case, when $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$; the modifications needed for the lattice case are routine.

By Lemma 2.27, it is sufficient to prove the theorem under the additional assumption that the initial distribution of X is a point mass measure δ_x. Because of this reduction, we can assume that $\mathbb{P} = \mathbb{P}_x$ and $\mathbb{E} = \mathbb{E}_x$.

Since $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, f is not center-tight, and $V_N := \mathrm{Var}(S_N) \to \infty$ (see Corollary 3.7 and Theorem 3.8). There is no loss of generality in assuming that $V_N \neq 0$ for all N.

Let $[\widehat{a}_N^R, \widehat{b}_N^R] := [\mathcal{F}_N'(-R) - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}_N'(R) - \frac{\mathbb{E}(S_N)}{V_N}]$. By Theorem 7.4(1), for R large enough, $\bigcap_N [\widehat{a}_N^R, \widehat{b}_N^R]$ contains a non-empty interval. Fix R like that.

Suppose $\frac{z_N-\mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$, and let $h_n^{\xi}(\cdot) := h_n(\cdot,\xi)$, $p_n(\xi)$, and $P_N(\xi)$ be as in §§7.3.2, 7.3.3. Then $\exists \xi_N \in [-(R+1),(R+1)]$ as in Lemma 7.19: $P_N'(\xi_N) = z_N,\ \xi_N = O\left(\frac{z_N-\mathbb{E}(S_N)}{V_N}\right)$ and $\widetilde{\mathbb{E}}^{\xi_N}(S_N) = z_N + O(1)$.

Define a Markov array $\widetilde{\mathsf{X}} := \{\widetilde{X}_n^{(N)} : 1 \le n \le N+1\}$ with the state spaces of X (i.e. $\mathfrak{S}_n^{(N)} = \mathfrak{S}_n$), the initial distributions of X (i.e. $\pi^{(N)} = \delta_x$), but with the transition probabilities

$$\widetilde{\pi}_{n,n+1}^{(N)}(x,\mathrm{d}y) := \mathrm{e}^{\xi_N f_n(x,y)} \frac{h_{n+1}(y,\xi_N)}{\mathrm{e}^{p_n(\xi_N)} h_n(x,\xi_N)} \cdot \pi_{n,n+1}(x,\mathrm{d}y).$$

Let $\widetilde{\mathsf{f}} = \{f_n^{(N)} : 1 \le n \le N+1, N \in \mathbb{N}\}$ where $f_n^{(N)} := f_n$, and set

$$\widetilde{S}_N := f_1(\widetilde{X}_1^{(N)}, \widetilde{X}_2^{(N)}) + \cdots + f_N(\widetilde{X}_N^{(N)}, \widetilde{X}_{N+1}^{(N)}).$$

Recall that $\mathrm{e}^{\xi_N f_n}$, h_n, and $\mathrm{e}^{p_n(\xi_N)}$ are uniformly bounded away from zero and infinity, see Lemma 7.14. So $\widetilde{\pi}_{n,n+1}^{(N)}(x,\mathrm{d}y)$ differ from $\pi_{n,n+1}(x,\mathrm{d}y)$ by densities which are bounded away from zero and ∞, uniformly in N.

By Corollary 2.10, $\widetilde{\mathsf{X}}$ is uniformly elliptic, and the hexagon measures associated to $\widetilde{\mathsf{X}}$ and X on Hex (N,n) are related by Radon-Nikodym derivatives which are bounded away from zero and ∞ uniformly in $5 \le n \le N < \infty$.

Thus, $d_n^{(N)}(\widetilde{\mathsf{f}},\eta) \asymp d_n(\mathsf{f},\eta)$ and $u_n^{(N)}(\widetilde{\mathsf{f}}) \asymp u_n(\mathsf{f})$ for $5 \le n \le N < \infty$. So,

- $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$ and (X,f) have the same co-ranges and essential ranges. In particular, $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$ is irreducible and non-lattice.
- $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$ is stably hereditary (see Example 4.12 in §4.2.3).
- $\widetilde{V}_N := \mathrm{Var}(\widetilde{S}_N) \xrightarrow[N\to\infty]{} \infty$ (because $\widetilde{V}_N \asymp \sum_{n=3}^N u_n^2 \asymp V_N \to \infty$).

By the choice of ξ_N, $\dfrac{z_N - \mathbb{E}(\widetilde{S}_N)}{\sqrt{V_N}} = O\Big(\dfrac{1}{\sqrt{V_N}}\Big) \xrightarrow[N\to\infty]{} 0$. Therefore $\widetilde{S}_N$ satisfies the local limit theorem (Theorem 5.1):

$$\mathbb{P}(\widetilde{S}_N - z_N \in (a,b)) \sim |a-b|/\sqrt{2\pi \widetilde{V}_N^{\xi_N}}. \tag{7.42}$$

The task now is to translate (7.42) into an asymptotic for $\mathbb{P}(S_N - z_N \in (a,b))$. By construction, the initial distributions of $\widetilde{\mathsf{X}}$ are $\pi^{(N)} = \pi = \delta_x$, therefore

$$\begin{aligned}\mathbb{P}[S_N - z_N \in (a,b)] &= \mathbb{E}_x(1_{(a,b)}(S_N - z_N)) = \mathrm{e}^{P_N(\xi_N)-\xi_N z_N} \times \\ &\times \mathbb{E}_x\left(\mathrm{e}^{\xi_N S_N} \frac{h_{N+1}^{\xi_N}(X_{N+1}^{(N)})}{\mathrm{e}^{P_N(\xi_N)} h_1^{\xi_N}(x)} \cdot \frac{h_1^{\xi_N}(x)}{h_{N+1}^{\xi_N}(X_{N+1}^{(N)})} \cdot \mathrm{e}^{\xi_N(z_N - S_N)} 1_{(a,b)}(S_N - z_N)\right) \\ &= \mathrm{e}^{P_N(\xi_N)-\xi_N z_N} \times h_1^{\xi_N}(x) \times \widetilde{\mathbb{E}}_x^{\xi_N}\left(h_{N+1}^{\xi_N}(\widetilde{X}_{N+1}^{(N)})^{-1} \phi_{a,b}(\widetilde{S}_N - z_N)\right),\end{aligned} \tag{7.43}$$

where $\phi_{a,b}(t) := 1_{(a,b)}(t)\mathrm{e}^{-\xi_N t}$.

The first term simplifies as follows: By construction $\frac{P_N'(\xi_N)}{V_N} = \frac{z_N}{V_N}$, so

$$\mathrm{e}^{P_N(\xi_N)-\xi_N z_N} = \exp\left[-V_N\left(\xi_N \frac{z_N}{V_N} - \frac{P_N(\xi_N)}{V_N}\right)\right] = \mathrm{e}^{-V_N H_N\left(\frac{z_N}{V_N}\right)}, \tag{7.44}$$

where $H_N(\eta)$ is the Legendre transform of $P_N(\xi)/V_N$.

To simplify the third term $\widetilde{\mathbb{E}}_x^{\xi_N}(\cdots)$, we sandwich $\phi_{a,b}$ in $L^1(\mathbb{R})$ between continuous functions with compact support, and sandwich $1/h_{N+1}^{\xi_N}$ between finite linear combinations of indicators of sets $\mathfrak{A}_i^{(N+1)}$ such that $\widetilde{\mathbb{P}}^{\xi_N}(\widetilde{X}_{N+1}^{\xi_N} \in \mathfrak{A}_i)$ is bounded away from zero. Then we apply the Mixing LLT (Theorem 5.4) to $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$, and obtain

$$\widetilde{\mathbb{E}}_x^{\xi_N}\left(\frac{\phi_{a,b}(\widetilde{S}_N - z_N)}{h_{N+1}^{\xi_N}(\widetilde{X}_{N+1}^{(N)})}\right) \sim \frac{\widetilde{\mathbb{E}}^{\xi_N}\left(h_{N+1}^{\xi_N}(\widetilde{X}_{N+1}^{\xi_N})^{-1}\right)}{\sqrt{2\pi \widetilde{V}_N^{\xi_N}}} \int_a^b e^{-\xi_N t} dt. \tag{7.45}$$

Since ξ_N is bounded, $\widetilde{V}_N^{\xi_N} \sim P_N''(\xi_N)$ as $N \to \infty$ (Lemma 7.18(3)). Since $H_N(\eta)$ is the Legendre transform of $P_N(\xi)/V_N$, $H_N''(z_N/V_N) = V_N/P_N''(\xi_N) \sim V_N/\widetilde{V}_N^{\xi_N}$, see (7.41). This leads to

$$\widetilde{V}_N^{\xi_N} \sim \frac{V_N}{H_N''(\frac{z_N}{V_N})} \quad \text{as} \quad N \to \infty. \tag{7.46}$$

Substituting (7.44), (7.45), and (7.46) in (7.43), we obtain the following:

$$\mathbb{P}[S_N - z_N \in (a,b)] \sim \frac{e^{-V_N \mathcal{I}_N(\frac{z_N}{V_N})}}{\sqrt{2\pi V_N}} |a-b| \times \frac{1}{|a-b|} \int_a^b e^{-\xi_N t} dt \times$$
$$\times \underbrace{\left[e^{V_N \mathcal{I}_N(\frac{z_N}{V_N}) - V_N H_N(\frac{z_N}{V_N})} \sqrt{H_N''\left(\frac{z_N}{V_N}\right)}\,\right]}_{\widehat{\rho}_N} \times \underbrace{\left[h_1^{\xi_N}(x)\widetilde{\mathbb{E}}^{\xi_N}\left[h_{N+1}^{\xi_N}\left(\widetilde{X}_{N+1}^{\xi_N}\right)^{-1}\right]\right]}_{\overline{\rho}_N}.$$

Let $\eta_N := \frac{z_N - \mathbb{E}(S_N)}{V_N}$, then $\xi_N = \xi_N(\eta_N)$ where $\xi_N : [\widehat{a}_N^R, \widehat{b}_N^R] \to [-(R+1), (R+1)]$ is defined implicitly by $P_N'(\xi_N(\eta)) = \eta_N V_N + \mathbb{E}(S_N)$. Lemma 7.19 shows that $\xi_N(\cdot)$ is well-defined, and that it satisfies the statement of Theorem 7.8(4b).

There exists a constant $L = L(R)$ such that $|\eta_N| \le L(R)$. Indeed, $\eta_N \in [\widehat{a}_N^R, \widehat{b}_N^R]$ and $|\widehat{a}_N^R|, |\widehat{b}_N^R| \le |\mathcal{F}'(\pm R) - \mathcal{F}'(0)| \le R \sup_{[-R,R]} \mathcal{F}_N''$, which is uniformly bounded by Theorem 7.3(3).

Let $\widehat{\rho}_N(\eta) := e^{V_N \mathcal{I}_N\left(\eta + \frac{\mathbb{E}(S_N)}{V_N}\right) - V_N H_N\left(\eta + \frac{\mathbb{E}(S_N)}{V_N}\right)} \sqrt{H_N''\left(\eta + \frac{\mathbb{E}(S_N)}{V_N}\right)}$. This is well-defined for $\eta \in [\widehat{a}_N^R, \widehat{b}_N^R]$, and by Lemma 7.25, there is a constant C such that $C^{-1} \le \widehat{\rho}_N(\eta) \le C$ for all N and $\eta \in [\widehat{a}_N^R, \widehat{b}_N^R]$. In addition, for every $\epsilon > 0$ there are $\delta, N_\epsilon > 0$ such that $e^{-\epsilon} \le \widehat{\rho}_N(\eta) \le e^{\epsilon}$ for all $N > N_\epsilon$ and $|\eta| \le \delta$. In particular, if $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\widehat{\rho}_N(\frac{z_N - \mathbb{E}(S_N)}{V_N}) \xrightarrow[N\to\infty]{} 1$.

Let $\overline{\rho}_N(\eta) := h_1(x, \xi(\eta))\widetilde{\mathbb{E}}^{\xi(\eta)}\left(h_{N+1}(\widetilde{X}_{N+1}^{\xi(\eta)}, \xi(\eta))^{-1}\right)$. (There is no dependence on x, because x is fixed.) This function is well-defined for all $\eta \in [\widehat{a}_N^R, \widehat{b}_N^R]$, and by Lemma 7.14, there is a constant C such that

$$C^{-1} \le \overline{\rho}_N(\eta) \le C \text{ for all } N \text{ and } \eta \in [\widehat{a}_N^R, \widehat{b}_N^R].$$

By Lemma 7.16 and the obvious identity $h_n(\cdot, 0) \equiv 1$, $\|h_n^\xi - 1\|_\infty \xrightarrow[\xi\to 0]{} 0$ uniformly in n. Since $|\xi(\eta)| \le C|\eta|$, for every $\epsilon > 0$ there are $\delta, N_\epsilon > 0$ such that $e^{-\epsilon} \le \overline{\rho}_N(\eta) \le e^{\epsilon}$ for all $N > N_\epsilon$, and $|\eta| \le \delta$.

Setting $\rho_N := \widehat{\rho}_N \cdot \overline{\rho}_N$, we obtain the theorem in the non-lattice case. The modifications needed for the lattice case are routine, and are left to the reader. □

7.3.8 Rough Bounds in the Reducible Case

Proof of Theorem 7.9: We use the same argument we used in the previous section, except that we will use the rough bounds of §6.2.5 (which do not require irreducibility), instead of the LLT for local deviations (which does).

Let $\hbar = 100K + 1$ where $K = \text{ess sup}(\mathrm{f})$. Looking at (7.43) and using Theorem 6.11 and the assumption that $\frac{z_N}{V_N} \in [\mathcal{F}_N'(\varepsilon), \mathcal{F}_N'(R)]$, we get that there exists a constant $\overline{c} = \overline{c}(R)$ and $\xi_N := \xi_N\left(\frac{z_N}{V_N}\right) \in [\varepsilon, R+1]$ such that for all N large enough,

$$\mathbb{P}(S_N - z_N \in [0, \hbar]) \geq \frac{\overline{c}\mathrm{e}^{-\xi_N \hbar}\hbar}{\sqrt{V_N}}\mathrm{e}^{-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)} \tag{7.47}$$

(Theorem 6.11 is applicable because $\hbar > 2\delta(\mathsf{f})$ due to Corollary 4.6).)

Since $\mathbb{P}(S_N \geq z_n) \geq \mathbb{P}(S_N - z_N \in [0, \hbar])$ the lower bound follows.

Likewise, applying Theorem 6.12, we conclude that there is a constant $C^* = C^*(R)$ such that for all N large enough we have, uniformly in $j \in \mathbb{N} \cup \{0\}$, $\mathbb{P}(S_N - z_N \in [\hbar j, \hbar(j+1)]) \leq \frac{C^* \mathrm{e}^{-\xi_N \hbar j}}{\sqrt{V_N}}\mathrm{e}^{-V_N \mathcal{I}_N\left(\frac{z_N}{V_N}\right)}$. Summing over j we obtain the upper bound. □

7.4 Large Deviations Thresholds

7.4.1 The Large Deviations Threshold Theorem

In this section we look closer into the conditions imposed on $\frac{z_N - \mathbb{E}(S_N)}{V_N}$ in the LLT for Large Deviations (Theorem 7.8), and ask what happens when they fail.

Let f be an additive functional on a uniformly elliptic Markov chain X. We assume throughout that f is non-center-tight, irreducible, and a.s. uniformly bounded. Without loss of generality, $G := G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$ or $\mathbb{Z}$. Our main result is:

Theorem 7.26 *There exist $\mathfrak{c}_- < 0 < \mathfrak{c}_+$ as follows. Suppose z_N is a sequence of numbers such that $\mathbb{P}[S_N - z_N \in G] = 1$ for all N, and $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to z$.*

(1) If $z \in (\mathfrak{c}_-, \mathfrak{c}_+)$, then z_N satisfies the assumptions and conclusions of the LLT for large deviations (Theorem 7.8), and $\exists$ a bounded sequence I_N s.t. for all intervals (a, b) which intersect G, $\mathbb{P}[S_N - z_N \in (a, b)] \asymp \dfrac{\mathrm{e}^{-V_N I_N}}{\sqrt{2\pi V_N}}$.

(2) If $z \notin [\mathfrak{c}_-, \mathfrak{c}_+]$, then the conditions of Theorem 7.8 fail, and $\exists N_k \uparrow \infty$ such that $\mathbb{P}[S_{N_k} - z_{N_k} \in (a, b)] \to 0$ faster than $\mathrm{e}^{-V_{N_k} I}$, for each $I > 0$.

The theorem does not extend to the case when $z = \mathfrak{c}_-$ or $\mathfrak{c}_+$, see Example 7.37. The numbers $\mathfrak{c}_-, \mathfrak{c}_+$ are called the **large deviations thresholds** of (X, f).

Corollary 7.27 *Let $\mathfrak{I}_a(z) = \limsup_{N\to\infty} V_N^{-1} |\log \mathbb{P}(S_N - \mathbb{E}(S_N) \in zV_N + [-a, a])|$. If a is bigger than the graininess constant of (X, f), then $\mathfrak{c}_+ = \sup\{z : \mathfrak{I}_a(z) < \infty\}$ and $\mathfrak{c}_- = \inf\{z : \mathfrak{I}_a(z) < \infty\}$.*

Thus, if $z \notin [\mathfrak{c}_-, \mathfrak{c}_+]$, then $\mathbb{P}[S_N - z_N \in (a, b)]$ decays "too fast" along some sub-sequence. Here is a simple scenario when this happens. Let

$$\mathfrak{r}_- := \limsup_{N\to\infty} \frac{\text{ess inf}(S_N - \mathbb{E}(S_N))}{V_N}, \ \mathfrak{r}_+ := \liminf_{N\to\infty} \frac{\text{ess sup}(S_N - \mathbb{E}(S_N))}{V_N}.$$

If $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to z$ and $z \notin [\mathfrak{r}_-, \mathfrak{r}_+]$, then $\exists N_k \to \infty$ such that $\mathbb{P}[S_{N_k} - z_{N_k} \in (a, b)]$ is eventually zero, and we must be in case (2) of Theorem 7.26. We call $\mathfrak{r}_\pm$ the **positivity thresholds** of (X, f). Clearly, $(\mathfrak{c}_-, \mathfrak{c}_+) \subset (\mathfrak{r}_-, \mathfrak{r}_+)$.

If $\mathfrak{c}_\pm = \mathfrak{r}_\pm$ then we say that (X, f) has a **full large deviations regime**. Otherwise, we say that the large deviations regime is **partial**. For examples, see §7.4.4. Note that even in the full regime case, our results do not apply to $z = \mathfrak{r}^\pm$.

In the partial case, $\mathbb{P}[S_N - z_N \in (a, b)]$ decays faster than expected when $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to z \in (\mathfrak{r}_-, \mathfrak{r}_+) \setminus (\mathfrak{c}_-, \mathfrak{c}_+)$, but the precise asymptotic behavior remains unknown. We are not aware of general results in this direction, even for sums of independent, non-identically distributed, random variables.

7.4.2 Admissible Sequences

Let f be a non-center-tight and a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain X, and recall that $\mathcal{F}_N(\xi) = (1/V_N)\log\mathbb{E}(e^{\xi S_N})$.

Theorem 7.8 is stated under the assumption that for some $R > 0$, for all N large,

$$\frac{z_N - \mathbb{E}(S_N)}{V_N} \in \left[\mathcal{F}_N'(-R) - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}_N'(R) - \frac{\mathbb{E}(S_N)}{V_N}\right]. \tag{7.48}$$

Sequences $\{z_N\}$ satisfying (7.48) for a specific R are called R**-admissible**, and sequences which satisfy (7.48) for some $R > 0$ are called **admissible**. Admissible sequences exist, see Theorem 7.4(1).

Why do we need admissibility? The proof of the LLT for large deviations uses a change of measure $\widetilde{\mathsf{X}}$, given by (7.3). The change of measure depends on parameters ξ_N, which are calibrated so that $\mathbb{E}(\widetilde{S}_N) = z_N + O(1)$. These parameters are roots of the equation $P_N'(\xi_N) = z_N$, where P_N are the functions from (7.24), and the admissibility condition is necessary and sufficient for these roots to exist:

Lemma 7.28 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*, such that* $V_N \to \infty$*. The following are equivalent:*

(1) $\{z_N\}$ *is admissible; (2) For some* $R > 0$*, for all* N *large enough,* $\exists \xi_N \in [-R, R]$ *s.t.* $P_N'(\xi_N) = z_N$*;*

(3) For some $R' > 0$*, for all* N *large enough,* $\exists \xi_N \in [-R', R']$ *s.t.* $\mathcal{F}_N'(\xi_N) = \frac{z_N}{V_N}$*.*

Proof (1)⇒(2) is Lemma 7.19(1).

Assume (2). P_N is strictly convex, therefore if $\exists \xi_N \in [-R, R]$ such that $P_N'(\xi_N) = z_N$, then $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in \left[\frac{P_N'(-R) - \mathbb{E}(S_N)}{V_N}, \frac{P_N'(R) - \mathbb{E}(S_N)}{V_N}\right] =: [\widehat{a}_N^R, \widehat{b}_N^R]$. By (7.29), $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^{R+1}, \widehat{b}_N^{R+1}] = \left[\frac{\mathcal{F}_N'(-R-1) - \mathbb{E}(S_N)}{V_N}, \frac{\mathcal{F}_N'(R+1) - \mathbb{E}(S_N)}{V_N}\right]$. Since $\mathcal{F}_N'$ is continuous, (3) follows with $R' := R + 1$. So (2)⇒(3).

Assume (3). $\mathcal{F}_N$ is a smooth convex function, therefore $\mathcal{F}_N'$ is continuous and increasing, whence $[\mathcal{F}_N'(-R'), \mathcal{F}_N'(R')] = \mathcal{F}_N'([-R', R'])$. By (3), for all $N \gg 1$, $\frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^{R'}, \widehat{b}_N^{R'}] := [\mathcal{F}_N'(-R') - \frac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}_N'(R') - \frac{\mathbb{E}(S_N)}{V_N}]$, and $\{z_N\}$ is R'-admissible. So (3)⇒(1). □

Lemma 7.29 *Let* f *be an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain, and assume that* $V_N \to \infty$*. Then* $\forall R > 0\ \exists \varepsilon = \varepsilon(R) > 0$ *such that if* $\{z_N\}$ *is* R*-admissible, and* $|\overline{z}_N - z_N| \le \varepsilon V_N$*, then* $\{\overline{z}_N\}$ *is* $(R+1)$*-admissible.*

Proof It is sufficient to prove the lemma under the assumption that $\mathbb{E}(S_N) = 0$. Suppose $\{z_N\}$ is R-admissible, and choose $\xi_N \in [-R, R]$ such that $\mathcal{F}_N'(\xi_N) = \frac{z_N}{V_N}$.

$\mathcal{F}_N$ are uniformly strictly convex on $[-(R+1), (R+1)]$ (Theorem 7.3), so there exists an $\varepsilon > 0$ such that $\mathcal{F}_N'(R+1) \ge \mathcal{F}_N'(R) + \varepsilon$ and $F_N'(-(R+1)) \le \mathcal{F}_N'(-R) - \varepsilon$ for all N. So, if $|\overline{z}_N - z_N| \le \varepsilon V_N$, then $\overline{z}_N / V_N \in [\mathcal{F}_N'(-R-1), \mathcal{F}_N'(R+1)]$. □

The following theorem characterizes the admissible sequences probabilistically. Recall the definition of the graininess constant $\delta(\mathsf{f})$ from (6.6).

Theorem 7.30 *Let* f *be a non center-tight a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*. The following are equivalent:*

(a) $\{z_N\}$ *is admissible.*

(b) $\exists \varepsilon > 0, \eta > 0$ *such that for every sequence* $\{\overline{z}_N\}$ *with* $|\overline{z}_N - z_N| \le \varepsilon V_N$ *and for every* a, b *such that* $|a|, |b| \le 10\delta(\mathsf{f}) + 1$ *and* $b - a > 2\delta(\mathsf{f})$*, for all* N *large enough,* $\mathbb{P}[S_N - \overline{z}_N \in (a, b)] \ge \eta^{V_N}$*.*

(c) $\exists \varepsilon > 0, \eta > 0$ *such that for all* N *large enough,* $\mathbb{P}(S_N \ge z_N + \varepsilon V_N) \ge \eta^{V_N}$ *and* $\mathbb{P}(S_N \le z_N - \varepsilon V_N) \ge \eta^{V_N}$*.*

Proof Note that under the assumptions of the theorem, $V_N \to \infty$ (Corollary 3.7).

(a)⇒(b): Let $\varepsilon := \varepsilon(R)$ be the constant from Lemma 7.29. Suppose $\{z_N\}$ is R-admissible, then $\overline{z}_N$ is $(R+1)$-admissible.

By Lemma 7.28, there is a bounded sequence of real numbers ξ_N such that $P'_N(\xi_N) = \overline{z}_N$. Consider the decomposition (7.43). Integrating over $x \in \mathfrak{S}_1$ and taking note of the uniform bounds for $h_k^{\xi_N}$ in Lemma 7.14, we obtain $\mathbb{P}[S_N - \overline{z}_N \in (a,b)] \geq \text{const.}\mathrm{e}^{P_N(\xi_N)-\xi_N\overline{z}_N}\widetilde{\mathbb{P}}[\widetilde{S}_N - \overline{z}_N \in (a,b)]$, where $\widetilde{\mathbb{P}}$ and $\widetilde{S}_N$ correspond to the change of measure (7.3).

The term $\mathrm{e}^{P_N(\xi_N)-\xi_N\overline{z}_N}$ is bounded below by a constant times $\mathrm{e}^{-V_N\mathcal{I}_N(\overline{z}_N/V_N)}$, see (7.44) and Lemma 7.25(2). By the choice of ξ_N, $\widetilde{\mathbb{E}}(\widetilde{S}_N) = \overline{z}_N + O(1)$, so the asymptotics of $\widetilde{\mathbb{P}}[\widetilde{S}_N - \overline{z}_N \in (a,b)]$ is given by Theorem 6.11, provided $|a-b| > 2\delta(\mathsf{f})$. In this case, $\mathbb{P}[S_N - \overline{z}_N \in (a,b)] \geq C\dfrac{\mathrm{e}^{-V_N\mathcal{I}_N(\frac{\overline{z}_N}{V_N})}}{\sqrt{2\pi V_N}}$ for all N large enough.

To finish the proof of (b), it is sufficient to show that $\mathcal{I}_N(\frac{\overline{z}_N}{V_N})$ is bounded from above. First we note that since $\{\overline{z}_N\}$ is $(R+1)$-admissible,

$$\frac{\overline{z}_N}{V_N} \in [a_N^{R+1}, b_N^{R+1}] \quad \text{for all } N. \tag{7.49}$$

By Theorem 7.4(2) and (3), $\mathcal{I}_N(\frac{\mathbb{E}(S_N)}{V_N}) = \mathcal{I}'_N(\frac{\mathbb{E}(S_N)}{V_N}) = 0$, and there exists $\rho > 0$ such that $0 \leq \mathcal{I}''_N < \rho$ on $[a_N^{R+1}, b_N^{R+1}]$ for all N. As $\frac{\mathbb{E}(S_N)}{V_N} = \mathcal{F}'_N(0) \in [a_N^{R+1}, b_N^{R+1}]$,

$$\left|\mathcal{I}_N\left(\frac{\overline{z}_N}{V_N}\right)\right| = \left|\mathcal{I}_N\left(\frac{\overline{z}_N}{V_N}\right) - \mathcal{I}_N\left(\frac{\mathbb{E}(S_N)}{V_N}\right)\right| \leq \frac{1}{2}\rho^2\left(\frac{\overline{z}_N - \mathbb{E}(S_N)}{V_N}\right)^2.$$

Equation (7.49) and Theorem 7.3 (1) and (3) tell us that

$$\left|\frac{\overline{z}_N - \mathbb{E}(S_N)}{V_N}\right| \leq C := \sup_N \max_{|\xi|\leq R+1}\left|\mathcal{F}'_N(\xi) - \frac{\mathbb{E}(S_N)}{V_N}\right| < \infty.$$

So $\mathcal{I}_N(\overline{z}_N/V_N) \leq \frac{1}{2}\rho^2C^2$, and part (b) follows.

(b)⇒(c): The bound $\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^{V_N}$ follows from part (b) with $\overline{z}_N = z_N + \varepsilon V_N$, $a = 0$, and $b = 2\delta(\mathsf{f}) + 1$. The lower bound is similar.

(c)⇒(a): If (c) holds, then $\mathbb{E}(\mathrm{e}^{RS_N}) \geq \mathbb{E}(\mathrm{e}^{RS_N}1_{[S_N\geq z_N+\varepsilon V_N]}) \geq \eta^{V_N}\mathrm{e}^{R(z_N+\varepsilon V_N)}$, whence $\mathcal{F}_N(R) \geq \log\eta + R(\frac{z_N}{V_N}+\varepsilon)$. If $R > 2\varepsilon^{-1}|\log\eta|$, then $\mathcal{F}_N(R) \geq R(\frac{z_N}{V_N}+\frac{\varepsilon}{2})$. Since $\mathcal{F}_N(0) = 0$ and $\mathcal{F}'_N$ is increasing, $\mathcal{F}'_N(R) > \frac{z_N}{V_N}$. Similarly one shows that $\mathcal{F}'_N(-R) < \frac{z_N}{V_N}$. So $\{z_N\}$ is admissible. □

Example 7.31 The following example shows that condition (c) with $\varepsilon = 0$ is *not* equivalent to admissibility.

Let $S_N = \sum_{n=1}^N X_n$ where X_n are iid random variables supported on $[\alpha,\beta]$ and such that X has an atom on the right edge: $\mathbb{P}(X=\beta) = \gamma > 0$. Then $\mathbb{P}[S_N \geq \beta N] = \mathbb{P}[S_N = \beta N] = \gamma^N$ while $\mathbb{P}[S_N \geq \beta N + 1] = 0$. Thus $\{\beta N\}$ is not admissible.

We note that the notion of R-admissibility depends on the initial distribution π_0 since $\mathbb{E}(S_N)$ depends on π_0. However, the notion of admissibility is independent of the initial distribution as the next result shows.

Lemma 7.32 *Let* f *be a uniformly bounded non center-tight additive functional of a uniformly elliptic Markov chain. If* z_N *is admissible with respect to the initial distribution* π_0*, then it is admissible with respect to any other initial distribution* $\widetilde{\pi}_0$.

Proof Let $\widetilde{\mathsf{X}}$ denote the Markov chain obtained from X by changing the initial distribution to $\widetilde{\pi}_0$. Objects associated with $(\widetilde{\mathsf{X}},\mathsf{f})$ will be decorated with a tilde.

Suppose z_N is R-admissible for (X,f). Then for all N large enough $\dfrac{z_N - \mathbb{E}(S_N)}{V_N} \in \left[\mathcal{F}'_N(-R) - \dfrac{\mathbb{E}(S_N)}{V_N}, \mathcal{F}'_N(R) + \dfrac{\mathbb{E}(S_N)}{V_N}\right]$.

By (7.29), $\dfrac{z_N - \mathbb{E}(S_N)}{V_N} \in \left[\dfrac{P'_N(-R-1) - \mathbb{E}(S_N)}{V_N}, \dfrac{P'_N(R+1) - \mathbb{E}(S_N)}{V_N}\right]$.

By Lemmas 7.18(3) and 7.21(3), there is a constant c_1 which only depends on R so that $\frac{P''_N}{V_N} > c_1$ on $[-R-3, R+3]$ for all N large enough. Necessarily, $\frac{P'_N(R+2)}{V_N} - \frac{P'_N(R+1)}{V_N} \geq c_1$ and $\frac{P'_N(-R-2)}{V_N} - \frac{P'_N(-R-1)}{V_N} \leq -c_1$. This leads to:

$$\frac{z_N - \mathbb{E}(S_N)}{V_N} \in \left[\frac{P'_N(-R-2) + c_1 V_N - \mathbb{E}(S_N)}{V_N}, \frac{P'_N(R+2) - c_1 V_N - \mathbb{E}(S_N)}{V_N}\right].$$

Applying the increasing map $t \mapsto (V_N/\widetilde{V}_N)\left(t + \frac{\mathbb{E}(S_N) - \widetilde{\mathbb{E}}(S_N)}{V_N}\right)$, we arrive at

$$\frac{z_N - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N} \in \left[\frac{P'_N(-R-2) + c_1 V_N - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}, \frac{P'_N(R+2) - c_1 V_N - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}\right].$$

Looking at (7.24), we see that $P'_N(\xi) - \widetilde{P}'_N(\xi) = \xi(\mathbb{E}(S_N) - \widetilde{\mathbb{E}}(S_N))$, (the function $\overline{P}_N(\xi)$ does not depend on the initial distribution, is the same for (X, f) and $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$, and cancels out). Next, Lemma 7.17(1) tells us that $\left|\mathbb{E}(S_N) - \widetilde{\mathbb{E}}(S_N)\right| \leq c_2$ for some constant c_2. Thus, $\exists c_3 = c_3(R)$ so that

$$|P'_N(R+2) - \widetilde{P}'_N(R+2)| < c_3, \text{ and } |P'_N(-R-2) - \widetilde{P}'_N(-R-2)| < c_3.$$

(X, f) is not center-tight, therefore $V_N \to \infty$. If N is so large that $c_1 V_N > c_3$, then

$$\frac{z_N - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N} \in \left[\frac{\widetilde{P}'_N(-R-2) - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}, \frac{\widetilde{P}'_N(R+2) - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}\right].$$

By (7.29), $\frac{z_N - \widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N} \in \left[\widetilde{\mathcal{F}}'_N(-R-3) - \frac{\widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}, \widetilde{\mathcal{F}}'_N(R+3) - \frac{\widetilde{\mathbb{E}}(S_N)}{\widetilde{V}_N}\right]$, whence the $(\widetilde{\mathsf{X}}, \mathsf{f})$-admissibility of z_N. □

We say that (X, f) and $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ are related by a change of measure *with bounded weights* if $\mathsf{X}, \widetilde{\mathsf{X}}$ have the same state spaces, $f_n \equiv \widetilde{f}_n$ for all n, and if for some $\overline{\varepsilon}$, the initial distributions and transition probabilities of X and $\widetilde{\mathsf{X}}$ are equivalent and related by $\overline{\varepsilon} \leq \frac{d\widetilde{\pi}_n(x,dy)}{d\pi_n(x,dy)} \leq \overline{\varepsilon}^{-1}$ for all n.

Lemma 7.33 *Suppose* f *is an a.s. uniformly bounded additive functional on a uniformly elliptic Markov chain* X*. If* (X, f) *and* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ *are related by the change of measure with bounded weights, and* $V_N \geq cN$ *for some* $c > 0$*, then* $\{z_N\}$ *is* (X, f)*-admissible iff* $\{z_N\}$ *is* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$*-admissible.*

Proof Since admissibility does not depend on the initial distribution we may suppose without the loss of generality that $\widetilde{\pi} = \pi$.

Since X is uniformly elliptic, $\widetilde{\mathsf{X}}$ is uniformly elliptic. By the exponential mixing bounds for uniformly elliptic chains both $\widetilde{V}_N := \text{Var}[S_N(\widetilde{\mathsf{f}})]$ and $V_N := \text{Var}[S_N(\mathsf{f})]$ are $O(N)$. Without loss of generality, $cN \leq V_N \leq c^{-1}N$.

Under the assumptions of the lemma, the structure constants of (X, f) are equal to the structure constants of $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ up to bounded multiplicative error. By Theorem 3.7, $\widetilde{V}_N \asymp V_N$ as $N \to \infty$. So $\exists \widetilde{c} > 0$ such that $\widetilde{c}N \leq \widetilde{V}_N \leq \widetilde{c}^{-1}N$.

Let $\{z_N\}$ be (X, f)-admissible. Then there are $\varepsilon > 0$ and $\eta > 0$ such that for all N large enough, $\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^N$, $\quad \mathbb{P}[S_N \leq z_N - \varepsilon V_N] \geq \eta^N$.

It follows that $\widetilde{\mathbb{P}}[S_N(\widetilde{\mathsf{f}}) \geq z_N + \widetilde{\varepsilon}\widetilde{V}_N] \geq \widetilde{\eta}^N$, $\quad \widetilde{\mathbb{P}}[S_N(\widetilde{\mathsf{f}}) \leq z_N - \widetilde{\varepsilon}\widetilde{V}_N] \geq \widetilde{\eta}^N$ where $\widetilde{\eta} = \eta\overline{\varepsilon}$ and $\widetilde{\varepsilon} := \widetilde{c}c\varepsilon$. Hence $\{z_N\}$ is $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$-admissible. □

Lemma 7.34 *Let* f *and* $\widetilde{\mathsf{f}}$ *be two a.s. uniformly bounded additive functionals on the same uniformly elliptic Markov chain. Suppose* $V_N := \text{Var}[S_N(\mathsf{f})] \to \infty$ *and*

$$\lim_{N\to\infty} \frac{\|S_N(\widetilde{\mathsf{f}}) - S_N(\mathsf{f})\|_\infty}{V_N} = 0. \tag{7.50}$$

Then $\{z_N\}$ *is* f*-admissible iff* $\{z_N\}$ *is* $\widetilde{\mathsf{f}}$*-admissible.*

Proof We write $S_N = S_N(\mathsf{f})$, and $\widetilde{S}_N = S_N(\mathsf{f})$. By the assumptions of the lemma, $\widetilde{V}_N := \text{Var}(\widetilde{S}_N) \sim V_N$ as $N \to \infty$.

Let $\{z_N\}$ be f-admissible. By Theorem 7.30(c), there are $\varepsilon > 0, \eta > 0$ such that $\mathbb{P}[S_N \geq z_N + \varepsilon V_N] \geq \eta^N$, $\mathbb{P}[S_N \leq z_N - \varepsilon V_N] \geq \eta^N$. It now follows from (7.50) that for large N, $\widetilde{\mathbb{P}}\left[\widetilde{S}_N \geq z_N + \frac{\varepsilon}{2}\widetilde{V}_N\right] \geq \eta^N$ and $\widetilde{\mathbb{P}}\left[\widetilde{S}_N \leq z_N - \frac{\varepsilon}{2}\widetilde{V}_N\right] \geq \eta^N$. Hence $\{z_N\}$ is $\widetilde{\mathsf{f}}$-admissible. □

7.4.3 Proof of the Large Deviations Threshold Theorem

Let f be an a.s. uniformly bounded and non-center-tight irreducible additive functional on a uniformly elliptic Markov chain X. Then $V_N \to \infty$, and the algebraic range is $\mathbb{R}$ or $t\mathbb{Z}$, with $t > 0$. It is sufficient to consider the cases $\mathbb{R}$ and $\mathbb{Z}$.

We call $z \in \mathbb{R}$ **reachable**, if at least one of the following conditions holds:

(1) The sequence $z_N:=\mathbb{E}(S_N)+V_N z$ is admissible. (2) $\exists$ an admissible sequence $\{z_N\}$ such that $\frac{z_N-\mathbb{E}(S_N)}{V_N} \to z$.

(3) Every sequence $\{z_N\}$ such that $\frac{z_N-\mathbb{E}(S_N)}{V_N} \to z$ is admissible.

The conditions are equivalent, by Lemma 7.29. If $z_N:=\mathbb{E}(S_N)+V_N z$ is R-admissible, we say that z is ***R*-reachable**.

We denote the set of R–reachable points by C_R and the set of reachable points by C. Since $\mathcal{F}_N'$ is monotone increasing, C_R and C are connected, and by Theorem 7.4(1), C contains a non-empty neighborhood of the origin. Therefore $\mathrm{int}(C) = (\mathfrak{c}_-', \mathfrak{c}_+')$ for some $\mathfrak{c}_-' < 0 < \mathfrak{c}_+'$. The plan is to show that Theorem 7.26 holds with $\mathfrak{c}_\pm := \mathfrak{c}_\pm'$.

Step 1. *Let* $\Im_a(z) = \limsup\limits_{N\to\infty} \dfrac{1}{V_N} |\log \mathbb{P}(S_N - \mathbb{E}(S_N)\in zV_N + [-a, a])|$ *with* $a \geq 3\delta(\mathsf{f})$ *(see* (6.6)*), then* $\mathfrak{c}_+' = \sup\{z : \Im_a(z) < \infty\}$ *and* $\mathfrak{c}_-' = \inf\{z : \Im_a(z) < \infty\}$.

Proof. If $z \in (\mathfrak{c}_-', \mathfrak{c}_+')$ then $z \in C$, $z_N := zV_N + \mathbb{E}(S_N)$ is admissible, and $\Im_a(z) < \infty$ by Theorem 7.30(b). So

$$\mathfrak{c}_+' \leq \sup\{z : \Im_a(z) < \infty\} =: S.$$

In particular, $S > 0$. We claim that S is the limit of $\overline{z}$ with the following property:

$$\exists 0 < \epsilon < \overline{z} \text{ such that } \Im_a(\overline{z} + 2\epsilon) < \infty. \tag{7.51}$$

If $S = \infty$, this is clear, and if $S < \infty$, then any $\dfrac{1}{2}\sup\{z : \Im_a(z) < \infty\} < \overline{z} < \sup\{z : \Im_a(z) < \infty\}$ satisfies (7.51).

Fix $\overline{z}$ as in (7.51). Necessarily $\exists \eta \in (0, 1)$ such that for all N large enough

$$\mathbb{P}[S_N - \mathbb{E}(S_N) \geq (\overline{z} + \epsilon)V_N] \geq \mathbb{P}(S_N - \mathbb{E}(S_N) \in (\overline{z} + 2\epsilon)V_N + [-a, a]) \geq \eta^{V_N};$$

$$\mathbb{P}[S_N - \mathbb{E}(S_N) \leq (\overline{z} - \epsilon)V_N] \geq \mathbb{P}[S_N - \mathbb{E}(S_N) \leq 0] \overset{!}{=} \frac{1}{2} + o(1) \geq \eta^{V_N}, \text{by the CLT.}$$

By Theorem 7.30(c), $\overline{z}V_N + \mathbb{E}(S_N)$ is admissible, and $\overline{z}$ is reachable.

We just proved that S is the supremum of reachable $\overline{z}$. It follows that $\mathfrak{c}_+' = S \equiv \sup\{z : \Im_a(z) < \infty\}$. The formula for $\mathfrak{c}_-'$ follows by considering $(\mathsf{X}, -\mathsf{f})$.

Step 2. *Theorem 7.26 holds with* $\mathfrak{c}_\pm := \mathfrak{c}_\pm'$.

Proof. Suppose first that $G_{alg}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$.

If $z \in (\mathfrak{c}_-', \mathfrak{c}_+')$, then $z_N := \mathbb{E}(S_N) + zV_N$ is R-admissible for some $R > 0$, and the conditions and conclusions of Theorem 7.8(2) and (4) are satisfied.

In particular, for some $R > 0$, $\frac{z_N-\mathbb{E}(S_N)}{V_N} \in [\widehat{a}_N^R, \widehat{b}_N^R]$ and $\log \rho_N(\frac{z_N-\mathbb{E}(S_N)}{V_N})$ and $\xi_N(\frac{z_N-\mathbb{E}(S_N)}{V_N})$ are uniformly bounded. Therefore $\mathbb{P}[S_N - z_N \in (a, b)] \asymp \dfrac{\mathrm{e}^{-V_N I_N}}{\sqrt{2\pi V_N}}$, with $I_N := \mathcal{I}_N\left(\dfrac{z_N}{V_N}\right)$.

By assumption, $\frac{z_N}{V_N} \in [a_N^R, b_N^R]$. By Theorem 7.4(2), there is a constant $\rho = \rho(R)$ such that for all N, $0 < \mathcal{I}_N'' \leq \rho$ on $[a_N^R, b_N^R]$. In addition, Theorem 7.4(3) clearly implies that $\mathcal{I}_N\left(\frac{\mathbb{E}(S_N)}{V_N}\right) = \mathcal{I}_N'\left(\frac{\mathbb{E}(S_N)}{V_N}\right) = 0$. Thus $0 \leq I_N \leq \frac{1}{2}\rho(\frac{z_N-\mathbb{E}(S_N)}{V_N})^2 \to \frac{1}{2}\rho z^2$, proving that I_N is bounded.

Now take $z \notin [\mathfrak{c}'_-, \mathfrak{c}'_+]$. By step 1, $\mathfrak{I}_a(z)=\infty$ for all $a>3\delta(\mathsf{X},\mathsf{f})$. Equivalently, $\limsup\limits_{N\to\infty} \frac{1}{V_N} |\log \mathbb{P}(S_N - z_N \in [-a,a])| = \infty$. It follows that $\exists N_k \uparrow \infty$ such that $\mathbb{P}(S_{N_k} - z_{N_k} \in [-a,a])$ decays to zero faster that $e^{-V_{N_k} I}$ for all $I > 0$.

In summary, if $G_{alg}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, then Theorem 7.26 holds with $\mathfrak{c}_\pm := \mathfrak{c}'_\pm$. The case when $G_{alg}(\mathsf{X},\mathsf{f}) = \mathbb{Z}$ is handled in the same way, and is left to the reader. □

7.4.4 Examples

Recall that $(\mathfrak{c}_-, \mathfrak{c}_+) \subset (\mathfrak{r}_-, \mathfrak{r}_+)$. We give examples of equality ("full large deviation regime") and strict inclusion ("partial large deviations regime"):

Example 7.35 (Equality) Let $S_N = X_1 + \cdots + X_N$ where X_n are bounded iid random variables with law X, expectation zero, and non-zero variance. In this case $\mathcal{F}_N(\xi) = \log \mathbb{E}(e^{\xi X})/\mathrm{Var}(X)$ for all N (Example 7.2), and it is not difficult to see that $\mathfrak{c}_- = \dfrac{\text{ess inf}(X)}{\mathrm{Var}(X)} = \mathfrak{r}_-$, $\quad \mathfrak{c}_+ = \dfrac{\text{ess sup}(X)}{\mathrm{Var}(X)} = \mathfrak{r}_+$. So we have full large deviation regime. See also Theorem 8.7.

Example 7.36 (Strict Inclusion) Let $X_n = (Y_n, Z_n)$ where $\{Y_n\}$, $\{Z_n\}$ are two independent sequences of iid random variables uniformly distributed on $[0,1]$. Fix a sequence of numbers $0 < p_n < 1$ such that $p_n \to 0$, and let $f_n(X_n) = \begin{cases} Z_n & \text{if } Y_n > p_n \\ 2 & \text{if } Y_n \le p_n. \end{cases}$ Then $f_n(X_n)$ are independent, but not identically distributed.

A calculation shows that $\mathbb{E}(S_N) = \frac{N}{2} + o(N)$ and $\mathrm{Var}(S_N) = \frac{N}{12} + o(N)$. Clearly ess sup $S_N = 2N$, so $\mathfrak{r}_+ = \lim\limits_{N\to\infty} \dfrac{2N - (N/2 + o(N))}{N/12} = 18$. We will show that $\mathfrak{c}_+ = 6$, proving that $(\mathfrak{c}_-, \mathfrak{c}_+) \neq (\mathfrak{r}_-, \mathfrak{r}_+)$.

$$\mathcal{F}_N(\xi) = \frac{1}{V_N} \log \prod_{n=1}^{N} \mathbb{E}\left(e^{\xi f_n(Y_n,Z_n)}\right) \sim \frac{12}{N} \sum_{n=1}^{N} \log \mathbb{E}\left(e^{\xi f_n(Y_n,Z_n)}\right) = \frac{12}{N} \sum_{n=1}^{N} \log\left(p_n e^{2\xi} + (1-p_n)\mathbb{E}(e^{\xi U[0,1]})\right) \to 12 \log \mathbb{E}\left(e^{\xi U[0,1]}\right),$$

because $p_n \to 0$. Hence $\mathcal{F}_N(\xi) \xrightarrow[N\to\infty]{} 12 \log\left(\frac{e^\xi - 1}{\xi}\right) \overset{!}{<} 12\xi$ for $\xi>0$, because $\dfrac{e^\xi - 1}{\xi} = \sum\limits_{n=0}^{\infty} \dfrac{\xi^n}{(n+1)!} < e^\xi$. Therefore for every $\xi > 0$, for every sufficiently large N,

$$\mathbb{E}\left(e^{\xi S_N}\right) \le e^{12 V_N \xi}. \tag{7.52}$$

Take some arbitrary $0 < z < z' < \mathfrak{c}_+$. By Corollary 7.27, we can choose z' so that $\mathfrak{I}_1(z') < \infty$, and there is some $\eta > 0$ such that for all N large enough, for all $\xi > 0$,

$$\eta^{V_N} \le \mathbb{P}(S_N - \mathbb{E}(S_N) \in z' V_N + [-1,1]) \le \mathbb{P}[S_N > z V_N + \mathbb{E}(S_N)] \le \mathbb{E}(e^{\xi[S_N - zV_N - \mathbb{E}(S_N)]}) \overset{(7.52)}{\le} e^{\xi V_N [12 - (z + \mathbb{E}(S_N)/V_N)]}.$$

Since η is fixed but ξ can be arbitrarily large, the term in the square brackets must be non-negative for all N large enough. As $\mathbb{E}(S_N)/V_N \to 6$, z must be no larger than 6. Since z can be chosen to be arbitrarily close to $\mathfrak{c}_+$, we get $\mathfrak{c}_+ \le 6$.

Next we show that $\mathfrak{c}_+ = 6$. It suffices to show that $z_N := zV_N + \mathbb{E}(S_N)$ is admissible for each $0 < z < 6$ (because then z_N satisfies the conditions of Theorem 7.8, and we are in the first case of Theorem 7.26).

By Theorem 7.30, it suffices to find $\varepsilon, \eta > 0$ such that for all N large enough,

$$\mathbb{P}(S_N \le z_N - \varepsilon V_N) \ge \eta^{V_N} \tag{7.53}$$

$$\mathbb{P}(S_N \ge z_N + \varepsilon V_N) \ge \eta^{V_N}. \tag{7.54}$$

(7.53) is because for $\varepsilon < z$, $\mathbb{P}(S_N \le z_N - \varepsilon V_N) \ge \mathbb{P}(S_N - \mathbb{E}(S_N) \le 0) \to \frac{1}{2}$, by the CLT. To see (7.54), we let $S'_N := Z_1 + \cdots + Z_N$, and note that $\mathbb{E}(S'_N) = \frac{N}{2} = \mathbb{E}(S_N) + o(N)$, $V'_N := \mathrm{Var}(S'_N) = \frac{N}{12} = V_N + o(N)$, and

$S_N \geq S'_N$ (because $f_n(X_n, Y_n) \geq Z_n$). Let $z'_N := zV'_N + \mathbb{E}(S'_N)$, then $z'_N = z_N + o(N)$, and so for all large N,

$$\mathbb{P}(S_N \geq z_N + \varepsilon V_N) \geq \mathbb{P}\left(S'_N \geq z_N + \varepsilon V_N\right) \geq \mathbb{P}\left(S'_N \geq z'_N + 2\varepsilon V'_N\right).$$

By Example 7.35, the large deviations thresholds for S'_N are the positivity thresholds of S'_N, which can be easily found to be ± 6. Since $0 < z < 6$, $\mathbb{P}\left(S'_N \geq z'_N + 2\varepsilon V'_N\right) \geq \eta_1^{V'_N} \geq \eta_2^{V_N}$ for some $\varepsilon, \eta_i > 0$ and all N large enough. This completes the proof of (7.54), and the proof that $\mathfrak{c}_+ = 6$.

Example 7.37 (Failure of Theorem 7.26 at $z = \mathfrak{c}_\pm$) This already happens in Example 7.1, when $S_N = X_1 + \cdots + X_N$, and X_i are iid random variables such that $X_i = -1, 0, 1$ with equal probabilities.

By Example 7.35, $\mathfrak{c}_\pm = \mathfrak{r}_\pm = \pm\frac{3}{2}$. Clearly, $\mathbb{E}(S_N) = 0$ and $V_N = \frac{2}{3}N$. So $z_N^\pm := N \pm 1$ satisfy $\frac{z_N^\pm - \mathbb{E}(S_N)}{V_N} \to \mathfrak{c}_+$. However, $\mathbb{P}[S_N - z_N^+ \in (-\frac{1}{2}, \frac{1}{2})] = 0$, which violates the first alternative of Theorem 7.26, and $\mathbb{P}[S_N - z_N^- \in (-\frac{1}{2}, \frac{1}{2})] = 3^{-N}N$ which violates the second alternative of the Theorem 7.26.

7.5 Notes and References

The reader should note the difference between the LLT for large deviations and the large deviations principle (LDP): LLT for large deviations give the asymptotics of $\mathbb{P}[S_N - z_N \in (a, b)]$ or $\mathbb{P}[S_N > z_N]$; The LDP gives the asymptotics of the *logarithm* of $\mathbb{P}[S_N > z_N]$, see Dembo & Zeitouni [40] and Varadhan [197].

The interest in precise asymptotics for $\mathbb{P}[S_N > z_N]$ in the regime of large deviations goes back to the first paper on large deviations, by Cramér [34]. That paper gave an asymptotic series expansion for $\mathbb{P}[S_N - \mathbb{E}(S_N) > x]$ for the sums of iid random variables. The first sharp asymptotics for $\mathbb{P}[S_N - z_N \in (a, b)]$ appear to be the work of Richter [167], [103, chapter 7] and Blackwell & Hodges [14].

These results were refined by many authors, with important contributions by Petrov [155], Linnik [133], Moskvin [146], Bahadur & Ranga Rao [11], Statulavicius [188] and Saulis [176]. For accounts of these and other results, we refer the reader to the books of Ibragimov & Linnik [103], Petrov [156], and Saulis & Statulevicius [177]. See also the survey of Nagaev [150].

Plachky and Steinebach [158] and Chaganty & Sethuraman [23, 24] proved LLT for large deviations for arbitrary sequences of random variables T_n (e.g. sums of dependent random variables), subject only to assumptions on the asymptotic behavior of the normalized log-moment generating functions of T_n and their Legendre-Fenchel transforms (their rate functions). Our LLT for large deviations are in the spirit of these results.

Corollary 7.10 is an example of a limit theorem conditioned on a large deviation. For other examples of such results, in the context of statistical physics, see [45].

We comment on some of the technical devices in the proofs. The "change of measure" trick discussed in section 7.3.1 goes back to Cramér [34] and is a standard idea in large deviations. In the classical homogeneous setup, a single parameter $\xi_N = \xi$ works for all times N, but in our inhomogeneous setup, we need to allow the parameter ξ_N to depend N. For other instances of changes of measure which involve a time dependent parameter, see Dembo & Zeitouni [39] and references therein.

The Gärtner-Ellis Theorem we used to prove Theorem 7.7 can be found in [69]. The one-dimensional case, which is sufficient for our purposes, is stated and proved in appendix A, together with historical comments.

Birkhoff's Theorem is proved in [13], and is discussed further in appendix B.

Results similar to Lemma 7.13 on the existence of the generalized eigenfunction h_n^ξ were proved by many authors in many different contexts, see for example [15], [67], [75], [90], [93], [113], [173]. The analytic dependence of the generalized eigenvalue and eigenvector on the parameter ξ was considered in a different context (the top Lyapunov exponent) by Ruelle [172] and Peres [154]. Our proof of Lemma 7.16 follows closely a proof in Dubois's paper [67].

For an account of the theory of real-analyticity for vector valued functions, see [49] and [199].

Chapter 8
Important Examples and Special Cases

Abstract In this chapter we consider several special cases where our general results take stronger form. These include sums of independent random variables, and homogeneous or asymptotically homogeneous or equicontinuous additive functionals.

8.1 Introduction

In the previous chapters we studied the LLT for general uniformly bounded additive functionals f on uniformly elliptic inhomogeneous Markov chains X. We saw that

- If $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, and $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, then $\mathbb{P}[S_N - z_N \in (a,b)] \sim \dfrac{|a-b|}{\sqrt{2\pi V_N}} \mathrm{e}^{-z^2/2}$.
- If $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, and $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to 0$, then $\mathbb{P}[S_N - z_N \in (a,b)] \sim \dfrac{|a-b|}{\sqrt{2\pi V_N}} \mathrm{e}^{-V_N \mathcal{I}_N\left(\frac{z}{V_N}\right)}$, where $\mathcal{I}_N$ are the rate functions, see §§7.2.2, 7.2.3.
- If $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, $z \in (\mathfrak{c}_-, \mathfrak{c}_+)$, and $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to z$, then

$$\mathbb{P}[S_N - z_N \in (a,b)] \sim \frac{|a-b|}{\sqrt{2\pi V_N}} \mathrm{e}^{-V_N \mathcal{I}_N\left(\frac{z}{V_N}\right)} \rho_N\left(\tfrac{z_N - \mathbb{E}(S_N)}{V_N}\right) \times \frac{1}{|a-b|} \int_a^b \mathrm{e}^{-t\xi_N\left(\frac{z_N - \mathbb{E}(S_N)}{V_N}\right)} \mathrm{d}t,$$

  where $\rho_N(\eta) \xrightarrow[\eta\to 0]{} 1$ and $\xi_N(\eta) \xrightarrow[\eta\to 0]{} 0$ uniformly in N (see §7.2.4), and $\mathfrak{c}_\pm$ are the large deviation thresholds of (X,f) (see §7.4).

We will now apply the general theory to special cases of interest. The point is to verify the condition $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$ and to find $\mathfrak{c}_\pm$.

8.2 Sums of Independent Random Variables

Throughout this section, let X_n be independent real-valued random variables, possibly with different distributions, such that for some K, $|X_n| \le K$ for all n. Let $S_N = X_1 + \cdots + X_N$, $V_N := \mathrm{Var}(S_N) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_N)$. This is a special case of the setup studied in the previous chapters: $\mathsf{X} := \{X_n\}$ is a uniformly elliptic inhomogeneous Markov chain (with ellipticity constant $\epsilon_0 = 1$), and $S_N \equiv S_N(\mathsf{f})$, for the uniformly bounded additive functional $f_n(x,y) := x$.

Proposition 8.1 *(X,f) is center-tight iff $\sum \mathrm{Var}(X_n) < \infty$. In this case, the limit $\lim_{N\to\infty}(S_N - \mathbb{E}(S_N))$ exists and is finite almost surely.*

Proof Corollary 3.9 says that (X,f) is center-tight iff $\sup_N U_N \equiv \sum_{n=3}^{\infty} u_n^2 < \infty$, where u_n are given by (2.26). In Proposition 2.20 we saw that $u_n^2 = 2(\mathrm{Var}(X_{n-1}) + \mathrm{Var}(X_n))$. Thus (X,f) is center-tight iff $\sum \mathrm{Var}(X_n) < \infty$. The a.s. existence of $\lim_{n\to\infty}(S_n - \mathbb{E}(S_n))$ is due to the two-series theorem (see also Theorem 3.12). □

Recall the following definition from Chapter 3: Given a real-valued random variable X and a number $\xi \in \mathbb{R}$,

$$\mathfrak{D}(X,\xi) := \min_{\theta\in\mathbb{R}} \mathbb{E}\left[\mathrm{dist}^2\left(X, \theta + \frac{2\pi}{\xi}\mathbb{Z}\right)\right]^{1/2}.$$

Proposition 8.2 *Suppose* $\sum \mathrm{Var}(X_n) = \infty$.

(1) (X,f) *has essential range* $\mathbb{R}$ *iff* $\sum_{n=1}^{\infty} \mathfrak{D}(X_n,\xi)^2 = \infty$ *for all* $\xi > 0$.

(2) Otherwise, the following holds:

(a) There is a minimal positive ξ *such that* $\sum_{n=1}^{\infty} \mathfrak{D}(X_n,\xi)^2 < \infty$. *(b)* $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$ *for* $t := 2\pi/\xi$.

(c) We can decompose $X_n = \mathbb{F}_n(X_n) + h_n(X_n)$ *where* $\mathbb{F}_n$ *are uniformly bounded measurable functions taking values in* $t\mathbb{Z}$, *and* $\sum \mathrm{Var}[h_n(X_n)] < \infty$.

(d) $\exists$ *constants* γ_N *such that* $S_N - \gamma_N$ *converges a.s. modulo* $t\mathbb{Z}$, *i.e.* $\lim_{N\to\infty} \exp\left[\frac{2\pi \mathrm{i}}{t}(S_N - \gamma_N)\right]$ *exists a.s.*

Notice that unlike the general case discussed in Lemma 4.16, the decomposition in (c) does not require a gradient.

Proof By Proposition 2.21, for every $\xi > 0$ there is a constant $C(\xi) > 1$ such that

$$d_n^2(\xi) = C(\xi)^{\pm 1}[\mathfrak{D}(X_{n-1},\xi)^2 + \mathfrak{D}(X_n,\xi)^2)]. \tag{8.1}$$

Therefore, $\sum \mathfrak{D}(X_n,\xi)^2 < \infty$ iff $\sum d_n^2(\xi) < \infty$. Hence the co-range of (X,f) is given by $H(\mathsf{X},\mathsf{f}) = \{\xi \in \mathbb{R} : \sum_{n=1}^{\infty} \mathfrak{D}(X_n,\xi)^2 < \infty\}$. Parts (1), (2)(a), and (2)(b) now follow from Theorems 4.3 and 4.4 on $H(\mathsf{X},\mathsf{f})$.

Next we prove (2)(c). Let $\xi > 0$ be as in (2)(a), and set $t := 2\pi/\xi$. For every n, choose $\theta_n \in \mathbb{R}$ such that

$$\mathbb{E}\left[\mathrm{dist}^2\left(X_n, \theta_n + t\mathbb{Z}\right)\right] < \mathfrak{D}_n(X_n,\xi)^2 + \frac{1}{2^n}.$$

We can decompose every $x \in \mathbb{R}$ into $x = \mathbb{F}_n(x) + h_n(x)$, where

$$\mathbb{F}_n(x) := \text{ the (minimal) closest point to } x - \theta_n \text{ in } t\mathbb{Z}, \quad h_n(x) := x - \mathbb{F}_n(x).$$

Necessarily, $|h_n(x) - \theta_n| = \mathrm{dist}(x, \theta_n + t\mathbb{Z})$, and $X_n = \mathbb{F}_n(X_n) + h_n(X_n)$ with $\mathbb{F}_n$ bounded and taking values in $t\mathbb{Z}$. We claim that h has summable variances. Recall that for a random variable Y, $\mathrm{Var}(Y) = \min_{\theta\in\mathbb{R}} \mathbb{E}[(Y-\theta)^2]$. Thus

$$\sum_{n=1}^{\infty} \mathrm{Var}(h_n(X_n)) = \sum_{n=1}^{\infty} \min_{\theta} \mathbb{E}[(h_n(X_n)-\theta)^2] \le \sum_{n=1}^{\infty} \mathbb{E}[(h_n(X_n) - \theta_n)^2] = \sum_{n=1}^{\infty} \mathbb{E}[\mathrm{dist}^2(x, \theta_n + t\mathbb{Z})] < \left[\sum_{n=1}^{\infty} \mathfrak{D}_n^2(X_n, t)\right] + 1 < \infty.$$

We proved (2)(c).

Let $\alpha_n := \mathbb{E}[h_n(X_n)]$. By the two-series theorem, the series $\sum_{n=1}^{\infty}(h_n(X_n) - \alpha_n)$ converges a.s. Therefore, if $\gamma_N := \alpha_1 + \cdots + \alpha_N$, then $\exp\left[\frac{2\pi \mathrm{i}}{t}(S_N - \gamma_N)\right] =$

$$\exp\left[\frac{2\pi \mathrm{i}}{t}(S_N(\mathbb{F}) + S_N(\mathsf{h}) - \gamma_N)\right] = \exp\left[\frac{2\pi \mathrm{i}}{t}(S_N(\mathsf{h}) - \gamma_N)\right] \xrightarrow[N\to\infty]{} \exp\left(\frac{2\pi \mathrm{i}}{t}\sum_{n=1}^{\infty}(h_n(X_n) - \alpha_n)\right),$$ whence (2)(d).□

We are now ready to state the non-lattice LLT for independent random variables:

Theorem 8.3 (Dolgopyat) *Let* X_n *be a sequence of uniformly bounded independent real-valued random variables such that* $\sum \mathrm{Var}(X_n) = \infty$.

(1) Suppose $\sum_{n=1}^{\infty} \mathfrak{D}(X_n,\xi)^2 = \infty$ *for all* $\xi > 0$. *Then for every* $z_N, z \in \mathbb{R}$ *s.t.* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, *for every* $a < b$,

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1 + o(1)]\frac{e^{-z^2/2}}{\sqrt{2\pi V_N}}|a-b| \text{ as } N \to \infty.$$

(2) Otherwise, there is a finite maximal t such that $\sum_{n=1}^{\infty} \mathfrak{D}\left(X_n, \frac{2\pi}{t}\right)^2 < \infty$*, and:*

(a) There exist constants γ_N *s.t.* $S_N - \gamma_N$ *converge a.s. modulo* $t\mathbb{Z}$*, (i.e.* $\lim_{N\to\infty} \exp\left[\frac{2\pi \mathrm{i}}{t}(S_N - \gamma_N)\right]$ *exists a.s.).*

(b) There is a bounded random variable $\mathfrak{F}(X_1, X_2, \ldots)$ *and a bounded sequence of numbers* b_N *s. t. for all* $z_N \in b_N + t\mathbb{Z}$ *such that* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z \in \mathbb{R}$, $\forall \phi \in C_c(\mathbb{R})$, $\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}[\phi(S_N - z_N)] = \frac{t\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(mt + \mathfrak{F})]$.

Proof In case 1, Proposition 8.2 says that $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$. So (X, f) is non-lattice and irreducible, and the first part of the theorem follows from Theorem 5.1. In case 2, Proposition 8.2 says that there is a minimal $\xi > 0$ such that $\sum \mathfrak{D}(X_n, \xi)^2 < \infty$, and for this ξ, $G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$, where $t := 2\pi/\xi$. In addition, there are constants γ_N such that $S_N - \gamma_N$ converge a.s. modulo $t\mathbb{Z}$.

By Theorem 6.3(a) there are bounded random variables $\mathfrak{F}(X_1, X_2, \ldots)$ and $b_N(X_1, X_{N+1})$ so that for all $\phi \in C_c(\mathbb{R})$ and $z'_n \in t\mathbb{Z}$ such that $\frac{z'_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$,

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}[\phi(S_N - z'_N - b_N)] = \frac{t\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \mathbb{E}[\phi(mt + \mathfrak{F})]. \tag{8.2}$$

To finish the proof, we need to show that (8.2) holds with *constant* b_N, because then we can take $z'_N := z_N - b_N$ for any $z_N \in b_N + t\mathbb{Z}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$.

There is no loss of generality in assuming that $\mathbb{E}(X_n) = 0$ for all n, and $t = 1$ (this can always be achieved by scaling). By Proposition 8.2(c), we can decompose $X_n = \mathbb{F}_n(X_n) + h_n(X_n)$ with $\mathbb{F}_n$ uniformly bounded taking values in $\mathbb{Z}$, and where h has summable variances. The absence of a gradient term in this decomposition places us in the "special case" (6.40), where the proof of Theorem 6.3(a) gives (8.2) with b_N constant and bounded. □

Next, we discuss the lattice LLT for sums of independent *integer*-valued random variables X_n. We say that X_n satisfies **Prokhorov's condition**, if

$$\prod_{k=1}^{\infty} \left(\max_{0\le m<t} \mathbb{P}(X_k = m \bmod t)\right) = 0 \text{ for all integers } t \ge 2. \tag{8.3}$$

Let m_k be the (smallest) most likely residue mod t for X_k. Then it is not difficult to see that Prokhorov's condition is equivalent to

$$\sum_k \mathbb{P}[X_k \ne m_k \bmod t] = \infty \text{ for all integers } t \ge 2. \tag{8.4}$$

Lemma 8.4 *Let* X_n *be a sequence of independent* $\mathbb{Z}$*-valued random variables, then the following are equivalent: (1) Prokhorov's condition* (8.4)*; (2)* (X, f) *is irreducible, with essential range* $\mathbb{Z}$*.*

Proof

(2) ⇒ (1): Fix an integer $t \ge 2$. Every $x \in \mathbb{Z}$ can be decomposed uniquely in the form $x = \{x\}_{t\mathbb{Z}} + [x]_{t\mathbb{Z}}$, where $\{x\}_{t\mathbb{Z}} \in [0, t)$ and $[x]_{t\mathbb{Z}} \in t\mathbb{Z}$. Set

- $y_k(x) :=$ the (smallest) integer in $m_k + t\mathbb{Z}$ closest to x, and $z_k(x) := x - y_k(x)$,
- $g_k(x) := (y_k(x) - m_k) + [x - y_k(x)]_{t\mathbb{Z}}$ (g_k takes values in $t\mathbb{Z}$),
- $h_k(x) := \{x - y_k(x)\}_{t\mathbb{Z}}$ (h_k takes values in $\mathbb{Z}$). Then

$$X_k = g_k(X_k) + h_k(X_k) + m_k.$$

The algebraic range of (X, g) is inside $t\mathbb{Z}$, and by the Borel-Cantelli lemma, (8.4) fails ⇔ $X_k \ne m_k \bmod t\mathbb{Z}$ finitely often a.s. ⇔ $h_k(X_k) \ne 0$ finitely often a.s. So if (8.4) fails, then $\sum_{k=0}^{\infty} |h_k(X_k)| < \infty$ almost surely. Hence h is center-tight. Since $G_{alg}(\mathsf{X}, \mathsf{f} - \mathsf{h}) = G_{alg}(\mathsf{X}, \mathsf{g}) \subset t\mathbb{Z}$, $G_{ess}(\mathsf{X}, \mathsf{f}) \ne \mathbb{Z}$ contradicting to (2).

(1) ⇒ (2): Fix an integer $t \geq 2$, $\theta \in [0,t)$, and let m_θ be the (smallest) closest integer to θ. Then $|m' - \theta| \geq \frac{1}{2}$ for $m' \neq m_\theta$, whence $\mathbb{E}[\mathrm{dist}^2(X_n, \theta + t\mathbb{Z})] \geq \dfrac{1}{4}\mathbb{P}(X_n \neq m_\theta \bmod t) \geq \dfrac{1}{4}[1 - \max\limits_{0\leq m<t} \mathbb{P}(X_n = m \bmod t)]$.

Passing to the infimum over θ, we obtain $\mathfrak{D}^2(X_n, \frac{2\pi}{t}) \geq \dfrac{1}{4}[1 - \max\limits_{0\leq m<t} \mathbb{P}(X_n = m \bmod t)] = \dfrac{1}{4}\mathbb{P}[X_n \neq m_n \bmod t]$.

By (8.1) and (8.4), we get $\sum\limits_{n=3}^{\infty} d_n^2(\frac{2\pi}{t}) \geq \mathrm{const} \sum\limits_{n=3}^{\infty} (\mathfrak{D}^2(X_{n-1}, \frac{2\pi}{t}) + \mathfrak{D}^2(X_n, \frac{2\pi}{t})) = \infty$.

We find that the co-range does *not* contain $2\pi/t$ for $t \in \{2,3,4,\ldots\}$. But it *does* contain 2π (because X_k are integer-valued). The only closed sub-group of $\mathbb{R}$ with these properties is $2\pi\mathbb{Z}$. So the co-range is $2\pi\mathbb{Z}$, and the essential range is $\mathbb{Z}$. Since X_k are integer-valued, we have irreducibility. □

Theorem 8.5 (Prokhorov) *Let X_n be a uniformly bounded sequence of integer-valued independent random variables. Assume* (8.3). *Then* $\sum \mathrm{Var}(X_n) = \infty$, *and for all $z_N \in \mathbb{Z}$ and $z \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \to z$, $\forall k \in \mathbb{Z}$,*

$$\mathbb{P}[S_N - z_N = k] = [1 + o(1)]\frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} \text{ as } N \to \infty.$$

Proof The theorem is a direct consequence of Lemma 8.4, and Theorem 5.2. □

Prokhorov also showed that (8.3) is a *necessary* condition for his LLT to apply to all uniformly bounded sequences of integer-valued independent random variables X_n' such that for some r, $X_n' = X_n$ for all $n \geq r$.We omit the proof, which given Lemma 8.4, follows from a lattice version of Theorem 6.5.

We now turn to the LLT for large deviations. Recall that the large deviations thresholds are the endpoints of the largest interval $(\mathfrak{c}_-, \mathfrak{c}_+)$ so that the LLT for large deviations in Chapter 7 applies to all sequences z_N such that $\frac{z_N - \mathbb{E}(S_N)}{V_N} \to z$ with $z \in (\mathfrak{c}_-, \mathfrak{c}_-)$ (see Theorem 7.26). Necessarily, $(\mathfrak{c}_-, \mathfrak{c}_+) \subset (\mathfrak{r}_-, \mathfrak{r}_+)$, where

$$\mathfrak{r}_- := \limsup_{N\to\infty} \frac{\mathrm{ess\ inf}[S_N - \mathbb{E}(S_N)]}{V_N}, \quad \mathfrak{r}_+ := \liminf_{N\to\infty} \frac{\mathrm{ess\ sup}[S_N - \mathbb{E}(S_N)]}{V_N}.$$

"Full large deviations regime" means that $(\mathfrak{c}_-, \mathfrak{c}_+) = (\mathfrak{r}_-, \mathfrak{r}_+)$.

We saw in Example 7.35 that sums of iid random variables have full large deviation regime. We will now give a sufficient condition for full regime in the case of non-identically distributed independent random variables.

A sequence of bounded real-valued independent random variables X_n is called **tame**, when one of the following conditions holds:

(a) $\liminf\limits_{N\to\infty} V_N/N > 0$, and $\forall\delta > 0$ there is an $\eta_\delta > 0$ such that for all n,
$$\mathbb{P}[X_n > \mathrm{ess\ sup}\, X_n - \delta\mathrm{Var}(X_n)] \geq \eta_\delta, \qquad \mathbb{P}[X_n < \mathrm{ess\ inf}\, X_n + \delta\mathrm{Var}(X_n)] \geq \eta_\delta. \tag{8.5}$$

(b) $\liminf\limits_{N\to\infty} V_N/N = 0$, and $\forall\delta > 0$ there is an $\eta_\delta > 0$ such that for all n,
$$\mathbb{P}[X_n > \mathrm{ess\ sup}\, X_n - \delta\mathrm{Var}(X_n)] \geq \eta_\delta^{\mathrm{Var}(X_n)}, \qquad \mathbb{P}[X_n < \mathrm{ess\ inf}\, X_n + \delta\mathrm{Var}(X_i)] \geq \eta_\delta^{\mathrm{Var}(X_n)}. \tag{8.6}$$

If $\mathrm{Var}(X_n) = O(1)$, then (8.6) implies (8.5), but it is not equivalent to it. For example, suppose there are sequences $n_k \to \infty$ and $\sigma_{n_k} \in (0,1)$ such that $\sigma_{n_k} \to 0$, and $X_{n_k} = \pm\sigma_{n_k}$ with probabilities $\frac{1}{2}$. Then $\mathbb{P}[X_{n_k} > \mathrm{ess\ sup}\, X_{n_k} - \delta\mathrm{Var}(X_{n_k})] = \frac{1}{2}$, and (8.5) holds. But (8.6) fails, because $\eta_\delta^{\mathrm{Var}(X_{n_k})} \to 1$ for all $\eta_\delta > 0$.

Example 8.6 Let X_n be bounded independent random variables with non-zero variance. The tameness assumption holds in each of the following cases:

(1) X_n are identically distributed;
(2) There are finitely many random variables $Y_1, \ldots, Y_k$ such that for all n there is some k such that $X_n = Y_k$ in distribution;
(3) X_n are uniformly bounded discrete random variables, and there is $\varepsilon_0 > 0$ such that every atom of X_n has mass bigger than or equal to ε_0.

Theorem 8.7 *Suppose X_n is a uniformly bounded and tame sequence of independent random variables such that $\sum \mathrm{Var}(X_n) = \infty$, then*

$$\mathfrak{c}_+ = \mathfrak{r}_+ = \liminf_{N\to\infty} \frac{1}{V_N} \sum_{n=1}^{N} \operatorname{ess\,sup}(X_n - \mathbb{E}(X_n)), \qquad \mathfrak{c}_- = \mathfrak{r}_- = \limsup_{N\to\infty} \frac{1}{V_N} \sum_{n=1}^{N} \operatorname{ess\,inf}(X_n - \mathbb{E}(X_n)).$$

Proof Without loss of generality, $\mathbb{E}(X_n) = 0$. By symmetry, $\mathfrak{c}_-(\mathsf{X}, \mathsf{f}) = -\mathfrak{c}_+(\mathsf{X}, -\mathsf{f})$ and $\mathfrak{r}_+(\mathsf{X}, -\mathsf{f}) = -\mathfrak{r}_-(\mathsf{X}, \mathsf{f})$. Therefore, it is sufficient to prove the inequalities for $\mathfrak{c}_+$.

We claim that $\mathfrak{c}_+ \le \mathfrak{r}_+ = \liminf\limits_{N\to\infty} \dfrac{1}{V_N} \sum\limits_{n=1}^{N} \operatorname{ess\,sup}(X_n)$. The first inequality is valid for general Markov chains (see §7.4); the second follows from the independence of X_n.

Next we show that $\mathfrak{c}_+ \ge \mathfrak{r}_+$. Since $V_N \to \infty$, $\mathfrak{c}_+ > 0$, (see Lemma 7.19(2)). Therefore, by what we just proved, $\mathfrak{r}_+ > 0$. Fix $0 < \delta < \frac{1}{4}\mathfrak{r}_+$ small, and choose some $\delta < z < \mathfrak{r}_+ - 3\delta$. For all large enough N, $zV_N + \delta V_N \le \sum\limits_{n=1}^{N} \operatorname{ess\,sup}(X_n) - \delta V_N$, so

$$\begin{aligned}\mathbb{P}[S_N \ge zV_N + \delta V_N] &\ge \mathbb{P}\Big[S_N \ge \sum_{n=1}^{N} \operatorname{ess\,sup}(X_n) - \delta V_N\Big] \\ &\ge \prod_{i=1}^{N} \mathbb{P}[X_i \ge \operatorname{ess\,sup}(X_i) - \delta \mathrm{Var}(X_i)], \text{ because } X_i \text{ are independent.}\end{aligned}$$

By the tameness assumption, there exists $0 < \eta < 1$ independent of N such that

$$\mathbb{P}[S_N \ge zV_N + \delta V_N] \ge \eta^{V_N}. \tag{8.7}$$

Indeed, in case (b), (8.7) is straightforward, and in case (a) we use the inequalities $\mathbb{P}[S_N \ge zV_N + \delta V_N] \ge \eta_\delta^N$ and $V_N \ge \text{const.}N$ for large N.

Next, by the CLT, since $z > \delta$, $\mathbb{P}[S_N \le zV_N - \delta V_N] \ge \mathbb{P}[S_N \le 0] \to \frac{1}{2}$. So for all N large enough,

$$\mathbb{P}[S_N \le zV_N - \delta V_N] \ge \eta^{V_N}. \tag{8.8}$$

By (8.7), (8.8) and Theorem 7.30, $z_N := zV_N \equiv zV_N + \mathbb{E}(S_N)$ is admissible.

As explained in §7.4.3, the admissibility of z_N means that z is reachable, and therefore $z \in [\mathfrak{c}_-, \mathfrak{c}_+]$. In particular, $z \le \mathfrak{c}_+$. Passing to the supremum over z we obtain $\mathfrak{r}_+ - 3\delta \le \mathfrak{c}_+$. Passing to the limit $\delta \to 0$, we obtain $\mathfrak{r}_+ \le \mathfrak{c}_+$. □

In the absence of the tameness condition, the identities $\mathfrak{c}_\pm = \mathfrak{r}_\pm$ may be false, see Example 7.36. We are not aware of general formulas for $\mathfrak{c}_\pm$ in such cases.

8.3 Homogenous Markov Chains

A Markov chain $\mathsf{X} = \{X_n\}$ is called **homogeneous**, if its state spaces and transition probabilities do not depend on n. In this case we let

$$\mathfrak{S}_n = \mathfrak{S}, \quad \pi_n(x, \mathrm{d}y) = \pi(x, \mathrm{d}y).$$

An additive functional on a homogeneous Markov chain is called **homogeneous**, if $\mathsf{f} = \{f_n\}$, where $f_n(x, y) = f(x, y)$ for all n.

If $f(x, y) = a(x) - a(y)$ with $a : \mathfrak{S} \to \mathbb{R}$ bounded and measurable, then f is called a **homogeneous gradient**.

A Markov chain is called **stationary**, if $\{X_n\}$ is a **stationary stochastic process**: For each n, the joint distribution of $(X_{1+k}, \ldots, X_{n+k})$ is the same for all $k \ge 0$. Homogeneity and stationarity are closely related:

- Every stationary Markov chain is equal in distribution to a homogeneous Markov chain with state space $\mathfrak{S} := \bigcap_{n\geq 1}\mathfrak{S}_n$, and transition kernel $\pi = \pi_1|_{\mathfrak{S}}$. Moreover, the initial distribution $\mu(E) := \mathbb{P}[X_1 \in E]$ must satisfy
$$\mu(E) = \int \pi(x, E)\mu(\mathrm{d}x) \quad (E \in \mathscr{B}(\mathfrak{S})) \tag{8.9}$$
(on the left we have $\mathbb{P}[X_1 \in E]$, and on the right $\mathbb{P}[X_2 \in E]$).
- Conversely, any homogeneous Markov chain with an initial distribution satisfying (8.9) is stationary. To see this iterate (8.9) to see that $\mathbb{P}[X_n \in E] = \mu(E)$ for all n, and then use (2.1).

Probability measures satisfying (8.9) are called **stationary**.

Every homogeneous Markov chain with a finite state space admits a stationary measure, by the Perron-Frobenius theorem. Some Markov chains on infinite state spaces, e.g. null recurrent Markov chains, do not have stationary measures. However, if the chain is uniformly elliptic, then a stationary measure always exists:

Lemma 8.8 *Let* X *be a uniformly elliptic homogeneous Markov chain. Then* X *admits a unique stationary initial distribution.*

Proof Fix $x \in \mathfrak{S}$, and consider the measure μ_n which describes the distribution of X_n given $X_1 = x$, i.e. $\mu_n(\phi) = \mathbb{E}(\phi(X_n)|X_1 = x)$.

By homogeneity and Proposition 2.13, there is a constant $0 < \theta < 1$ such that for every bounded measurable $\phi : \mathfrak{S} \to \mathbb{R}$,

$$\begin{aligned}\mu_{n+1}(\phi) &= \int \mathbb{E}(\phi(X_{n+1})|X_2 = y)\pi(x, \mathrm{d}y) = \int \mathbb{E}(\phi(X_n)|X_1 = y)\pi(x, \mathrm{d}y)\\ &= \int \left[\mu_n(\phi) + O(\theta^n)\right]\pi(x, \mathrm{d}y) = \mu_n(\phi) + O(\theta^n).\end{aligned}$$

Necessarily $\{\mu_n(\phi)\}$ is a Cauchy sequence, and $\mu(\phi) := \lim\limits_{n\to\infty} \mu_n(\phi)$ exists for every bounded measurable ϕ.

Let $\widehat{\mu}$ denote the set function $\widehat{\mu}(E) := \mu(1_E)$. We claim that $\widehat{\mu}$ is σ-additive. Finite additivity is clear, because $\phi \mapsto \mu(\phi)$ is linear. Next suppose A is a disjoint union of measurable sets $\{A_l\}$. By finite additivity,

$$\widehat{\mu}(A) = \left[\sum_{l=1}^{L} \widehat{\mu}(A_l)\right] + \widehat{\mu}\left(\bigcup_{l=L+1}^{\infty} A_l\right). \tag{8.10}$$

Let ϵ_0 and ν be the ellipticity constant and background measure with respect to which our chain satisfies the ellipticity condition. (By definition, ν_n depends on n but since the chain is homogeneous, $\nu := \nu_3$ will work for all n.) Applying Proposition 2.8 to X conditioned on $X_1 = x$, we obtain that $\dfrac{d\mu_n}{d\nu} \leq \epsilon_0^{-1}$ for $n \geq 3$. Hence,

$$0 \leq \widehat{\mu}\left(\bigcup_{l=L+1}^{\infty} A_l\right) = \mu\left(1_{\bigcup_{l=L+1}^{\infty} A_l}\right) = \lim_{n\to\infty} \mu_n\left(1_{\bigcup_{l=L+1}^{\infty} A_l}\right) \leq \epsilon_0^{-1}\nu\left(\bigcup_{l=L+1}^{\infty} A_l\right).$$

Since ν is σ-additive, $\lim\limits_{L\to\infty} \nu\left(\bigcup_{l=L+1}^{\infty} A_l\right)$=0. It follows that $\lim\limits_{L\to\infty} \widehat{\mu}\left(\bigcup_{l=L+1}^{\infty} A_l\right)$=0. Looking at (8.10), we find that $\widehat{\mu}(A) = \sum\limits_{l=1}^{\infty} \widehat{\mu}(A_l)$. So $\widehat{\mu}$ is a well-defined measure.

A standard approximation argument shows that for any bounded measurable function ϕ, $\mu(\phi) = \int \phi \mathrm{d}\widehat{\mu}$. It follows that for every bounded measurable $\phi : \mathfrak{S} \to \mathbb{R}$,

$$\lim_{n\to\infty} \int \phi \mathrm{d}\mu_n = \int \phi \mathrm{d}\widehat{\mu}. \tag{8.11}$$

Take an arbitrary measurable set $E \subset \mathfrak{S}$, then

$$\widehat{\mu}(E) = \mu(1_E) = \lim_{n\to\infty} \mu_n(1_E) = \lim_{n\to\infty} \mathbb{E}_x[1_E(X_{n+1})] = \lim_{n\to\infty} \mathbb{E}_x[\mathbb{E}(1_E(X_{n+1})|X_n)]$$
$$= \lim_{n\to\infty} \iint \pi(y,\mathrm{d}z)1_E(z)\mu_n(\mathrm{d}y) = \lim_{n\to\infty} \int \pi(y,E)\mu_n(\mathrm{d}y) \stackrel{!}{=} \int \pi(y,E)\widehat{\mu}(\mathrm{d}y),$$

see (8.11). This proves that $\widehat{\mu}$ is stationary. So a stationary measure exists.

Suppose $\widetilde{\mu}$ is another stationary measure, and let $\{\widetilde{X}_n\}$ denote the Markov chain with initial distribution $\widetilde{\mu}$ and transition kernel π. This chain is stationary.

Hence $\widetilde{\mu}(\phi)=\mathbb{E}(\phi(\widetilde{X}_1))=\mathbb{E}(\phi(\widetilde{X}_n)) = \int \mathbb{E}(\phi(X_n)|X_1=y)\widetilde{\mu}(\mathrm{d}y)$ for all n. By (2.11),

$$\widetilde{\mu}(\phi) = \lim_{n\to\infty} \int \mathbb{E}(\phi(X_n)|X_1 = y)\widetilde{\mu}(\mathrm{d}y) = \lim_{n\to\infty} \int [\mu_n(\phi) + O(\theta^n)]\widetilde{\mu}(\mathrm{d}y) = \mu(\phi).$$

So the stationary measure is unique. □

We will now discuss the LLT for uniformly elliptic homogeneous Markov chains. Nagaev gave the first proof of this result (in the regime of local deviations). His proof is described in a special case, in the next section. Nagaev's proof is homogeneous in character; Here we will explain how to deduce the result from the inhomogeneous theory we developed in the previous chapters.

We will always assume that the chain is equipped with its unique stationary initial distribution, given by the previous lemma. Non-stationary uniformly elliptic homogeneous chains are "asymptotically homogeneous" in the sense of §8.5, and will be discussed there.

Theorem 8.9 *Let* f *denote an a.s. uniformly bounded homogeneous additive functional on a uniformly elliptic stationary homogeneous Markov chain* X.

(1) **Asymptotic Variance:** *The limit* $\sigma^2 = \lim_{N\to\infty} \frac{1}{N}\mathrm{Var}(S_N)$ *exists, and* $\sigma^2 = 0$ *iff* f *is the a.s. sum of a homogeneous gradient and a constant.*

(2) **CLT:** *If* $\sigma^2 > 0$, *then* $\frac{S_N - \mathbb{E}(S_N)}{\sqrt{N}}$ *converges in probability as* $N \to \infty$ *to the Gaussian distribution with mean zero and variance* σ^2.

(3) **LLT:** *If* $\sigma^2 > 0$ *then exactly one of the following options holds:*

(a) $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$. *In this case, if* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{N}} \to z$, *then for every interval* (a,b),

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1+o(1)]\frac{\mathrm{e}^{-z^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2 N}}(b-a), \text{ as } N\to\infty;$$

(b) $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$ *with* $t > 0$. *In this case, there are* $\kappa \in \mathbb{R}$ *and a bounded measurable function* $a : \mathfrak{S} \to \mathbb{R}$ *such that*

$$f(X_1,X_2) + a(X_1) - a(X_2) + \kappa \in t\mathbb{Z} \text{ a.s.}$$

Proof. Let $V_N := \mathrm{Var}(S_N)$ and $f_k := f(X_k, X_{k+1})$, and assume without loss of generality that $\mathbb{E}[f(X_1,X_2)] = 0$. By stationarity, $\mathbb{E}(f_n) = 0$ for all n.

Proof of Part (1): $V_N = \mathbb{E}(S_N^2) = \sum_{n=1}^{N} \mathbb{E}(f_n^2) + 2 \sum_{1\le m<n\le N} \mathbb{E}(f_n f_m)$. By stationarity, $\mathbb{E}(f_n f_m) = \mathbb{E}(f_0 f_{n-m})$, so

$$\frac{1}{N}V_N = \mathbb{E}(f_0^2) + 2\sum_{k=1}^{N-1} \mathbb{E}(f_0 f_k)(1 - \tfrac{k}{N}).$$

$|\mathbb{E}(f_0 f_m)|$ decays exponentially (Prop. 2.13), so $\sum |\mathbb{E}(f_0 f_k)| < \infty$, whence

$$\sigma^2 := \lim_{N\to\infty} \frac{1}{N}\mathrm{Var}(S_N) = \mathbb{E}(f_0^2) + 2\sum_{k=1}^{\infty} \mathbb{E}(f_0 f_k). \tag{8.12}$$

(This identity for σ^2 is called the **Green-Kubo formula**. Note that had $f_n(X_n, X_{n+1})$ been uncorrelated, then $\mathrm{Var}(S_N)$ would have been trivially equal to $N\mathbb{E}(f_0^2)$. The term $\sum_{k=1}^{\infty} \mathbb{E}(f_0 f_k)$ is the correction needed for the dependent case.)

Let u_n denote the structure constants of (X, f). The stationarity assumption implies that u_n is independent of n, say $u_n = u$ for all n. It follows that $U_N \equiv u_3^2 + \cdots + u_N^2 = (N-2)u^2$. Now we have two cases:

- $u > 0$: In this case by Theorem 3.7, $V_N \asymp U_N \asymp N$, whence $\sigma^2 > 0$.
- $u = 0$: In this case, $\mathrm{Var}(S_N) = O(1)$ by Theorem 3.7, whence $\sigma^2 = 0$ and f is center-tight. By the gradient lemma (Lemma 3.14),
$$f(X_1, X_2) = a_2(X_2) - a_1(X_1) + \kappa$$
for some $a_1, a_2 : \mathfrak{S} \to \mathbb{R}$ bounded and measurable and $\kappa \in \mathbb{R}$. In the homogeneous case, we may take $a_1 \equiv a_2$, see (3.6) in the proof of the gradient lemma. So $f(X_1, X_2) = a(X_2) - a(X_1) + \kappa$ a.s.

Part (2) follows from Dobrushin's CLT (Theorem 3.10).

Proof of Part (3): By stationarity, the structure constants $d_n(\xi)$ are independent of n, and they are all equal to $d(\xi) := \mathbb{E}(|e^{i\xi\Gamma} - 1|^2)^{1/2}$, where Γ is the balance of a random hexagon at position 3. So $D_N(\xi) = \sum_{k=3}^{N} d_k^2(\xi) = (N-2)d^2(\xi)$.

If $d(\xi) \neq 0$ for all $\xi \neq 0$, then $D_N(\xi) \to \infty$ for all $\xi \neq 0$, and $H(\mathsf{X}, \mathsf{f}) = \{0\}$. By Theorem 4.4. $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$ and f is irreducible. The non-lattice LLT now follows from Theorem 5.1.

If $d(\xi) = 0$ for some $\xi \neq 0$, then $D_N(\xi) = 0$ for all N, ξ is in the co-range of (X, f), and the reduction lemma says that there exist $\kappa_n \in \mathbb{R}$ and uniformly bounded measurable $a_n : \mathfrak{S} \to \mathbb{R}$ and $h_n(X_n, X_{n+1})$ such that $\sum h_n(X_n, X_{n+1})$ converges a.s., and $f(X_n, X_{n+1}) + a_n(X_n) - a_{n+1}(X_{n+1}) + h_n(X_n, X_{n+1}) + \kappa_n \in \frac{2\pi}{\xi}\mathbb{Z}$ a.s.

Let $A_n(X_n, X_{n+1}, \ldots) := a_n(X_n) + \sum_{k\geq n} h_k(X_k, X_{k+1})$. Then for all n

$$f_n(X_n, X_{n+1}) + A_n(X_n, X_{n+1}, \ldots) - A_{n+1}(X_{n+1}, X_{n+2}, \ldots) + \kappa_n \in \frac{2\pi}{\xi}\mathbb{Z} \text{ a.s.} \tag{8.13}$$

To finish the proof, we need to replace $A_i(X_i, X_{i+1}, \ldots)$ by functions of the form $a(X_i)$. This is the purpose of the following proposition:

Proposition 8.10 *Let* X *be a uniformly elliptic stationary homogeneous Markov chain with state space* $(\mathfrak{S}, \mathscr{B})$*. Let* $f : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ *be a measurable function such that* $\operatorname{ess\,sup} |f(X_1, X_2)| < \infty$*. If there are measurable functions* $A_n : \mathfrak{S}^{\mathbb{N}} \to \mathbb{R}$ *and* $\kappa_n \in \mathbb{R}$ *satisfying* (8.13)*, then there are* $\kappa \in \mathbb{R}$ *and a measurable* $a : \mathfrak{S} \to \mathbb{R}$ *s.t.*

$$f(X_n, X_{n+1}) + a(X_n) - a(X_{n+1}) + \kappa \in \frac{2\pi}{\xi}\mathbb{Z} \text{ a.s. for all } n.$$

Proof We assume for notational simplicity that $\xi = 2\pi$.

Let $\Omega := \mathfrak{S}^{\mathbb{N}}$, equipped with the σ-algebra $\mathscr{F}$ generated by the **cylinder sets**

$$[A_1, \ldots, A_n] := \{x \in \mathfrak{S}^{\mathbb{N}} : x_i \in A_i \ (i = 1, \ldots, n)\} \ \ (A_i \in \mathscr{B}(\mathfrak{S})).$$

Let m be the probability measure on $(\Omega, \mathscr{F})$ defined by

$$m[A_1, \ldots, A_n] = \mathbb{P}[X_1 \in A_1, \ldots, X_n \in A_n].$$

Let $\sigma : \Omega \to \Omega$ denote the left-shift map, $\sigma[(x_n)_{n\geq 1}] = (x_{n+1})_{n\geq 1}$. The stationarity of X translates to the shift invariance of m: $m \circ \sigma^{-1} = m$.

STEP 1 (**Zero-One Law**): *Let $\sigma^{-n}\mathscr{F} := \{\sigma^{-n}(A) : A \in \mathscr{F}\}$, then for every $A \in \bigcap_{n\geq 1} \sigma^{-n}\mathscr{F}$, either $m(A) = 0$ or $m(A) = 1$.*

Proof of the Step. Fix a cylinder $A := [A_1, \dots, A_\ell]$.

By uniform ellipticity, for every cylinder $B = [B_1, \dots, B_N]$,

$$m(A \cap \sigma^{-(\ell+1)}B) = m([A_1, \dots, A_\ell, *, B_1, \dots, B_N]) \geq \epsilon_0 m(A)m(B).$$

Applying this to the cylinders $B = [\mathfrak{S}, \dots, \mathfrak{S}, C_1, \dots, C_n]$, we find that

$$m(A \cap \sigma^{-(\ell+k)}[C_1, \dots, C_n]) \geq \epsilon_0 m(A)m[C_1, \dots, C_n] \text{ for all } k \geq 1.$$

Since this holds for all $C_i \in \mathscr{B}(\mathfrak{S})$, we have by the monotone class theorem that

$$m(A \cap \sigma^{-(\ell+k)}E) \geq \epsilon_0 m(A)m(E) \text{ for every } \mathscr{F}\text{–measurable } E \text{ and } k \geq 1. \tag{8.14}$$

Suppose $E \in \bigcap_{k\geq 1} \sigma^{-n}\mathscr{F}$, and let A be an arbitrary cylinder of length ℓ. By the assumption on E, $E = \sigma^{-n}E_n$ with $E_n \in \mathscr{F}$ and $n > \ell$. So

$$m(A \cap E) = m(A \cap \sigma^{-n}E_n) \geq \epsilon_0 m(A)m(E_n) = \epsilon_0 m(A)m(E).$$

We see that $m(E|A) \equiv \frac{m(A\cap E)}{m(A)} \geq \epsilon_0 m(E)$ for all cylinders A. So

$$\mathbb{E}(1_E|X_1, \dots, X_\ell) \geq \epsilon_0 m(E) \text{ for all } \ell,$$

and by the martingale convergence theorem, $1_E \geq \epsilon_0 m(E)$ a.e. So $m(E) = 0$ or 1.

STEP 2: *Identify f with a function $f : \Omega \to \mathbb{R}$ s.t. $f[(x_i)_{i\geq 1}] = f(x_1, x_2)$. Then there exist $A : \Omega \to \mathbb{R}$ measurable and $\kappa \in \mathbb{R}$ s.t. $f + A - A \circ \sigma + \kappa \in \mathbb{Z}$ almost surely.*

Proof of the Step. The assumptions of the proposition say that there exist $A_n : \Omega \to \mathbb{R}$ measurable and $\kappa_n \in \mathbb{R}$ s.t.

$$f \circ \sigma^n + A_n \circ \sigma^n - A_{n+1} \circ \sigma^{n+1} + \kappa_n \in \mathbb{Z} \ m\text{-a.e. for every } n.$$

Let $w_n := \mathrm{e}^{2\pi \mathrm{i} A_n}$ and $c_n := \mathrm{e}^{-2\pi \mathrm{i}\kappa_n}$, then $\mathrm{e}^{2\pi \mathrm{i} f\circ\sigma^n} \frac{w_n\circ\sigma^n}{w_{n+1}\circ\sigma^{n+1}} = c_n$ m-a.s. Since $m \circ \sigma^{-1} = m$, we get $\mathrm{e}^{2\pi \mathrm{i} f} \frac{w_n}{w_{n+1}\circ\sigma} = c_n$ m-almost everywhere. So

$$\begin{aligned} w_n &= c_n \mathrm{e}^{-2\pi \mathrm{i} f} w_{n+1} \circ \sigma = c_n c_{n+1} \mathrm{e}^{-2\pi \mathrm{i}(f + f\circ\sigma)} w_{n+2} \circ \sigma^2 \\ &= \cdots = c_n \cdots c_{n+k-1} \mathrm{e}^{-2\pi \mathrm{i} \sum_{j=0}^{k-1} f\circ\sigma^j} w_{n+k} \circ \sigma^k. \end{aligned}$$

Dividing the identities for w_n and w_{n+1} (with the same k), we obtain,

$$w_n/w_{n+1} = (c_n/c_{n+k})(w_{n+k}/w_{n+k+1}) \circ \sigma^k \text{ for all } k.$$

Hence w_n/w_{n+1} is $\sigma^{-k}\mathscr{F}$–measurable for all k. By the zero-one law, w_n/w_{n+1} is constant almost surely. In particular, there exists a constant c such that $A_2 - A_1 \in c + \mathbb{Z}$ m–a.e., and the step follows with $A := A_1$ and $\kappa := \kappa_1 - c$.

STEP 3: *There exists $a : \Omega \to \mathbb{R}$ constant on cylinders of length one such that $f + a - a \circ \sigma + \kappa \in \mathbb{Z}$ m-a.e.*

Proof of the Step. The **transfer operator** of $\sigma : \Omega \to \Omega$ is the operator $L : L^1(\Omega) \to L^1(\Omega)$ which describes the action of σ on mass densities on Ω: $\sigma_*[\varphi \mathrm{d}\mu] = L\varphi \mathrm{d}\mu$. Formally, $L\varphi := \frac{\mathrm{d}m_\varphi \circ \sigma^{-1}}{\mathrm{d}m}$, where $m_\varphi := \varphi \mathrm{d}m$. We will need the following facts:

(a) If φ depends only on the first m-coordinates, then $L\varphi$ depends only on the first $(m-1) \vee 1$–coordinates. Specifically, $(L\varphi)[(y_i)_{i\geq 1}] = \Phi(y_1, \dots, y_{m-1})$ where

$$\Phi(y_1,\ldots,y_{m-1}) := \mathbb{E}[\varphi(X_1,\ldots,X_m)|X_{i+1}=y_i\ (1\le i\le m-1)];$$

(b) $L\varphi$ is characterized by the condition $\int \psi L\varphi \mathrm{d}m = \int \psi\circ\sigma\varphi \mathrm{d}m\ \forall\psi\in L^\infty(\mathfrak{S})$;
(c) $L(\varphi\psi\circ\sigma)=\psi L\varphi\ \forall\varphi\in L^1,\psi\in L^\infty$;
(d) $L1=1$;
(e) $\forall\varphi\in L^\infty$, $L^n\varphi \xrightarrow[n\to\infty]{} \int\varphi\mathrm{d}m$ in L^1.

Part (b) is standard. Parts (c) and (d) follow from (b) and the σ-invariance of m. Part (a) follows from (b), and the following chain of identities:

$$\begin{aligned}\int \psi L\varphi \mathrm{d}m &= \int \psi \mathrm{d}m_\varphi\circ\sigma^{-1} = \int \psi\circ\sigma\,\varphi\mathrm{d}m = \mathbb{E}[\psi(X_2,X_3,\ldots)\varphi(X_1,\ldots,X_m)]\\ &= \mathbb{E}(\psi(X_2,X_3,\ldots)\mathbb{E}[\varphi(X_1,\ldots,X_m)|X_i,\ i\ge 2])\\ &\overset{!}{=} \mathbb{E}(\psi(X_2,X_3,\ldots)\mathbb{E}[\varphi|X_2,\ldots,X_m]) \overset{!!}{=} \int\psi\Phi\mathrm{d}m.\end{aligned}$$

The first marked equality is a standard calculation; the second uses stationarity.

Part (e) can be proved using the following argument of M. Lin. It is enough to consider $\varphi\in L^\infty$ such that $\int\varphi\mathrm{d}m=0$.

For such functions,

$$\begin{aligned}\|L^n\varphi\|_1 &= \int \mathrm{sgn}(L^n\varphi)L^n\varphi\,\mathrm{d}m = \int \mathrm{sgn}(L^n\varphi)\circ\sigma^n\cdot\varphi\,\mathrm{d}m\\ &= \int \mathrm{sgn}(L^n\varphi)\circ\sigma^n\mathbb{E}(\varphi|\sigma^{-n}\mathscr{F})\,\mathrm{d}m \le \int |\mathbb{E}(\varphi|\sigma^{-n}\mathscr{F})|\mathrm{d}m.\end{aligned}$$

$\mathbb{E}(\varphi|\sigma^{-n}\mathscr{F})$ is uniformly bounded (by $\|\varphi\|_\infty$), and by the martingale convergence theorem, it converges a.e. to

$$\mathbb{E}\left(\varphi\,\middle|\,\bigcap_{n=1}^{\infty}\sigma^{-n}\mathscr{F}\right) \overset{!}{=} \mathbb{E}(\varphi|\{\varnothing,\Omega\}) = \mathbb{E}(\varphi)=0.$$

The marked equality relies on the zero-one law.

Let $w:=\mathrm{e}^{2\pi\mathrm{i}A}$ where $A:\Omega\to\mathbb{R}$ is as in step 2, and assume w.l.o.g. that $\kappa=0$ (else absorb it into f). Set $S_n=f+f\circ\sigma+\cdots+f\circ\sigma^{n-1}$, then $\mathrm{e}^{-2\pi\mathrm{i}f}=w/w\circ\sigma$, whence $\mathrm{e}^{-2\pi\mathrm{i}S_n}=w/w\circ\sigma^n$. By (c) and (e), for all $\varphi\in L^1(\Omega)$,

$$wL^n(\mathrm{e}^{-2\pi\mathrm{i}S_n}\varphi) = L^n(\mathrm{e}^{-2\pi\mathrm{i}S_n}w\circ\sigma^n\varphi) = L^n(w\varphi) \xrightarrow[n\to\infty]{L^1} \int w\varphi\mathrm{d}m.$$

Since $|w|=1$ a.e., $\exists m\ge 2$ and $\exists\varphi=\varphi(x_1,\ldots,x_m)$ bounded measurable so that $\int w\varphi\mathrm{d}m\ne 0$. For this φ, we have

$$w^{-1} = \lim_{n\to\infty}\frac{L^n(\mathrm{e}^{-2\pi\mathrm{i}S_n}\varphi)}{\int w\varphi\mathrm{d}m} \text{ in } L^1.$$

We claim that the right-hand-side depends only on the first coordinate. This is because $\mathrm{e}^{-2\pi\mathrm{i}f}\varphi$ is function of the first m coordinates, whence by (a), $L(\mathrm{e}^{-2\pi\mathrm{i}f}\varphi)$ is a function of the first $(m-1)\vee 1$ coordinates. Applying this argument again we find that $L^2(\mathrm{e}^{-2\pi\mathrm{i}S_2}\varphi) = L[\mathrm{e}^{-2\pi\mathrm{i}f}L(\mathrm{e}^{-2\pi\mathrm{i}f}\varphi)]$ is a function of the first $(m-2)\vee 1$ coordinates. Continuing by induction, we find that $L^n(\mathrm{e}^{-2\pi\mathrm{i}S_n}\varphi)$ is a function of $(m-n)\vee 1$-coordinates, and eventually only of the first coordinate.

Since w^{-1} is an L^1-limit of a functions of the first coordinate, w is equal a.e. to a function of the first coordinate. Write $w[(x_i)_{i\ge1}]=\exp[2\pi\mathrm{i}a(x_1)]$ a.e.

By construction $\mathrm{e}^{2\pi\mathrm{i}f}w/w\circ\sigma=1$, so $f(X_1,X_2)+a(X_1)-a(X_2)\in\mathbb{Z}$ a.s. By stationarity, $f(X_n,X_{n+1})+a(X_n)-a(X_{n+1})\in\mathbb{Z}$ a.s. for all n. □

We now consider the LLT for large deviations. Given the work done in Chapter 7, what remains to be done is to determine the large deviations thresholds $\mathfrak{c}_\pm$.

Lemma 8.11 *Let* f *be an a.s. uniformly bounded homogeneous additive functional on a uniformly elliptic stationary homogeneous Markov chain, and assume the asymptotic variance* σ^2 *is strictly positive. Then the positivity thresholds* $\mathfrak{r}_\pm$ *satisfy*

$$\mathfrak{r}_- = \lim_{N\to\infty} \frac{\text{ess inf}[S_N - \mathbb{E}(S_N)]}{\sigma^2 N}\ ,\ \mathfrak{r}_+ := \lim_{N\to\infty} \frac{\text{ess sup}[S_N - \mathbb{E}(S_N)]}{\sigma^2 N}.$$

Proof Let $a_n := \text{ess sup}[S_n - \mathbb{E}(S_n)]$. By stationarity,

$$a_{n+m} \le a_n + a_m,$$

therefore $\lim(a_n/n)$ exists. By Theorem 8.9, $V_N \sim \sigma^2 N$, and the formula for $\mathfrak{r}_+$ follows. The formula for $\mathfrak{r}_-$ follows by symmetry, by considering $(\mathsf{X}, -\mathsf{f})$. □

Theorem 8.12 *Let* f *be an a.s. bounded homogeneous additive functional on a uniformly elliptic stationary homogeneous Markov chain. If* f *is* not *the a.s. sum of a homogeneous gradient and a constant, then* (X, f) *has full large deviations regime:* $\mathfrak{c}_\pm = \mathfrak{r}_\pm$.

Proof Let $f_n := f(X_n, X_{n+1})$. Subtracting a constant from f, we can arrange $\mathbb{E}[f(X_1, X_2)] = 0$. By stationarity, $\mathbb{E}(f_n) = 0$ and $\mathbb{E}(S_N) = 0$ for all n, N.

By the assumptions of the theorem, the asymptotic variance σ^2 is positive. Without loss of generality, $\sigma^2 = 1$. This can always be achieved by replacing f by f/σ.

We will prove that $\mathfrak{c}_+ = \mathfrak{r}_+$. The inequality $\mathfrak{c}_+ \le \mathfrak{r}_+$ is always true, so we focus on $\mathfrak{c}_+ \ge \mathfrak{r}_+$. By Lemma 8.11, for every $\varepsilon > 0$, for all sufficiently large M,

$$\delta_M := \mathbb{P}[S_M \ge (\mathfrak{r}_+ - \varepsilon)M] > 0.$$

Let $K := \text{ess sup}\,|\mathsf{f}|$. $S_{\ell(M+2)} = S_M + (f_{M+1} + f_{M+2}) + \sum_{k=(M+2)+1}^{\ell(M+2)} f_k$, so

$$\begin{aligned}
&\mathbb{P}[S_{\ell(M+2)} \ge \ell M(\mathfrak{r}_+ - \varepsilon) - 2\ell K] \\
&\ge \mathbb{P}\left[S_M \ge M(\mathfrak{r}_+ - \epsilon) \text{ and } \sum_{k=1}^{(\ell-1)(M+2)} f_{(M+2)+k} \ge (\ell-1)M(\mathfrak{r}_+ - \epsilon) - 2(\ell-1)K\right].
\end{aligned}$$

We now appeal to (8.14): Let $\sigma(X_i, \dots, X_j)$ denote the σ-field generated by $X_i, \dots, X_j$. If $E \in \sigma(X_1, \dots, X_{M+1})$ and $F \in \sigma(X_{M+3}, \dots, X_{\ell(M+2)+1})$, then $\mathbb{P}[E \cap F] \ge \epsilon_0 \mathbb{P}(E)\mathbb{P}(F)$. Thus by stationarity,

$$\begin{aligned}
&\mathbb{P}[S_{\ell(M+2)} \ge \ell M(\mathfrak{r}_+ - \varepsilon) - 2\ell K] \\
&\ge \epsilon_0 \delta_M \times \mathbb{P}[S_{(\ell-1)(M+2)} \ge (\ell-1)M(\mathfrak{r}_+ - \varepsilon) - 2(\ell-1)K] \\
&\ge (\epsilon_0\delta_M)^2 \times \mathbb{P}[S_{(\ell-2)(M+2)} \ge (\ell-2)M(\mathfrak{r}_+ - \varepsilon) - 2(\ell-2)K] \ge \cdots \ge (\epsilon_0\delta_M)^\ell.
\end{aligned}$$

Rearranging terms, we see that for all ℓ sufficiently large,

$$\mathbb{P}\left[S_{\ell(M+2)} \ge \ell(M+2)\left(\frac{M}{M+2}(\mathfrak{r}_+ - \varepsilon) - \frac{2K}{M+2}\right)\right] \ge [(\epsilon_0\delta_M)^{\frac{1}{M+2}}]^{\ell(M+2)}.$$

Recall that M is chosen after the choice of ε and K. So we may take M so large that $\frac{M}{M+2}\left((\mathfrak{r}_+ - \varepsilon) - \frac{2K}{M+2}\right) \ge \mathfrak{r}_+ - 2\varepsilon$. Let $\eta := (\epsilon_0\delta_M)^{\frac{2}{M+2}}$, then

$$\mathbb{P}\left[S_{\ell(M+2)} \geq \ell(M+2)(\mathfrak{r}_+ - 2\varepsilon)\right] \geq \eta^{\frac{1}{2}\ell(M+2)}.$$

Next, for all ℓ sufficiently large, for all $N \in [\ell(M+2), (\ell+1)(M+2))$, we have $S_N \geq S_{\ell(M+2)} - K(M+2)$. Therefore, for all N sufficiently large,

$$\mathbb{P}[S_N \geq N(\mathfrak{r}_+ - 3\varepsilon) + \varepsilon V_N] \geq \eta^{\frac{1}{2}N} \overset{!}{\geq} \eta^{V_N}, \text{ because } V_N \sim \sigma^2 N = N. \tag{8.15}$$

Theorem 7.26 says that $\mathfrak{c}_+ > 0$, so $\mathfrak{r}_+ \geq \mathfrak{c}_+ > 0$. If $\varepsilon < \mathfrak{r}_+/4$, then by the CLT,

$$\mathbb{P}[S_N \leq N(\mathfrak{r}_+ - 3\varepsilon) - \varepsilon V_N] \geq \mathbb{P}[S_N \leq 0] \sim \frac{1}{2} \geq \eta^{V_N}. \tag{8.16}$$

Looking at (8.15), (8.16) and Theorem 7.30, we deduce that $z_N := N(\mathfrak{r}_+ - 3\varepsilon)$ is admissible. By Theorem 7.8, $\mathbb{P}[S_N - z_N \in (a,b)]$ satisfy the LLT for large deviations, and by Theorem 7.26, $\mathfrak{r}_+ - 3\varepsilon = \lim_{N\to\infty} \frac{z_N - \mathbb{E}(S_N)}{V_N} \in [\mathfrak{c}_-, \mathfrak{c}_+]$. So $\mathfrak{r}_+ \leq \mathfrak{c}_+ + 3\varepsilon$. Since ε was arbitrary, $\mathfrak{c}_+ \geq \mathfrak{r}_+$. Thus $\mathfrak{c}_+ = \mathfrak{r}_+$.

By symmetry, $\mathfrak{c}_-(\mathsf{X},\mathsf{f}) = -\mathfrak{c}_+(\mathsf{X},-\mathsf{f}) = -\mathfrak{r}_+(\mathsf{X},-\mathsf{f}) = \mathfrak{r}_-(\mathsf{X},\mathsf{f})$. □

*8.4 One-Step Homogeneous Additive Functionals in L^2

In this work we focus on *bounded* functionals on *two-step* uniformly elliptic Markov chains. We will now deviate from this convention, and consider *unbounded* homogeneous additive functionals with finite variance, on stationary homogeneous Markov chains with the *one-step* uniform ellipticity condition:

$$\pi(x, \mathrm{d}y) = p(x,y)\mu(\mathrm{d}y)\,,\ \epsilon_0 \leq p(x,y) \leq \epsilon_0^{-1}\,,\ \mu(E) = \mathbb{P}[X_n \in E].$$

There is an obvious overlap with the setup of the previous section, but we will give a very different proof of the local limit theorem. This proof, due to Nagaev, is specific to the homogeneous case. But its ideas are of such importance, that we decided to include it, despite its definite homogeneous character.

For simplicity, we will restrict our attention to **one-step additive functionals** $f(x,y) = f(x)$, i.e. $S_N = f(X_1) + f(X_2) + \cdots + f(X_N)$. The "one-step" assumptions on X and f are not essential, but the one-step theory has special appeal, because it enables a more explicit characterization of the cases when $\sigma^2 = 0$ or $G_{ess}(\mathsf{X},\mathsf{f}) = t\mathbb{Z}$. Specifically, no gradient terms are needed, as in Theorem 8.9.

Theorem 8.13 (Nagaev) *Let* $\mathsf{f} : \mathfrak{S} \to \mathbb{R}$ *denote a one-step square integrable homogeneous additive functional, on a stationary homogeneous Markov chain* X *with the one-step ellipticity condition.*

(1) **Asymptotic Variance:** *The limit* $\sigma^2 = \lim\limits_{N\to\infty} \frac{1}{N}\mathrm{Var}(S_N)$ *exists, and* $\sigma^2 = 0$ *iff* f *is equal a.s. to a constant.*

(2) **CLT:** *If* $\sigma^2 > 0$*, then* $\frac{S_N - \mathbb{E}(S_N)}{\sqrt{N}}$ *converges in probability as* $N \to \infty$ *to the Gaussian distribution with mean zero and variance* σ^2.

(3) **LLT:** *If* $\sigma^2 > 0$*, then one of the two statements holds:*

(a) $\nexists c, t \in \mathbb{R}$ *such that* $f(X_1) \in c + t\mathbb{Z}$ *a.s. In this case, if* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{N}} \to z$*, then for every non-empty interval* (a,b)*,*

$$\mathbb{P}[S_N - z_N \in (a,b)] = [1 + o(1)]\frac{\mathrm{e}^{-z^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2 N}}(b-a), \text{ as } N \to \infty.$$

(b) $\exists t > 0$ *maximal and* $c \in \mathbb{R}$ *such that* $f(X_1) \in c + t\mathbb{Z}$ *a.s. In this case, if* $z_N \in c + t\mathbb{Z}$ *and* $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{N}} \to z$*, then for every* $k \in \mathbb{Z}$*,*

$$\mathbb{P}[S_N - z_N = kt] = [1 + o(1)]\frac{t\mathrm{e}^{-z^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2 N}}, \text{ as } N \to \infty.$$

The proof of this theorem is given in the following sections.

Asymptotic Variance and Irreducibility.

Proposition 8.14 *Suppose* X *is a stationary homogeneous Markov chain with the one-step ellipticity condition. Let* f *be a one-step square integrable homogeneous additive functional, and set* $S_n = f(X_1) + \cdots + f(X_n)$.

(1) The limit $\sigma^2 = \lim_{N\to\infty} \frac{1}{N}\mathrm{Var}(S_N)$ *exists, is finite, and*

$$\sigma^2 = \mathrm{Var}[f(X_1)] + 2\sum_{n=1}^{\infty} \mathrm{Cov}(f(X_1), f(X_{n+1})).$$

(2) $\sigma^2 = 0$ *iff* $f(X_1)$ *is constant almost surely.*

Proof Part (1) is proved exactly as in the case of bounded additive functionals, but using (2.13) instead of (2.11).

If $f(X_1)$ is constant, then $f(X_n)$ is constant for all n by stationary, and $\sigma^2 = 0$. We now suppose $\sigma^2 = 0$, and prove that $f(X_1)$ is constant a.e.

Without loss of generality, $\mathbb{E}[f(X_1)] = 0$ (otherwise subtract a constant from f).

Let $\mathfrak{S}$ denote the state space of X, equipped with the (stationary) initial distribution of X_1, and define an operator $\mathcal{L} : L^2(\mathfrak{S}) \to L^2(\mathfrak{S})$ by

$$(\mathcal{L}\varphi)(x) := \int_{\mathfrak{S}} \varphi(y)p(x,y)\mu(\mathrm{d}y) = \mathbb{E}[\varphi(X_2)|X_1 = x].$$

It is not difficult to see, by induction, that $(\mathcal{L}^n\varphi)(x) = \mathbb{E}[\varphi(X_{n+1})|X_1 = x]$. Let

$$\psi(x) := \sum_{n=0}^{\infty}(\mathcal{L}^n f)(x) = \sum_{n=0}^{\infty}\mathbb{E}[f(X_{n+1})|X_1 = x].$$

The sum converges in L^2 by (2.12).

Since $\mathcal{L}$ is bounded on $L^2(\mathfrak{S})$, $f(X_1) = \psi(X_1) - (\mathcal{L}\psi)(X_1)$ a.s. Therefore, if $\sigma^2 = 0$, then by part 1,

$$\begin{aligned}
0 = \sigma^2 &= \mathbb{E}\Big[f(X_1)^2 + 2f(X_1)\sum_{n=1}^{\infty} f(X_{n+1})\Big]\\
&= \mathbb{E}\Big(f(X_1)^2 + 2f(X_1)\sum_{n=1}^{\infty}\mathbb{E}[f(X_{n+1})|X_1 = x]\Big)\\
&= \mathbb{E}\left[((\psi - \mathcal{L}\psi)^2 + 2(\psi - \mathcal{L}\psi)\mathcal{L}\psi)(X_1)\right] = \mathbb{E}[(\psi - \mathcal{L}\psi)(\psi - \mathcal{L}\psi + 2\mathcal{L}\psi)(X_1)]\\
&= \mathbb{E}[(\psi^2 - (\mathcal{L}\psi)^2)(X_1)] = \mathbb{E}\left(\psi(X_1)^2 - \mathbb{E}[\psi(X_2)|X_1]^2\right)\\
&= \mathbb{E}\left(\psi(X_2)^2 - \mathbb{E}[\psi(X_2)|X_1]^2\right), \text{ by stationarity}\\
&= \mathbb{E}\left[\mathbb{E}\left(\psi(X_2)^2 - \mathbb{E}[\psi(X_2)|X_1]^2\Big|X_1\right)\right] = \mathbb{E}\left[\mathrm{Var}(\psi(X_2)|X_1)\right].
\end{aligned}$$

Necessarily, $\mathrm{Var}(\psi(X_2)|X_1) = 0$ a.s., whence $\psi(X_2) = \mathbb{E}(\psi(X_2)|X_1)$ almost surely. Recalling that $f = \psi - \mathcal{L}\psi$, we see that

$$f(X_1) = \psi(X_1) - (\mathcal{L}\psi)(X_1) = \psi(X_1) - \mathbb{E}[\psi(X_2)|X_1] = \psi(X_1) - \psi(X_2).$$

Rearranging terms, we find that

$$\psi(X_2) = \psi(X_1) - f(X_1) \text{ almost surely.}$$

We claim that $\psi(X_1)$ must be equal a.e. to a constant function. Assume the contrary, and choose some (t_0, s_0) in the support of the distribution of the random vector $(f(X_1), \psi(X_1))$. Then the support of $\psi(X_2)$ contains $s_0 - t_0$.

Since $\psi(X_1)$ is not constant a.e., and X is stationary, the support of the distribution of $\psi(X_2)$ contains more than one point. Choose some $s_1 \neq s_0 - t_0$ in this support.

Fix $0 < \varepsilon < \frac{1}{3}|(s_0 - t_0) - s_1|$. By the definition of the support of a measure, the following events have positive probability:

$$A := [|f(X_1) - t_0| < \varepsilon,\ |\psi(X_1) - s_0| < \varepsilon]\ ,\ B := [|\psi(X_2) - s_1| < \varepsilon]\ .$$

By the one-step ellipticity condition, $E := A \cap B$ has positive measure (bounded below by $\epsilon_0 \mathbb{P}(A)\mathbb{P}(B)$). But on E, $\psi(X_2) \neq \psi(X_1) - f(X_1)$, because

- $\psi(X_2)$ is ε-close to s_1,
- $\psi(X_1) - f(X_1)$ is 2ε-close to $s_0 - t_0$,
- $\text{dist}(s_1, t_0 - s_0) > 3\varepsilon$.

We obtain a contradiction to the a.s. equality $\psi(X_2) = \psi(X_1) - f(X_1)$.

Thus $\psi(X_1) = \text{const.}$. By stationarity, $\psi(X_2) = \text{const.}$, and $f(X_1) = 0$ a.s. □

Remark. If we replace the one-step ellipticity condition by a weaker condition which implies the convergence in norm of $\sum \mathcal{L}^n$ (e.g. uniform ellipticity), then we can only claim that $\sigma^2 = 0 \Leftrightarrow f(X_1) = \psi(X_1) - \psi(X_2) + \text{const.}$ for some $\psi \in L^2$.

Next, we calculate the co-range $H(\mathsf{X}, \mathsf{f}) := \{\xi \in \mathbb{R} : \sum d_n^2(\xi) < \infty\}$. Note that this is well-defined even when f is unbounded.

Lemma 8.15 *Suppose* X *is a stationary homogeneous Markov chain with the one-step ellipticity condition. Let* f *be a (possibly unbounded) homogeneous additive functional of the form* $f = f(x)$. *Then*

$$H(\mathsf{X}, \mathsf{f}) = \{0\} \cup \left\{ \frac{2\pi}{t} : t \neq 0 \quad \textit{and} \quad \exists c \in \mathbb{R} \textit{ s.t. } f(X_1) \in c + t\mathbb{Z} \textit{ a.s.} \right\}. \tag{8.17}$$

Proof It is clear that 0 belongs to both sides.

The inclusion $\supset$ is straightforward. To see $\subset$, suppose $\xi \in H(\mathsf{X}, \mathsf{f}) \setminus \{0\}$ and take $t := 2\pi/\xi$. In the stationary case, $d_n(\xi)$ are all equal to $d(\xi) = \mathbb{E}_{m_{\text{Hex}}}(|e^{\frac{2\pi i}{t}\Gamma} - 1|^2)$, where Hex is the space of position 3 hexagons and Γ is the balance of a random hexagon defined by (2.25). So

$$\xi \in H(\mathsf{X}, \mathsf{f}) \Leftrightarrow \sum_{n=3}^{\infty} d_n^2(\xi) < \infty \Leftrightarrow d(\xi) = 0 \Leftrightarrow \mathbb{E}_{m_{\text{Hex}}}(|e^{\frac{2\pi i}{t}\Gamma} - 1|^2) = 0.$$

So $\Gamma \in t\mathbb{Z}$ m_{Hex}-a.e. in Hex (n).

Fix t_0, s_0 in the support of the distribution of $f(X_1)$. By stationarity, t_0, s_0 are in the supports of the distributions of $f(X_n)$ for all n. So $\mathbb{P}[|f(X_n) - r| < \varepsilon] > 0$ for every $\varepsilon > 0$, $n \in \mathbb{N}$, and $r \in \{s_0, t_0\}$. By the one-step ellipticity condition,

$$m_{\text{Hex}} \left\{ \begin{pmatrix} y_1 & \begin{matrix} y_2 & x \\ y_4 & y_5 \end{matrix} & y_3 \end{pmatrix} \in \text{Hex}\,(n) : |f(y_i) - s_0| < \varepsilon\, , |f(x) - t_0| < \varepsilon \right\} > 0.$$

The balance of each hexagon in this set is 4ε-close to $t_0 - s_0$ (note that while in the case where f_n depend on two variables the balance of each hexagon contains six terms, but in the present case there only four terms since $f(y_1)$ cancels out). Since $\Gamma \in t\mathbb{Z}$ m_{Hex}-a.e., $\text{dist}(t_0 - s_0, t\mathbb{Z}) \leq 4\varepsilon$, and since ε can be chosen arbitrarily small,

$$t_0 \in s_0 + t\mathbb{Z}.$$

We now fix s_0 and take t_0 to be a general point in the support of the distribution of $f(X_1)$. The conclusion is that the support of the distribution of $f(X_1)$ is contained in $s_0 + t\mathbb{Z}$. Equivalently, $f(X_1) \in s_0 + t\mathbb{Z}$ almost surely. □

Perturbations of Linear Operators with Spectral Gap. Before continuing the proof of Theorem 8.13 we collect some definitions and facts from the theory of bounded linear operators. Let $(\mathfrak{X}, \|\cdot\|)$ denote a Banach space over $\mathbb{C}$.

- The **spectral radius** of a bounded linear operator L is $\rho(L) := \lim_{n\to\infty} \sqrt[n]{\|L^n\|}$.
- An **eigenvalue** of L is a number $\lambda \in \mathbb{C}$ such that $Lu = \lambda u$ for some non-zero $u \in \mathfrak{X}$. An eigenvalue is **simple** if $\dim\{u \in \mathfrak{X} : Lu = \lambda u\} = 1$.
- The **spectrum** of L is $\mathrm{spec}(L) := \{\lambda \in \mathbb{C} : \lambda I - L$ has no *bounded* inverse$\}$ (here and throughout, I denote the identity).

Every eigenvalue λ belongs to the spectrum (because $\ker(\lambda I - L) \neq 0$), but there could be points in $\mathrm{spec}(L)$ which are not eigenvalues (because $\lambda I - L$ could be invertible with an unbounded inverse).

Classical results in functional analysis say that the spectrum is always compact, non-empty, and

$$\rho(L) = \max\{|z| : z \in \mathrm{spec}(L)\}.$$

We will say that a bounded linear operator $L : \mathfrak{X} \to \mathfrak{X}$ has **spectral gap with simple leading eigenvalue λ**, if $\lambda \neq 0$ and $L = \lambda P + N$, where

(a) P and N are bounded linear operators such that $PN = NP = 0$;
(b) $P^2 = P$, $PL = LP = \lambda P$, and $\dim\{Pu : u \in \mathfrak{X}\} = 1$;
(c) $\rho(N) < |\lambda|$.

The operator P is called the **eigenprojection of λ**.

Lemma 8.16 *Suppose L has spectral gap with simple leading eigenvalue λ and eigenprojection P. Then:*

(1) $\exists 0 < \theta < 1$ such that $\|\lambda^{-n}L^n - P\| = O(\theta^n)$ as $n \to \infty$.
(2) λ is a simple eigenvalue of L, and $\mathrm{spec}(L) = K \cup \{\lambda\}$ where K is a compact subset of $\{z \in \mathbb{C} : |z| < |\lambda| - \gamma\}$ for some $\gamma > 0$ ("the gap").
(3) If L has spectral gap with simple leading eigenvalue λ' and eigenprojection P', then $\lambda' = \lambda$ and $P' = P$.

Proof $L^n = (\lambda P + N)^n = \lambda^n P + N^n$ (the mixed terms vanish and $P^n = P$ for all n). Thus, $\|\lambda^{-n}L^n - P\| = \|\lambda^{-n}N^n\|$. Since $\sqrt[n]{\|N^n\|} \to \rho(N) < |\lambda|$, $\|\lambda^{-n}L^n - P\| = O(\theta^n)$ for every $\rho(N)/|\lambda| < \theta < 1$.

The number λ is an eigenvalue, because $L(Pu) = \lambda Pu$ for every $u \in \mathfrak{X}$, and there are some u such that $Pu \neq 0$. It is a simple, because if $Lu = \lambda u$, then

$$u = \lambda^{-n}L^n u \to Pu,$$

whence u is in the image of P, a space of dimension one.

Since all eigenvalues are in the spectrum, $\lambda \in \mathrm{spec}(L)$. To finish the proof, we will show that $K := \mathrm{spec}(L)\setminus\{\lambda\}$ is contained in $\{z \in \mathbb{C} : |z| \leq \rho(N)\}$. It is sufficient to show that $zI - L$ has a bounded inverse for all $z \neq \lambda$ s.t. $|z| > \rho(N)$.

Let $\mathfrak{X}_1 := \ker(P) := \{u \in \mathfrak{X} : Pu = 0\}$ and let $\mathfrak{X}_2 := \mathrm{im}(P) := \{Pu : u \in \mathfrak{X}\}$. It is not difficult to see that

$$\mathfrak{X} = \mathfrak{X}_1 \oplus \mathfrak{X}_2 \text{ and } L(\mathfrak{X}_i) \subset \mathfrak{X}_i.$$

Indeed, the projections on $\mathfrak{X}_1, \mathfrak{X}_2$ are, respectively, $I - P$ and P.

Fix $z \neq \lambda$ such that $|z| > \rho(N)$.

- On $\mathfrak{X}_1$, $zI - L = zI - N$, and $\rho(L|_{\mathfrak{X}_1}) = \rho(N|_{\mathfrak{X}_1}) < |z|$. It follows that $z \notin \mathrm{spec}(L|_{\mathfrak{X}_1})$, and $(zI - L)|_{\mathfrak{X}_1}$ has a bounded inverse $(zI - N)^{-1}|_{\mathfrak{X}_1}$.
- On $\mathfrak{X}_2$, $zI - L = zI - \lambda I$, so $(zI - L)^{-1}|_{\mathfrak{X}_2} = (z - \lambda)^{-1}I$, a bounded linear operator.

Thus $(zI - L)^{-1} = (zI - N)^{-1}(I - P) + (z - \lambda)^{-1}P$, a bounded linear operator. □

The next result says that if we perturb a linear operator with spectral gap "smoothly," then the perturbed operator has spectral gap for small values of the perturbation. Moreover, the leading eigenvalue depends "smoothly" on the perturbation parameter. We begin by clarifying what we mean by "smooth."

Let $t \mapsto L_t$ be a function from a real open neighborhood U of zero, to the space of bounded linear operators on $\mathfrak{X}$.

- $t \mapsto L_t$ is **continuous** on U, if for all $t \in U$, $\|L_{t+h} - L_t\| \xrightarrow[h\to 0]{} 0$.
- $t \mapsto L_t$ is **differentiable** on U, if for all $t \in U$, there is a bounded linear operator L'_t (called the **derivative** of L_t at t) such that
$$\left\| \frac{L_{t+h} - L_t}{h} - L'_t \right\| \xrightarrow[h\to 0]{} 0.$$
If $t \mapsto L'_t$ is continuous on U, then we say that $t \mapsto L_t$ is $\boldsymbol{C^1}$**-smooth**, or just $\boldsymbol{C^1}$.
- By induction, $t \mapsto L_t$ is called $\boldsymbol{C^r}$**-smooth** if $t \mapsto L_t$ is differentiable on U, with C^{r-1}-smooth derivative. In this case, the $\boldsymbol{r^{\text{th}}}$**-derivative** of L_t is the bounded linear operator $L_t^{(r)}$ obtained inductively from $L_t^{(r)} := (L_t^{(r-1)})'$, $L_t^{(1)} := L'_t$.

Theorem 8.17 (Perturbation Theorem) *Fix $r \geq 1$ and $a > 0$. Suppose $L_t : \mathfrak{X} \to \mathfrak{X}$ is a bounded linear operator for each $|t| < a$, and $t \mapsto L_t$ is C^r-smooth with $r \geq 1$. If L_0 has spectral gap with simple leading eigenvalue λ_0 and eigenprojection P_0, then there exists some $0 < \kappa < a$ such that:*

(1) L_t has spectral gap with simple leading eigenvalue for each $|t| < \kappa$;
(2) The leading eigenvalue λ_t and eigenprojection P_t of L_t are C^r-smooth on $(-\kappa, \kappa)$;
(3) There exists $\gamma > 0$ such that $\rho(L_t - \lambda_t P_t) < |\lambda_t| - \gamma$ for all $|t| < \kappa$.

For the proof of this theorem, see Appendix C.

Nagaev's Perturbation Operators. We now return to the discussion of one-step additive functionals on one-step uniformly elliptic stationary homogeneous Markov chains.

Let $\mathfrak{X} := \{u : \mathfrak{S} \to \mathbb{C} : u \text{ is measurable, and } \|u\| := \sup_x |u(x)| < \infty\}$, and define a bounded linear operator $\mathcal{L}_t : \mathfrak{X} \to \mathfrak{X}$ by
$$(\mathcal{L}_t u)(x) := \int_{\mathfrak{S}} e^{itf(y)} u(y) p(x, y) \mu(dy) = \mathbb{E}(e^{itf(X_2)} u(X_2) | X_1 = x).$$

Lemma 8.18 *Let $S'_n := f(X_2) + \cdots + f(X_{n+1})$, then*
$$\mathbb{E}[e^{itS'_n} u(X_{n+1}) | X_1] = (\mathcal{L}_t^n u)(X_1) \quad (u \in \mathfrak{X}). \tag{8.18}$$

Proof This is a special case of Lemma 5.5, with $f_n(x, y) = f(y)$. □

Lemma 8.19 *If X is uniformly elliptic, then $\mathcal{L}_0$ has spectral gap with simple leading eigenvalue $\lambda_0 = 1$, and eigenprojection $P_0 u = \mathbb{E}[u(X_1)] 1_{\mathfrak{S}}$.*

Proof By definition, $(\mathcal{L}_0 u)(x) = \mathbb{E}[u(X_2) | X_1 = x]$. Let $P_0 u := \mathbb{E}(u(X_1)) 1_{\mathfrak{S}}$ and $N_0 := \mathcal{L}_0(I - P_0)$. These are bounded linear operators, and it is straightforward to verify that $\dim\{Pu : u \in \mathfrak{X}\} = 1$, $P_0^2 = P_0$, $P_0 \mathcal{L}_0 = \mathcal{L}_0 P_0 = P_0$, $\mathcal{L}_0 = P_0 + N_0$, and $N_0 \mathcal{L}_0 = \mathcal{L}_0 N_0$. It remains to check that $\rho(N_0) < 1$.

First, notice that $(I - P_0)^2 = I - 2P_0 + P_0^2 = I - P_0$. By induction, $(I - P_0)^n = I - P_0$. Since $\mathcal{L}_0$ commutes with P_0,
$$N_0^n = (\mathcal{L}_0(I - P_0))^n = \mathcal{L}_0^n (I - P_0)^n = \mathcal{L}_0^n (I - P_0).$$
For every $u \in \mathfrak{X}$, $(I - P_0)u = u - \mathbb{E}(u(X_1))$, and by stationarity, this is the same as $u - \mathbb{E}(u(X_{n+1}))$. This and (8.18) (with $t = 0$) leads to
$$\|N_0^n u\| = \|\mathbb{E}(u(X_{n+1}) - \mathbb{E}(u(X_{n+1})) | X_1 = x)\| \leq C_{mix} \theta^n \|u\|,$$
where C_{mix} and $0 < \theta < 1$ are constants which only depend on the ellipticity constant of X, see (2.11). So $\|N_0^n\| \leq 2C_{mix}\theta^n$, and $\rho(N_0) \leq \theta < 1$. □

Lemma 8.20 *If* X *has the one-step ellipticity condition and* $\mathbb{E}[f(X_1)^2] < \infty$*, then* $t \mapsto \mathcal{L}_t$ *is* C^2*-smooth on* $\mathbb{R}$*, and*

$$(\mathcal{L}_0' u)(x) = \mathrm{i}\mathbb{E}[f(X_2)u(X_2)|X_1 = x]\ ;\ (\mathcal{L}_0'' u)(x) = -\mathbb{E}[f(X_2)^2 u(X_2)|X_1 = x].$$

Proof Define an operator on $\mathfrak{X}$ by $(\mathcal{L}_t' u)(x) = \mathbb{E}[\mathrm{i}f(X_2)\mathrm{e}^{\mathrm{i}tf(X_2)}u(X_2)|X_1 = x]$. Clearly, $\|\mathcal{L}_t'\| \leq \mathbb{E}(|f(X_2)|) \leq \|f\|_2 < \infty$. For every $u \in \mathfrak{X}$, for every $x \in \mathfrak{S}$,

$$\left|\frac{1}{h}(\mathcal{L}_{t+h}u - \mathcal{L}_t u)(x) - (\mathcal{L}_t' u)(x)\right| = \left|\mathbb{E}\left(\frac{\mathrm{e}^{\mathrm{i}hf(X_2)} - 1 - \mathrm{i}hf(X_2)}{h}\mathrm{e}^{\mathrm{i}tf(X_2)}u(X_2)\middle|X_1 = x\right)\right|$$
$$\leq \mathbb{E}\left(\left|\frac{\mathrm{e}^{\mathrm{i}hf(X_2)} - 1 - \mathrm{i}hf(X_2)}{h^2 f(X_2)^2}hf(X_2)^2\right|1_{[f(X_2)\neq 0]}\middle|X_1 = x\right)\|u\| \leq M|h|\mathbb{E}(f(X_2)^2|X_1 = x)\|u\|,$$

where $M := \sup_{s\in\mathbb{R}\setminus\{0\}}\left|\frac{\mathrm{e}^{\mathrm{i}s} - 1 - \mathrm{i}s}{s^2}\right|$. By the one-step ellipticity condition, $\epsilon_0 \leq p(x, y) \leq \epsilon_0^{-1}$. So

$$\mathbb{E}(f^2(X_2)|X_1 = x) = \int p(x, y)f(y)^2\mu(\mathrm{d}y) \leq \epsilon_0^{-1}\int f(y)^2\mu(\mathrm{d}y) \leq \epsilon_0^{-1}\mathbb{E}(f^2).$$

It follows that $\left\|\frac{1}{h}(\mathcal{L}_{t+h} - \mathcal{L}_t) - \mathcal{L}_t'\right\| \leq M\epsilon_0^{-1}\|f\|_2^2|h| \xrightarrow[h\to 0]{} 0$. So $t \mapsto \mathcal{L}_t$ is differentiable, with derivative $\mathcal{L}_t'$.

Next we define the operator $\mathcal{L}_t'' u = -\mathbb{E}[f(X_2)^2\mathrm{e}^{\mathrm{i}tf(X_2)}u(X_2)|X_1 = x]$.

By one-step ellipticity, $\|\mathcal{L}_t''\| \leq \epsilon_0^{-2}\mathbb{E}(f(X_2)^2) = \|f\|_2^2 < \infty$. Fix $u \in \mathfrak{X}$. Then:

$$\left|\frac{1}{h}(\mathcal{L}_{t+h}' u - \mathcal{L}_t' u)(x) - (\mathcal{L}_t'' u)(x)\right| = \left|\mathbb{E}\left(\mathrm{i}f(X_2)\frac{\mathrm{e}^{\mathrm{i}hf(X_2)} - 1 - \mathrm{i}hf(X_2)}{h}\mathrm{e}^{\mathrm{i}tf(X_2)}u(X_2)|X_1 = x\right)\right|$$
$$\leq \mathbb{E}\left(f(X_2)^2 q(hf(X_2))|X_1 = x\right)\|u\|,\ \text{where } q(s) := \begin{cases}\left|\frac{\mathrm{e}^{\mathrm{i}s}-1-\mathrm{i}s}{s}\right| & s \neq 0\\ 0 & s = 0.\end{cases}$$

We now apply the one-step ellipticity condition as before, and deduce that

$$\left\|\frac{1}{h}(\mathcal{L}_{t+h}' - \mathcal{L}_t') - \mathcal{L}_t''\right\| \leq \epsilon_0^{-1}\mathbb{E}[f(X_2)q(hf(X_2))] \xrightarrow[h\to 0]{!} 0. \tag{8.19}$$

(To see (!) note that $q(hf(X_2)) \xrightarrow[h\to 0]{} 0$ pointwise, q is bounded, and $\mathbb{E}(|f(X_2)|) < \infty$.) It follows that $\mathcal{L}_t'$ is differentiable, with derivative $\mathcal{L}_t''$.

To finish the proof of C^2-smoothness, we check that $t \mapsto \mathcal{L}_t''$ is continuous.

$$(\mathcal{L}_{t+h}'' u - \mathcal{L}_t'' u)(x) = -\mathbb{E}\left(f(X_2)^2[\mathrm{e}^{\mathrm{i}hf(X_2)} - 1]\mathrm{e}^{\mathrm{i}tf(X_2)}u(X_2)\middle|X_1 = x\right).$$

As before, this leads to $\|\mathcal{L}_{t+h}'' - \mathcal{L}_t''\| \leq \epsilon_0^{-1}\mathbb{E}\left(f(X_2)^2\left|\mathrm{e}^{\mathrm{i}hf(X_2)} - 1\right|\right)$. By the dominated convergence theorem, $\left\|\mathcal{L}_{t+h}'' - \mathcal{L}_t''\right\| \xrightarrow[h\to 0]{} 0$. □

Proposition 8.21 *If* X *has the one-step ellipticity condition, and* $\mathbb{E}[f(X_1)^2] < \infty$*, then there exists* $\kappa > 0$ *such that:*

(1) For every $|t| < \kappa$*,* $\mathcal{L}_t$ *has spectral gap with simple leading eigenvalue* λ_t *and eigenprojection* P_t*.*
(2) $t \mapsto \lambda_t$ *and* $t \mapsto P_t$ *are* C^2 *on* $(-\kappa, \kappa)$*.*
(3) Let $N_t := \mathcal{L}_t - \lambda_t P_t$*, then there exists* $\gamma > 0$ *such that* $\rho(N_t) < |\lambda_t| - \gamma$ *for all* $|t| < \kappa$*.*
(4) $\lambda_0 = 1$*,* $\lambda_0' := \left.\frac{\mathrm{d}\lambda_t}{\mathrm{d}t}\right|_{t=0} = \mathrm{i}\mathbb{E}[f(X_1)]$*, and* $\lambda_0'' := \left.\frac{\mathrm{d}^2\lambda_t}{\mathrm{d}t^2}\right|_{t=0} = -\left[\sigma^2 + \mathbb{E}(f(X_1))^2\right]$*, where* $\sigma^2 = \lim_{N\to\infty}\frac{1}{N}\mathrm{Var}(S_N)$*.*
(5) Suppose $\sigma^2 > 0$ *and* $\mathbb{E}(f) = 0$*. Then there is a constant* $c > 0$ *such that for every* $|t| < \kappa$*,* $|\lambda_t| \leq \mathrm{e}^{-ct^2}$*.*

Proof Parts (1)–(3) follow from the perturbation theorem, and the previous lemmas.

We prove part (4). Given $|t| < \kappa$, let λ_t and P_t be the leading eigenvalue and eigenprojection for $\mathcal{L}_t$. By Lemmas 8.19 and 8.16(3),

$$\lambda_0 = 1 \quad \text{and} \quad P_0 u = \mathbb{E}[u(X_1)]1_{\mathfrak{S}}.$$

It is straightforward to show that if two operator-valued functions A_t, B_t are differentiable, then $(A_t + B_t)' = A_t' + B_t'$ and$(A_t B_t)' = A_t' B_t + A_t B_t'$.

Differentiating the identity $L_t P_t = \lambda_t P_t$, we obtain $L_t' P_t + L_t P_t' = \lambda_t' P_t + \lambda_t P_t'$. Multiplying both sides on the left by P_t gives

$$P_t L_t' P_t + \lambda_t P_t P_t' = \lambda_t' P_t + \lambda_t P_t P_t' \quad (\because\ P_t L_t = L_t P_t = \lambda_t P_t\ ,\ P_t^2 = P_t).$$

Therefore $P_t L_t' P_t = \lambda_t' P_t$. Substituting $t = 0$ and recalling the formulas for P_0 and L_0', we obtain that $\lambda_0' = \mathrm{i}\mathbb{E}[f]$.

Next, we differentiate both sides of the identity $\mathcal{L}_t^n P_t = \lambda_t^n P_t$ twice:

$$(\mathcal{L}_t^n)'' P_t + 2(\mathcal{L}_t^n)' P_t' + \mathcal{L}_t^n P_t'' = (\lambda_t^n)'' P_t + 2(\lambda_t^n)' P_t' + \lambda_t^n P_t''.$$

Now we multiply on the left by P_t, substitute $t = 0$, and cancel $P_0 \mathcal{L}_0^n P_0'' = \lambda_0^n P_0 P_0''$:

$$P_0 (\mathcal{L}_0^n)'' P_0 + 2 P_0 (\mathcal{L}_0^n)' P_0' = (\lambda_0^n)'' P_0 + 2(\lambda_0^n)' P_0 P_0'. \tag{8.20}$$

Recall that $\mathcal{L}_t^n u = \mathbb{E}(\mathrm{e}^{\mathrm{i}tS_n'} u(X_{n+1})|X_1)$, where $S_n' = \sum_{k=1}^{n} f_{k+1}(X_{k+1})$. One can prove exactly as in Lemma 8.20 that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathcal{L}_t^n u = \mathrm{i}\mathbb{E}(S_n' u(X_{n+1})|X_1)\ ,\quad \left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} \mathcal{L}_t^n u = -\mathbb{E}[(S_n')^2 u(X_{n+1})|X_1].$$

Also $(\lambda_0^n)' = n\lambda_0^{n-1}\lambda_0' = n\lambda_0' = n\mathbb{E}[f]$, and

$$(\lambda_0^n)'' = n(n-1)\lambda_0^{n-1}(\lambda_0')^2 + n\lambda_0^{n-1}\lambda_0'' = n(n-1)(\mathrm{i}\mathbb{E}[f])^2 + n\lambda_0''.$$

Substituting this in (8.20), we obtain (in the special case $u \equiv 1$),

$$-\mathbb{E}[(S_n')^2] + 2\mathbb{E}[S_n' P_0' 1] = n(n-1)(\mathrm{i}\mathbb{E}(f))^2 + n\lambda_0'' + 2n\mathbb{E}(f)\mathbb{E}[P_0' 1].$$

By exponential mixing and stationarity, $2\mathbb{E}[S_n' P_0' 1] = 2n\mathbb{E}(f)\mathbb{E}[P_0' 1] + O(1)$. Substituting this in the above, dividing by n, and passing to the limit, we obtain

$$\lambda_0'' = -\mathbb{E}(f)^2 - \lim_{n\to\infty} \frac{1}{n}\left(\mathbb{E}[(S_n')^2] - n^2\mathbb{E}(f)^2\right) = -\mathbb{E}(f)^2 - \lim_{n\to\infty}\frac{1}{n}\mathrm{Var}(S_n).$$

This proves part (4).

Suppose $\sigma^2 > 0$ and $\mathbb{E}(f) = 0$. Make κ so small that $|\lambda_t'' - \lambda_0''| < \frac{1}{2}\sigma^2$ for $|t| < \kappa$. Then

$$\begin{aligned}\lambda_t &= \lambda_0 + \int_0^t \left(\lambda_0' + \int_0^s \lambda_\eta'' \mathrm{d}\eta\right)\mathrm{d}s = \lambda_0 + \int_0^t \left(\lambda_0' + \int_0^s \left(\lambda_0'' + (\lambda_\eta'' - \lambda_0'')\right)\mathrm{d}\eta\right)\mathrm{d}s \\ &= 1 + t\lambda_0' + \frac{1}{2}\lambda_0'' t^2 + \int_0^t\!\!\int_0^s (\lambda_\eta'' - \lambda_0'')\mathrm{d}\eta\mathrm{d}s.\end{aligned}$$

Thus $\lambda_t = 1 - \frac{1}{2}\sigma^2 t^2 + \varepsilon(t)$, where $\varepsilon(t)$ is the double integral. Clearly, $|\varepsilon(t)| \le \frac{1}{4}\sigma^2 t^2$. To get part (5), we choose κ and c so that $1 - \frac{1}{4}\sigma^2 t^2 < \mathrm{e}^{-ct^2}$ for $|t| < \kappa$. □

Lemma 8.22 *For every* $[a, b] \subset \mathbb{R} \setminus H(\mathsf{X}, \mathsf{f})$ *there exists* $\gamma > 0$ *such that* $\rho(\mathcal{L}_t) < 1 - \gamma$ *for all* $t \in [a, b]$.

Proof Fix $[a,b] \subset \mathbb{R} \setminus H(\mathsf{X},\mathsf{f})$ and $t \in [a,b]$ (necessarily, $t \neq 0$). Recall the notation $\{z\}_{t\mathbb{Z}} = t\{z/t\}$, and define $g_t(x) := \{f(x)\}_{\frac{2\pi}{t}\mathbb{Z}}$. Then $e^{itg_t} = e^{itf}$. Hence $d_n(t,\mathsf{f}) = d_n(t,\mathsf{g}_t)$, and $(\mathcal{L}_t\varphi)(x) = \int e^{itg_t(y)}p(x,y)\varphi(y)\mu(dy)$.

Notice that ess sup $|\mathsf{g}_t| \leq |t| \leq \max(|a|,|b|)$. Therefore Lemma 5.6 applies, and there is a constant $\widetilde{\varepsilon}$, which depends only on a, b and the ellipticity constant of X, such that

$$\|\mathcal{L}_t^5\| \leq e^{-\widetilde{\varepsilon}d^2(t,\mathsf{g}_t)} = e^{-\widetilde{\varepsilon}d^2(t,\mathsf{f})} = e^{-\widetilde{\varepsilon}d^2(t)}.$$

Here $d^2(t)$ is the structure constant of f.

Since $t \notin H(\mathsf{X},\mathsf{f})$, $\sum d_n^2(t,\mathsf{f}) = \infty$. As $d_n(t,\mathsf{f}) = d(t)$ for all n, $d(t) > 0$. So

$$\rho(\mathcal{L}_t) \leq \exp\left[-\frac{1}{5}\widetilde{\varepsilon}d(t)^2\right] < 1 - \gamma_t \text{ for some } \gamma_t > 0.$$

Choose some n_t such that $\|\mathcal{L}_t^{n_t}\| < (1-\gamma_t)^{n_t}$. By the continuity of $t \mapsto \mathcal{L}_t$, there is an open neighborhood U_t of t, where

$$\|\mathcal{L}_s^{n_t}\| < (1-\gamma_t)^{n_t} \text{ for all } s \in U_t.$$

Recall that $\rho(\mathcal{L}_s) = \lim_{n\to\infty} \sqrt[n]{\|\mathcal{L}_s^n\|} \overset{!}{=} \inf \sqrt[n]{\|\mathcal{L}_s^n\|}$ (because $\|\mathcal{L}_s^{kn}\| \leq \|\mathcal{L}_s^k\|^n$). In particular, $\rho(\mathcal{L}_s) \leq \sqrt[n_t]{\|\mathcal{L}_s^{n_t}\|}$, and therefore $\rho(\mathcal{L}_s) \leq 1 - \gamma_t$ for all $s \in U_t$.

Since $[a,b]$ is compact, we can cover it by finitely many $U_{t_1}, \ldots, U_{t_n}$, and obtain the bound $\rho(\mathcal{L}_s) < 1 - \min\{\gamma_{t_1}, \ldots, \gamma_{t_n}\}$ on $[a,b]$. □

Proof of Theorem 8.13. Throughout this proof we assume that X is a stationary homogeneous Markov chain with the one-step ellipticity condition, and we let $f : \mathfrak{S} \to \mathbb{R}$ be a function such that $\mathbb{E}[f(X_1)^2] < \infty$.

We also suppose that $f(X_1)$ is not equal a.s. to a constant, and therefore by Proposition 8.14, $\frac{1}{N}\mathrm{Var}(S_N) \xrightarrow[N\to\infty]{} \sigma^2 \neq 0$.

Proposition 8.23 (CLT) *For every $a < b$,*

$$\mathbb{P}\left[\frac{S_N - \mathbb{E}(S_N)}{\sqrt{N}} \in (a,b)\right] \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^b e^{-t^2/2\sigma^2}dt.$$

Proof It is sufficient to prove the proposition in the special case when $\mathbb{E}[f(X_1)] = 0$. In this case, by Lévy's continuity theorem, it is sufficient to show that for each t,

$$\mathbb{E}\left(e^{itS_N/\sqrt{N}}\right) \to e^{-\frac{1}{2}\sigma^2t^2},$$

the characteristic function of the centered Gaussian distribution with variance σ^2.

Fix t and suppose n is large enough so that $|t|/\sqrt{n} < \kappa$, where $(-\kappa,\kappa)$ is the interval of perturbation parameters which satisfy the conclusions of Proposition 8.21. By stationarity and (8.18),

$$\begin{aligned}
\mathbb{E}\left(e^{itS_n/\sqrt{n}}\right) &= \mathbb{E}\left(e^{itS'_n/\sqrt{n}}\right) = \mathbb{E}[(\mathcal{L}_{t/\sqrt{n}}^n 1)(X_1)] = \mathbb{E}\left[(\lambda_{t/\sqrt{n}}P_{t/\sqrt{n}} + N_{t/\sqrt{n}})^n(X_1)\right] \\
&= \mathbb{E}\left[(\lambda_{t/\sqrt{n}}^n P_{t/\sqrt{n}}1 + N_{t/\sqrt{n}}^n 1)(X_1)\right] = \lambda_{t/\sqrt{n}}^n \mathbb{E}[(P_{t/\sqrt{n}}1)(X_1)] + O\left(\|N_{t/\sqrt{n}}^n 1\|\right) \\
&= \lambda_{t/\sqrt{n}}^n\left[1 + O\left(\|P_{t/\sqrt{n}} - P_0\|\right) + O\left(|\lambda_{t/\sqrt{n}}|^{-n}\|N_{t/\sqrt{n}}^n 1\|\right)\right].
\end{aligned}$$

Observe that $\|P_{t/\sqrt{n}} - P_0\| \to 0$ (because $s \mapsto P_s$ is C^2); and $\lambda_{t/\sqrt{n}}^{-n}\|N_{t/\sqrt{n}}^n 1\| \to 0$ (because $\rho(N_s) < |\lambda_s| - \gamma \leq (1-\gamma)|\lambda_s|$ for all $|s| < \kappa$). Thus

$$\mathbb{E}(e^{itS_n/\sqrt{n}}) = [1 + o(1)]\lambda_{t/\sqrt{n}}^n.$$

It remains to show that $\lim_{n\to\infty} \lambda^n_{t/\sqrt{n}} = e^{-\frac{1}{2}\sigma^2 t^2}$.

Recall that $s \mapsto \lambda_s$ is C^2, therefore $\lambda_t = \lambda_0 + t\lambda_0' + \frac{1}{2}\lambda_0'' t^2 + o(t^2) = 1 - \frac{1}{2}\sigma^2 t^2 + o(t^2)$, as $t \to 0$. So $\lambda^n_{t/\sqrt{n}} = (1 - \frac{t^2}{2n}(\sigma^2 + o(1)))^n \xrightarrow[n\to\infty]{!} e^{-\frac{1}{2}\sigma^2 t^2}$.

To justify the limit, calculate $\mathrm{Log}(\lambda^n_{t/\sqrt{n}})$ for some branch of the complex logarithm function, which is holomorphic on a complex neighborhood of 1. ◻

Proposition 8.24 (Non-Lattice LLT) *Suppose there are no $t, c \in \mathbb{R}$ such that $f(X_1) \in c + t\mathbb{Z}$ almost surely, then for every $a < b$ and $z_N, z \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{N}} \to z$, $\mathbb{P}[S_N - z_N \in (a,b)] = [1 + o(1)]\frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 N}}|a - b|$.*

Proof Without loss of generality, $\mathbb{E}(f(X_1)) = 0$, whence $\mathbb{E}(S_n) = 0$ for all n.

As we saw in §5.2.1, it is sufficient to show that

$$\frac{1}{2\pi}\int_{-L}^{L} e^{-i\xi z_n}\widehat{\phi}(\xi)\mathbb{E}(e^{i\xi S_n})d\xi \sim \frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 n}}\widehat{\phi}(0) \text{ as } n \to \infty, \tag{8.21}$$

for every $L > 0$, and each $\phi \in L^1(\mathbb{R})$ such that $\mathrm{supp}(\widehat{\phi}) \subset [-L, L]$.

Choose κ as in Prop. 8.21, and fix $R > 0$, arbitrarily large. We divide $[-L, L]$ into $\left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$, $[-\kappa,\kappa] \setminus \left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$, and $[-L, L] \setminus [-\kappa, \kappa]$, and consider the contribution to the integral from each of these regions.

Contribution of $\left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$: This is the part of the integral governed by the CLT.

$$\begin{aligned}\frac{1}{2\pi}\int_{-R/\sqrt{n}}^{R/\sqrt{n}} e^{-i\xi z_n}\widehat{\phi}(\xi)\mathbb{E}(e^{i\xi S_n})d\xi &= \frac{1}{2\pi\sqrt{n}}\int_{-R}^{R} e^{-i\xi\frac{z_n}{\sqrt{n}}}\widehat{\phi}\left(\frac{\xi}{\sqrt{n}}\right)\mathbb{E}(e^{i\xi\frac{S_n}{\sqrt{n}}})d\xi \\ \sim \frac{1}{2\pi\sqrt{n}}\int_{-R}^{R} e^{-i\xi z}\widehat{\phi}(0)e^{-\frac{1}{2}\sigma^2\xi^2}d\xi &= \frac{\widehat{\phi}(0)}{2\pi\sqrt{n}}\left(\int_{-\infty}^{\infty} e^{-i\xi z}e^{-\frac{1}{2}\sigma^2\xi^2}d\xi + o_{R\to\infty}(1)\right).\end{aligned}$$

So the contribution of $[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}]$ is $\dfrac{e^{-z^2/2\sigma^2}\widehat{\phi}(0)}{\sqrt{2\pi\sigma^2 n}}[1 + o_{n\to\infty}(1) + o_{R\to\infty}(1)]$.

Contribution of $[-\kappa,\kappa] \setminus \left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$: For ξ in this region, by Prop. 8.21(5),

$$\begin{aligned}&|\mathbb{E}(e^{i\xi S_n})| = |\mathbb{E}(e^{i\xi S_n'})| = |\mathbb{E}[(\mathcal{L}_\xi^n 1)(X_1)]| = |\mathbb{E}[(\lambda_\xi^n P_t 1 + N_\xi^n 1)(X_1)]| \\ &\le |\lambda_\xi|^n(\|P_\xi\| + |\lambda_\xi|^{-n}\|N_\xi^n\|) \le e^{-c\xi^2 n}[O(1) + o(1)] \le \text{const.}e^{-c\xi^2 n}.\end{aligned}$$

Therefore this contribution is $O(\|\widehat{\phi}\|_\infty \int_{R/\sqrt{n}}^{\kappa} e^{-cn\xi^2}d\xi) = \frac{1}{\sqrt{n}}o_{R\to\infty}(1)$.

Contribution of $[-L, L] \setminus [-\kappa, \kappa]$: By the assumptions of the proposition, there are no $c, t \in \mathbb{R}$ such that $f(X_1) \in c + t\mathbb{Z}$ a.s. By Proposition 8.15, $H(\mathsf{X}, \mathsf{f}) = \{0\}$.

By Lemma 8.22, there is $\gamma > 0$ such that $\rho(\mathcal{L}_\xi) < 1 - \gamma$ for all $\xi \in [-L, L] \setminus [-\kappa, \kappa]$. It follows that $\left|\mathbb{E}\left(e^{i\xi S_n}\right)\right| \le \left\|\mathcal{L}_\xi^n\right\| <$ const. $\left(1 - \frac{\gamma}{2}\right)^n$. Therefore, the contribution of $[-L, L] \setminus [-\kappa, \kappa]$ is bounded by

$$\frac{1}{2\pi}\int_{[\kappa \le |\xi| \le L]} |\widehat{\phi}(\xi)||\mathbb{E}(e^{i\xi S_n})|d\xi = O\left(\left(1 - \frac{\gamma}{2}\right)^n\right) = o\left(\frac{1}{\sqrt{n}}\right).$$

Collecting all these contributions, we find that for each $R > 1$,

$$\frac{1}{2\pi}\int_{-L}^{L} e^{-i\xi z_n}\widehat{\phi}(\xi)\mathbb{E}(e^{i\xi S_n})d\xi = \frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 n}}\widehat{\phi}(0)\,[1 + o_{R\to\infty}(1)] + o\left(\frac{1}{\sqrt{n}}\right).$$

Taking $R \to \infty$, we obtain (8.21). □

Proposition 8.25 (Lattice LLT) *Suppose $t \in \mathbb{R}$ is the maximal number such that for some c, $f(X_1) \in c + t\mathbb{Z}$ almost surely (necessarily $t \neq 0$). Then for every $z_N \in t\mathbb{Z}$ such that $\frac{z_N - \mathbb{E}(S_N)}{\sqrt{n}} \to z$, for every $k \in \mathbb{Z}$,* $\mathbb{P}[S_N - z_N = kt] = [1 + o(1)]\dfrac{t\mathrm{e}^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 N}}$.

Proof Without loss of generality, $\mathbb{E}(f(X_1)) = 0$ and $t = 1$. As we saw in §5.2.1, it is sufficient to show that for every $k \in \mathbb{Z}$ and $z_n \in \mathbb{Z}$ such that $\frac{z_n}{\sqrt{n}} \to z$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}tz_n}\mathrm{e}^{-\mathrm{i}\xi k}\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_n})\mathrm{d}\xi \sim \frac{\mathrm{e}^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 n}}. \tag{8.22}$$

To see (8.22), we take κ as in Proposition 8.21, split $[-\pi, \pi]$ into $[-R/\sqrt{n}, R/\sqrt{n}]$, $[-\kappa, \kappa] \setminus [-R/\sqrt{n}, R/\sqrt{n}]$, and $[-\pi, \pi] \setminus [-\kappa, \kappa]$, and consider the contribution of each of these regions to the integral.

Contribution of $\left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$: As in the non-lattice case,

$$\frac{1}{2\pi}\int_{-R/\sqrt{n}}^{R/\sqrt{n}} \mathrm{e}^{-\mathrm{i}\xi z_N}\mathrm{e}^{-\mathrm{i}\xi k}\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_n})\mathrm{d}\xi \sim \frac{\mathrm{e}^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2 n}}[1 + o_{R\to\infty}(1)].$$

Contribution of $[-\kappa, \kappa] \setminus \left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$: As in the lattice case, in this region, there exists $c' > 0$ such that

$$|\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_n})| \leq \text{const.}\mathrm{e}^{-c'\xi^2 n},$$

and therefore the contribution of $[-\kappa, \kappa] \setminus \left[-\frac{R}{\sqrt{n}}, \frac{R}{\sqrt{n}}\right]$ is bounded by $\frac{1}{\sqrt{n}} o_{R\to\infty}(1)$.

Contribution of $[-\pi, \pi] \setminus [-\kappa, \kappa]$: By assumption, $t = 1$ is the maximal t for which there is a constant c such that $f(X_1) \in c + t\mathbb{Z}$ a.s. Therefore, by Proposition 8.15, $H(\mathsf{X}, \mathsf{f}) = 2\pi\mathbb{Z}$, and $[-\pi, \pi] \setminus [-\kappa, \kappa]$ is a compact subset of $H(\mathsf{X}, \mathsf{f})^c$.

By Lemma 8.22, there is a $\gamma > 0$ such that

$$|\mathbb{E}(\mathrm{e}^{\mathrm{i}\xi S_n})| \leq \|\mathcal{L}_\xi^n\| < \text{const.}(1 - \gamma)^n \text{ for all } n \geq 1,\ \xi \in [-\pi, \pi] \setminus [-\kappa, \kappa].$$

It follows that the contribution of $[-\pi, \pi] \setminus [-\kappa, \kappa]$ is $o(1/\sqrt{n})$.
Summing these contributions, and passing to the limit as $R \to \infty$, we obtain (8.22). □

8.5 Asymptotically Homogeneous Markov Chains

A Markov chain $\widetilde{\mathsf{X}}$ is called **asymptotically homogeneous (with limit X)**, if it has state spaces $\mathfrak{S}_n = \mathfrak{S}$, and transition probabilities

$$\widetilde{\pi}_{n,n+1}(x, \mathrm{d}y) = [1 + \varepsilon_n(x, y)]\pi(x, \mathrm{d}y), \text{ where } \sup|\varepsilon_n| \xrightarrow[n\to\infty]{} 0,$$

and $\pi(x, \mathrm{d}y)$ is the transition probability of a homogeneous Markov chain X on $\mathfrak{S}$. An additive functional $\widetilde{\mathsf{f}} = \{\widetilde{f}_n\}_{n\geq 1}$ on $\widetilde{\mathsf{X}}$ is called **asymptotically homogeneous (with limit *f*)**, if

$$\sup|\widetilde{f}_n - f| \xrightarrow[n\to\infty]{} 0 \text{ for some } f : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}.$$

Clearly, all homogeneous Markov chains and homogeneous additive functionals are asymptotically homogeneous.

If $\widetilde{\mathsf{X}}$ is uniformly elliptic, then X must also be uniformly elliptic. Conversely, if X is uniformly elliptic, then some truncation $\widetilde{\mathsf{X}}_r := \{X_n\}_{n\geq r}$ is uniformly elliptic. To see this, take r so large that $\sup|\varepsilon_n| < \frac{1}{2}$ for all $n \geq r$.

Suppose $\widetilde{\mathsf{X}}$ is uniformly elliptic. Then X is uniformly elliptic, and by Lemma 8.8, X has a unique stationary measure μ. There is no loss of generality in assuming that X itself is stationary, otherwise we change the initial distribution to μ. By Corollary 2.9, we can write

$$\widetilde{\pi}_{n,n+1}(x, dy) = \widetilde{p}_n(x, y)\mu(dy) \, , \; \widetilde{p}_n(x, y) := [1 + \varepsilon_n(x, y)]p(x, y),$$

where $\mu(dx)$ is the stationary measure of X, $\sup|\varepsilon_n| \to 0$, and where for some constant $\epsilon_0 > 0$, $0 \leq p \leq \epsilon_0^{-1}$, and $\int p(x, y)p(y, z)\mu(dy) \geq \epsilon_0$.

Henceforth, we assume that $\widetilde{\mathsf{X}}$ is a uniformly elliptic asymptotically homogeneous Markov chain with a *stationary* limit X, and $\widetilde{\mathsf{f}}$ is a uniformly bounded asymptotically homogeneous additive functional on $\widetilde{\mathsf{X}}$, with limit f. Let $\mathsf{f} := \{f_n\}$, where $f_n = f$.

Theorem 8.26 *Let* $\mathsf{g} := \widetilde{\mathsf{f}} - \mathsf{f}$.

(1) If (X, g) *is center-tight, then* $G_{ess}(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = G_{ess}(\mathsf{X}, \mathsf{f})$.
(2) Otherwise, $G_{ess}(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \mathbb{R}$.
(3) In particular, $G_{ess}(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \mathbb{R}$ *whenever it is* not true *that for some bounded measurable function* $a(x)$ *and* $c, t \in \mathbb{R}$, $f(X_1, X_2) + a(X_1) - a(X_2) + c \in t\mathbb{Z}$ *a.s. with respect to the stationary law of* X. *In this case,* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ *satisfies the non-lattice LLT* (5.1).

Proof Let $\widetilde{\mathsf{X}}_r := \{\widetilde{X}_n\}_{n\geq r}$ and $\widetilde{\mathsf{f}}_r := \{\widetilde{f}_n\}_{n\geq r}$. Similarly, define $\mathsf{X}_r, \mathsf{f}_r$. Discarding a finite number of terms does not change the essential range (since any functional which is identically zero for large n is center-tight). Therefore

$$G_{ess}(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = G_{ess}(\widetilde{\mathsf{X}}_r, \widetilde{\mathsf{f}}_r) \text{ and } G_{ess}(\mathsf{X}_r, \widetilde{\mathsf{f}}_r) = G_{ess}(\mathsf{X}, \widetilde{\mathsf{f}}).$$

Pick r so large, that $\dfrac{1}{2} \leq \dfrac{\widetilde{p}_n(x, y)}{p(x, y)} \leq 2$ for all $n \geq r$. Then $\widetilde{\mathsf{X}}_r$ and X_r are related by a change of measure with bounded weights, and therefore (by Example 4.12),

$$G_{ess}(\widetilde{\mathsf{X}}_r, \widetilde{\mathsf{f}}_r) = G_{ess}(\mathsf{X}_r, \widetilde{\mathsf{f}}_r).$$

It follows that $G_{ess}(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = G_{ess}(\mathsf{X}, \widetilde{\mathsf{f}})$.

If g is center-tight then $G_{ess}(\mathsf{X}, \widetilde{\mathsf{f}}) = G_{ess}(\mathsf{X}, \mathsf{f} + \mathsf{g}) = G_{ess}(\mathsf{X}, \mathsf{f})$, and we obtain the first part of the theorem.

Now suppose that g is not center-tight. By Theorem 4.4, to see that $G_{ess}(\mathsf{X}, \widetilde{\mathsf{f}}) = \mathbb{R}$, it is sufficient to show that $H(\mathsf{X}, \widetilde{\mathsf{f}}) = \{0\}$. Equivalently:

$$D_N(\xi, \widetilde{\mathsf{f}}) \xrightarrow[N\to\infty]{} \infty \text{ for all } \xi \neq 0,$$

where D_N are the structure constants associated to X, see §2.3.2.

Recall also the structure constants $d_n(\xi, \mathsf{f})$. Since X is stationary, $\{X_n\}_{n\geq 1}$ is a stationary stochastic process, and $d_n(\xi, \mathsf{f})$ are all equal. Call their common value $\mathfrak{d}(\xi)$. By Lemma 2.16(2) we have

$$\mathfrak{d}^2(\xi) = d_n^2(\xi, \mathsf{f}) \leq 8\left[d_n(\xi, \widetilde{\mathsf{f}})^2 + d_n(\xi, \mathsf{g})^2\right].$$

Asymptotic homogeneity says that $\|g_n\|_\infty \xrightarrow[n\to\infty]{} 0$, therefore $d_n^2(\xi, \mathsf{g}) \xrightarrow[n\to\infty]{} 0$. So for some $n_0 = n_0(\xi)$, for every $n \geq n_0$, $d_n^2(\xi, \widetilde{\mathsf{f}}) \geq \dfrac{\mathfrak{d}^2(\xi)}{10}$. Hence

$$D_N^2(\xi, \widetilde{\mathsf{f}}) \equiv \sum_{n=3}^{N} d_n^2(\xi, \widetilde{\mathsf{f}}) \geq \frac{(N - n_0)\mathfrak{d}^2(\xi)}{10} \to \infty, \text{ whenever } \mathfrak{d}(\xi) \neq 0.$$

Next we consider the case when $\mathfrak{d}(\xi) = 0$ and $\xi \neq 0$. In this case for a.e. hexagon $P \in \text{Hex}(n)$, $e^{i\xi\Gamma(\mathsf{f},P)} = 1$, where $\Gamma(\mathsf{f},\cdot)$ denotes the balance for f. Hence

$$e^{i\xi\Gamma(\widetilde{\mathsf{f}},\cdot)} = e^{i\xi\Gamma(\mathsf{g},\cdot)},$$

and so $d_n(\xi,\widetilde{\mathsf{f}}) = d_n(\xi,\mathsf{g})$.

Let $\gamma_N := \max_{n\geq N} \text{ess sup}\,|g_n|$, and fix $\tau_0 > 0$ s.t. $|e^{it} - 1|^2 \geq \frac{1}{2}t^2$ for all $|t| < \tau_0$. If $0 < |\eta| \leq \tau_0(6\gamma_N)^{-1}$, then (4.14) tells us that for every $n \geq N$,

$$d_n^2(\eta,\mathsf{g}) \geq \frac{\eta^2}{2}u_n^2(\mathsf{g}) \text{ for all } n > N+3.$$

By assumption, g is not center-tight, so $\sum u_n^2(\mathsf{g}) = \infty$. It follows that $D_N(\eta,\mathsf{g}) \to \infty$ for all $0 < |\eta| \leq \tau_0(6\gamma_N)^{-1}$.

By assumption, $\gamma_N \to 0$, so $D_N(\xi,\mathsf{g}) \to \infty$, whence

$$D_N^2(\xi,\widetilde{\mathsf{f}}) = \sum_{n=3}^{N} d_n^2(\xi,\widetilde{\mathsf{f}}) = \sum_{n=3}^{N} d_n^2(\xi,\mathsf{g}) = D_N^2(\xi,\mathsf{g}) \to \infty$$

also when $\mathfrak{d}(\xi) = 0$ (but $\xi \neq 0$). This completes the proof that $G_{ess}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = \mathbb{R}$, whenever g is not center-tight.

We proved parts (1) and (2) of the theorem. We now prove part (3). Suppose f is not a.s. a homogeneous gradient plus a constant, modulo some group $t\mathbb{Z}$. By Theorem 8.9, $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$. If $\mathsf{g} := \widetilde{\mathsf{f}} - \mathsf{f}$ is center-tight, then $G_{ess}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, by the first part of the theorem. If g is not center-tight, then $G_{ess}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = \mathbb{R}$ by the second part of the theorem. In both cases $G_{ess}(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = \mathbb{R}$. □

Lemma 8.27 *If* f *is not the sum of a homogeneous gradient and a constant, then the variance of* $S_N(\widetilde{\mathsf{f}})$ *with respect to* $\widetilde{\mathsf{X}}$ *satisfies* $\text{Var}[S_N(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})] \asymp N$ *as* $N \to \infty$.

Proof Choose r so large that $\sup|\varepsilon_n(x,y)| < \frac{1}{2}$ for all $n \geq r$. Then for $n \geq r+3$, the hexagon measures of $\widetilde{\mathsf{X}}$ and X on Hex (n) differ by a density uniformly bounded away from zero and infinity (see Example 4.12). It follows that

$$u_n^2(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) \asymp u_n^2(\mathsf{X},\widetilde{\mathsf{f}}).$$

Next, by the assumption $\sup|\widetilde{f}_n - f_n| \xrightarrow[n\to\infty]{} 0$, $|u_n^2(\mathsf{X},\widetilde{\mathsf{f}}) - u_n^2(\mathsf{X},\mathsf{f})| \to 0$. Therefore

$$\frac{1}{N}\sum_{n=3}^{N} u_n^2(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) \asymp \frac{1}{N}\sum_{n=3}^{N} u_n^2(\mathsf{X},\widetilde{\mathsf{f}}) \asymp \frac{1}{N}\sum_{n=3}^{N} u_n^2(\mathsf{X},\mathsf{f}) + o(1).$$

By Theorem 3.6, $\frac{1}{N}\sum_{n=3}^{N} u_n^2(\mathsf{X},\mathsf{f}) \asymp \frac{1}{N}[V_N + O(1)]$, where V_N is the variance of $S_N(\mathsf{f})$ with respect to the (stationary homogeneous) limit Markov chain X.

By Theorem 8.9(1) and the assumption that f is not a homogeneous gradient plus a constant, $V_N \sim \sigma^2 N$ where $\sigma^2 \neq 0$. So $\frac{1}{N}\sum_{n=3}^{N} u_n^2(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) \asymp 1$.

The lemma now follows from the variance estimates in Theorem 3.6. □

Next we discuss the large deviation thresholds for $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$ (see §7.4).

Theorem 8.28

(a) If f *is not a homogeneous gradient plus a constant, then* $(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}})$ *has full large deviations regime, and* $\mathfrak{c}_\pm(\widetilde{\mathsf{X}},\widetilde{\mathsf{f}}) = \mathfrak{c}_\pm(\mathsf{X},\mathsf{f}) = \mathfrak{r}_\pm(\mathsf{X},\mathsf{f})$.

(b) If f *is a homogeneous gradient plus a constant, and* $\widetilde{\mathsf{f}}$ *is not center-tight, then* $\mathfrak{c}_+(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = +\infty$, $\mathfrak{c}_-(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = -\infty$.

Proof Suppose f is not a homogeneous gradient plus a constant. By the previous lemma, $\widetilde{V}_N := \mathrm{Var}(\widetilde{\mathsf{X}}, S_N(\widetilde{\mathsf{f}})) \asymp N$, and therefore

$$\frac{\|S_N(\widetilde{\mathsf{f}}) - S_N(\mathsf{f})\|_\infty}{\widetilde{V}_N} = O\Big(\frac{1}{N}\sum_{n=1}^{N}\|\widetilde{f}_n - f_n\|_\infty\Big) \xrightarrow[N\to\infty]{} 0.$$

By Lemma 7.34, $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ and $(\widetilde{\mathsf{X}}, \mathsf{f})$ have the same admissible sequences. So

$$\mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \mathsf{f}).$$

Given r, let $\widetilde{\mathsf{X}}_r := \{\widetilde{X}_n\}_{n\ge r}$ and $\mathsf{X}_r := \{X_n\}_{n\ge r}$. It is not difficult to see, using Theorem 7.26 with intervals much larger than r ess sup $|f|$, that

$$\mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \mathsf{f}) = \mathfrak{c}_\pm(\widetilde{\mathsf{X}}_r, \mathsf{f}_r) \text{ and } \mathfrak{c}_\pm(\mathsf{X}, \mathsf{f}) = \mathfrak{c}_\pm(\mathsf{X}_r, \mathsf{f}_r).$$

By asymptotic homogeneity, we can choose r so large so that $\sup|\varepsilon_n(x, y)| < \frac{1}{2}$ for all $n \ge r$. In this case $\widetilde{\mathsf{X}}_r$ and X_r are related by a change of measure with bounded weights. By Lemma 7.33, $\mathfrak{c}_\pm(\widetilde{\mathsf{X}}_r, \mathsf{f}_r) = \mathfrak{c}_\pm(\mathsf{X}_r, \mathsf{f}_r)$.

In summary, $\mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \mathsf{f}) = \mathfrak{c}_\pm(\widetilde{\mathsf{X}}_r, \mathsf{f}_r) = \mathfrak{c}_\pm(\mathsf{X}_r, \mathsf{f}_r) = \mathfrak{c}_\pm(\mathsf{X}, \mathsf{f})$. Similarly, one shows that $\mathfrak{r}_\pm(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \mathfrak{r}_\pm(\mathsf{X}, \mathsf{f})$. By Theorem 8.12, $\mathfrak{c}_\pm(\mathsf{X}, \mathsf{f}) = \mathfrak{r}_\pm(\mathsf{X}, \mathsf{f})$, and the proof of part (a) is complete.

In the proof of part (b) we may assume that $\mathsf{f} = 0$ and $\mathbb{E}(\widetilde{f}_n) = 0$ for all n, since adding a homogeneous gradient and centering does not change $\mathfrak{c}_\pm$.

Write $S_N = S_N(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$, $\quad S_{n_1,n_2} = \sum_{k=n_1}^{n_2-1} \widetilde{f}_k(\widetilde{X}_k, \widetilde{X}_{k+1})$, $\quad \widetilde{V}_N = \mathrm{Var}(S_N(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}))$. Since $\widetilde{\mathsf{f}}$ is not center-tight, $\widetilde{V}_N \to \infty$.

Divide the interval $[0, N]$ into blocks

$$[n_1, n_2] \cup \{n_2 + 1\} \cup [n_3, n_4] \cup \cdots \cup [n_{k_N}, n_{k_N+1}] \cup \{n_{k_N+1} + 1\} \cup [n_{k_N+2}, N], \tag{8.23}$$

where for i odd, n_{i+1} is the minimal $n > n_i + 1$ such that $\mathrm{Var}(S_{n_i,n}) \ge 1$, and for i even, $n_{i+1} := n_i + 2$. We denote the maximal odd i with $n_i + 1 \le N$ by k_N.

Since $\|\widetilde{f}_n\|_\infty \equiv \|\widetilde{f}_n - f_n\|_\infty \to 0$, we have $\lim\limits_{\ell\to\infty}\lim\limits_{N\to\infty}\min\{n_{j+1} - n_j : \ell \le j \le k_N \text{ odd}\} = \infty$. From the identity $\mathrm{Var}(S_{n_i,n+1}) = \mathrm{Var}(S_{n_i,n}) + \mathrm{Var}(\widetilde{f}_n) + 2\mathrm{Cov}(S_{n_i,n}, \widetilde{f}_n)$ and the mixing estimate (2.13), we see that there is $M_0 > 1$ s.t.

$$1 \le \mathrm{Var}(S_{n_i,n_{i+1}}) \le M_0 \text{ for odd } i \le k_N, \text{ and } \mathrm{Var}(S_{n_{k_N+2},N}) \le M_0.$$

Similarly, as $\|\widetilde{f}_n\|_\infty \to 0$, $\lim\limits_{j\to\infty \text{ odd}} \mathrm{Var}(S_{n_j,n_{j+1}}) = 1$ and $\lim\limits_{j\to\infty \text{ odd}} \mathrm{Var}(S_{n_j,n_{j+1}+1}) = 1$.

Next we claim that

$$\lim_{N\to\infty}\frac{1}{\widetilde{V}_N}\sum_{j\le k_N \text{ odd}} \mathrm{Var}(S_{n_j,n_{j+1}+1}) = 1. \tag{8.24}$$

Clearly, $\widetilde{V}_N = \sum\limits_{j\le k_N+2 \text{ odd}} \mathrm{Var}(S_{n_j,n_{j+1}+1}) + 2\sum\limits_{i<j\le k_N+2 \text{ odd}} \mathrm{Cov}(S_{n_i,n_{i+1}+1}, S_{n_j,n_{j+1}+1})$ (with the convention that $n_{k_N+3} := N - 1$). Since $\mathrm{Var}(S_{n_{k_N+2},N}) \le M_0$ and $\widetilde{V}_N \to \infty$, to prove (8.24) it suffices to show that

$$\lim_{j\to\infty}\sum_{i=1}^{j-1}\Big|\mathrm{Cov}(S_{n_i,n_{i+1}+1}, S_{n_j,n_{j+1}+1})\Big| = 0. \tag{8.25}$$

By Proposition 2.13, the LHS of (8.25) is at most

$$\sum_{p\le n_j, q\ge n_j} C_{mix}\theta^{q-p}\|f_p\|_\infty\|f_q\|_\infty \le \sum_{p\le 0, q\ge 0} C_{mix}\theta^{q-p}\sup_{p\in\mathbb{N}}\|f_p\|_\infty \sup_{q\ge n_j}\|f_q\|_\infty,$$

which tends to 0, since the last term tends to 0. This proves (8.25) and hence (8.24).

Let β_N denote the number of blocks in the decomposition (8.23). Since $\lim\limits_{j\to\infty \text{ odd}} \mathrm{Var}(S_{n_j,n_{j+1}+1}) = 1$, (8.24) implies that for large N,

$$\widetilde{V}_N/2 \le \beta_N \le 2\widetilde{V}_N.$$

Let $M_j = \max\limits_{n_j \le l \le n_{j+1}} \|\widetilde{f}_l\|_\infty$, then $M_j \to 0$. Applying Dobrushin's CLT to the array

$$\widetilde{f}_l/M_j, \quad n_j \le l \le n_{j+1}, \quad j \text{ is odd}, \quad j \ge 1,$$

we obtain that $S_{n_j,n_{j+1}}/\sqrt{\mathrm{Var}(S_{n_j,n_{j+1}})}$ is asymptotically normal. So for each $z > 0$ there exists $\eta(z) > 0$ such that for all j large enough, except perhaps the last one,

$$\mathbb{P}(S_{n_j,n_{j+1}} \ge 3z) \ge \eta(z). \tag{8.26}$$

For $j = k_N + 2$ (the last one), $\mathbb{P}[|S_{j,N}| \le 2] \ge \frac{3}{4}$, by the Chebyshev inequality.

An ellipticity argument similar to the one we used in the proof of Theorem 8.12 now shows that as $N \to \infty$, $\mathbb{P}(S_N \ge \beta_N z) \ge c\epsilon_0^{\beta_N}\eta^{\beta_N}$.

(The constant c incorporates the contribution of the blocks (with small j or $j = k + 2$) where (8.26) fails.)

Recalling that $\widetilde{V}_N \le 2\beta_N$, we obtain $\mathbb{P}(S_N \ge \widetilde{V}_N z) \ge c\epsilon_0^{\beta_N}\eta(2z)^{\beta_N}$. Next, by the CLT for S_N, $\mathbb{P}[S_N \le \widetilde{V}_N z] \ge \mathbb{P}[S_N \le 0] \to \frac{1}{2}$. Theorem 7.30 now tells us that

$$z_N := (z - \varepsilon)\widetilde{V}_N \text{ is admissible for each } \varepsilon > 0.$$

It follows that $z - \varepsilon := \lim \frac{z_N - \mathbb{E}(S_N)}{\widetilde{V}_N}$ is inside $(\mathfrak{c}_-, \mathfrak{c}_+)$ for each $\varepsilon > 0$, whence $\mathfrak{c}^+(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) \ge z$. Since z is arbitrary, $\mathfrak{c}^+(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = +\infty$. Similarly, $\mathfrak{c}_-(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = -\infty$. □

It is worthwhile to spell out the results of the present section in the special case when the limit f is identically equal to zero.

Corollary 8.29 (Asymptotically Negligible Functionals) *Suppose* $\sup|\widetilde{f}_n| \to 0$ *and* $\mathbb{E}(\widetilde{f}_n) = 0$. *Then:*

- *Either* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ *is center-tight, and* $\sum\limits_{n=1}^{\infty} \widetilde{f}_n$ *converges almost surely;*
- *or* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ *is not center-tight, satisfies the non-lattice LLT* (5.1) *and* $(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})$ *has full large deviations regime with thresholds* $\mathfrak{c}_\pm(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}}) = \pm\infty$.

Proof The non center-tight case follows from the previous results, with $\mathsf{f} \equiv 0$.

In the center-tight case, the results of Chapter 3 tell us that

$$\widetilde{f}_n(x, y) = a_{n+1}(y) - a_n(x) + h_n(x, y) + c_n \quad \text{where} \quad \sum_n \mathrm{Var}(h_n) < \infty.$$

Moreover, we can obtain such a decomposition with $\|a_n\|_\infty \to 0$, see (3.6). Changing a_n if necessary, we may also assume that $\mathbb{E}(a_n) = 0$, in which case $\mathbb{E}(\widetilde{f}_n) = 0 = \mathbb{E}(h_n + c_n)$. Therefore the additive functional $\widetilde{\mathsf{h}} = \mathsf{h} + \mathsf{c}$ has zero mean and finite variance. Hence by Theorem 3.12,

$$\sum_{n=1}^{\infty} (h_n + c_n) \text{ converges almost surely.}$$

In summary $S_N(\widetilde{\mathsf{f}}) - a_N + a_1$ converges almost surely, and hence $S_N(\widetilde{\mathsf{f}}) - a_N$ converges almost surely. Since $\lim\limits_{N\to\infty} a_N = 0$, the proof is complete. □

Corollary 8.30 *Suppose* $\sup|\widetilde{f}_n| \to 0$ *and* $\mathbb{E}(\widetilde{f}_n) = 0$. *Then:*

- *Either* $S_N(\widetilde{\mathsf{f}})$ *converges a.s. to some random variable* $\mathcal{S}$, *in which case for each function* $\phi \in C_c(\mathbb{R})$,

$$\lim_{N\to\infty} \mathbb{E}(\phi(S_N)) = \mathbb{E}(\phi(\mathcal{S}));$$

- *or* $S_N(\widetilde{\mathsf{f}})$ *satisfies a non-lattice LLT, in which case, for every* $\phi \in C_c(\mathbb{R})$,

$$\lim_{N\to\infty} \sqrt{V_N}\mathbb{E}(\phi(S_N)) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \phi(s)\mathrm{d}s.$$

8.6 Equicontinuous Additive Functionals

In this section we examine the consequences of topological assumptions on f and X. We will say that (X,f) is **equicontinuous**, if the following three assumptions hold:

(T) $\mathfrak{S}_n$ are complete separable metric spaces.

(Σ) • $\pi_{n,n+1}(x,\mathrm{d}y) = p_n(x,y)\mu_{n+1}(\mathrm{d}y)$ where $\epsilon_0^{-1} \le p_n \le \epsilon_0$ for some $\epsilon_0 > 0$ independent of n, and
• for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every n, for every ball $B \subset \mathfrak{S}_n$ with radius ε, $\mu_n(B) > \delta$.

(U) For some $K < \infty$, ess sup $|\mathsf{f}| < K$, and for every $\varepsilon > 0$ $\exists\delta > 0$ such that

$$\sup\left\{|f_n(x',x'') - f_n(y',y'')| : n \in \mathbb{N}, \begin{array}{l} x',y' \in \mathfrak{S}_n \text{ s.t. } d(x',y') \le \delta \\ x'',y'' \in \mathfrak{S}_{n+1} \text{ s.t. } d(x'',y'') \le \delta \end{array}\right\} < \varepsilon.$$

In particular, by (Σ), X is uniformly elliptic (even one-step uniformly elliptic).

Theorem 8.31 *Suppose* (X,f) *is equicontinuous, and* $\mathfrak{S}_n$ *are connected. Then either* (X,f) *is center-tight, or* $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$, *and* (X,f) *satisfies the non-lattice LLT* (5.1).

Proof Assume (X,f) is not center-tight. Choose $c_1 > 0$ such that $|e^{i\theta} - 1|^2 = 4\sin^2\left(\frac{\theta}{2}\right) \ge c_1\theta^2$ for all $|\theta| \le 0.1$. Fix $\xi \neq 0$.

We consider the following two cases:

(I) $\exists N_0$ such that $|\xi\Gamma(P)| < 0.1$ for every hexagon $P \in \mathrm{Hex}\,(n)$ and $n \ge N_0$.
(II) $\exists n_k \uparrow \infty$ and hexagons $P_{n_k} \in \mathrm{Hex}\,(n_k)$ such that $|\xi\Gamma(P_{n_k})| \ge 0.1$.

In case (I), for all $n \ge N_0$, $d_n^2(\xi) = \mathbb{E}(|e^{i\xi\Gamma} - 1|^2) \ge c_1\mathbb{E}(\Gamma^2) \equiv c_1 u_n^2$. By non center-tightness, $\sum u_n^2 = \infty$, whence $\sum d_n^2(\xi) = \infty$.

In case (II), for every k there is a position n_k hexagon P_{n_k} with $|\xi\Gamma(P_{n_k})| \ge 0.1$. There is also a position n_k hexagon P'_{n_k} with balance zero, e.g. $\left(a \begin{array}{c} b\ c \\ b\ c \end{array} d\right)$. We would like to apply the intermediate value theorem to deduce the existence of $\overline{P}_{n_k} \in \mathrm{Hex}\,(n_k)$ such that $0.05 < \xi\Gamma(\overline{P}_{n_k}) < 0.1$. To do this we note that:

- Because of (Σ), the space of admissible position n_k hexagons is homeomorphic to $\mathfrak{S}_{n_k-2}\times\mathfrak{S}^2_{n_k-1}\times\mathfrak{S}^2_{n_k}\times\mathfrak{S}_{n_k}$.
- The product of connected topological spaces is connected.
- Real-valued continuous functions on connected topological spaces satisfy the intermediate value theorem.
- The balance of hexagon depends continuously on the hexagon.

So $\overline{P}_{n_k}$ exists. Necessarily, $|e^{i\xi\Gamma(\overline{P}_{n_k})} - 1| \ge c_1\xi^2\Gamma^2(\overline{P}_{n_k}) \ge c_1\xi^2\cdot 0.05^2 =: c_2$.

By the equicontinuity of f, $\exists\epsilon > 0$ such that $|e^{i\xi\Gamma(P)} - 1| > \frac{1}{2}c_2$ for every hexagon whose coordinates are in the ϵ-neighborhood of the coordinates of $\overline{P}_{n_k}$. By (Σ), this collection of hexagons has measure $\ge \delta$ for some $\delta > 0$ independent of k. So $d_{n_k}^2(\xi) \ge \frac{1}{2}c_2\delta$. Summing over all k, we find that $\sum d_{n_k}^2(\xi) = \infty$. Since $\xi \neq 0$ was arbitrary, $H(\mathsf{X},\mathsf{f}) = \{0\}$, and this implies that $G_{ess}(\mathsf{X},\mathsf{f}) = \mathbb{R}$. □

Theorem 8.32 *Suppose* $(\mathsf{X}, \mathfrak{f})$ *is equicontinuous, and* $V_N \geq \text{const.}N$ *for some positive constant. Then* $(\mathsf{X}, \mathfrak{f})$ *has full large deviations regime, and*

$$\mathfrak{c}_- = \limsup_{N\to\infty} \frac{\inf S_N - \mathbb{E}(S_N)}{V_N}\ ,\ \mathfrak{c}_+ = \liminf_{N\to\infty} \frac{\sup S_N - \mathbb{E}(S_N)}{V_N}, \tag{8.27}$$

where the infima and suprema are taken over all of $\mathfrak{S}_1 \times \cdots \times \mathfrak{S}_{N+1}$.

Example 7.36 shows that the equicontinuity assumption on $\mathfrak{f}$ cannot be removed.

Proof Changing $\mathfrak{f}$ by constants, it is possible to assume without loss of generality that $\mathbb{E}(S_N) = 0$ for all N.

The inequalities $0 < \mathfrak{c}_+ \leq \mathfrak{r}_+ \leq \liminf V_N^{-1} \sup S_N$ are always true, therefore to show the identity for $\mathfrak{c}_+$ it is sufficient to prove that $\mathfrak{c}_+ \geq \liminf V_N^{-1} \sup S_N$.

Fix $z, \varepsilon > 0$ such that $0 < z + \varepsilon < \liminf\limits_{N\to\infty} V_N^{-1} \sup S_N$. For all sufficiently large N there is a sequence $\widetilde{x}_1, \ldots, \widetilde{x}_{N+1}$ s.t.

$$\sum_{j=1}^{N} f_j(\widetilde{x}_j, \widetilde{x}_{j+1}) \geq (z+\varepsilon)V_N.$$

Let $B(x, r)$ denote the open ball with center x and radius r. By (U) and the fact that $V_N \geq \text{const.}N$, there is $r > 0$ such that if $X_j \in B(\widetilde{x}_j, r)$ for $2 \leq j \leq N+1$, then for all N large enough, $\sum\limits_{j=1}^{N} f_j(X_j, X_{j+1}) \geq (z + \varepsilon/2)V_N$.

By (Σ), all balls in $\mathfrak{S}_n$ with radius r have μ_n-measure bounded below by δ, for some positive constant $\delta = \delta(r)$. Therefore

$$\begin{aligned}&\mathbb{P}[X_2 \in B(\widetilde{x}_2, r), \ldots, X_{N+1} \in B(\widetilde{x}_{N+1}, r)]\\ &= \int\limits_{\mathfrak{S}_1} \int\limits_{B(\widetilde{x}_2, r)} \cdots \int\limits_{B(\widetilde{x}_{N+1}, r)} p_1(x_1, x_2) \cdots p_n(x_N, x_{N+1}) \pi_1(\mathrm{d}x_1)\mu_2(\mathrm{d}x_2) \cdots \mu_{N+1}(\mathrm{d}x_{N+1})\\ &\geq \epsilon_0^N \delta^N.\end{aligned}$$

Hence there is $0 < \eta < 1$ such that for all N large enough,

$$\mathbb{P}(S_N \geq (z + \varepsilon/3)V_N) \geq \eta^N.$$

Next, by the CLT and the assumptions that $\mathbb{E}(S_N) = 0$ and $z > 0$, if ε is small enough and N is large enough, then

$$\mathbb{P}[S_N \leq (z - \varepsilon)V_N] \geq \mathbb{P}[S_N \leq 0] = \frac{1}{2} + o(1) \geq \eta^N.$$

By Theorem 7.30, $z_N = zV_N$ is admissible. In particular (see §7.4.2), $z \leq \mathfrak{c}_+$. Passing to the supremum over z, we obtain that $\mathfrak{c}_+ \geq \liminf V_N^{-1} \sup S_N$, whence $\mathfrak{c}_+ = \liminf V_N^{-1} \sup S_N$. The identity for $\mathfrak{c}_-$ follows by symmetry. □

By (8.27), for each N, there is a finite sequence $\mathbf{x}^{\pm}(N) := (x^{\pm}_{1,N}, x^{\pm}_{2,N}, \ldots, x^{\pm}_{N+1,N})$ such that $\mathfrak{c}_{\pm} = \lim\limits_{N\to\infty} \dfrac{1}{V_N} \sum\limits_{n=1}^{N} \left(f_n(x^{\pm}_{n,N}, x^{\pm}_{n+1,N}) - \mathbb{E}(S_N) \right)$.

Our next result says that when $\mathfrak{S}_i$ are all compact, one can choose $\mathbf{x}^{\pm}(N)$ consistently, i.e. to have a common extension to an infinite sequence. Such infinite sequences appear naturally in statistical mechanics, where they are called "ground states."

An infinite sequence $\mathbf{x} = (x_1, x_2, \ldots) \in \prod_{i=1}^{\infty} \mathfrak{S}_i$ is called a **maximizer**, if $\forall N \geq 3$, for all $y_i \in \mathfrak{S}_i$,

$$\sum_{n=1}^{N-1} f_n(y_n, y_{n+1}) + f_N(y_N, x_{N+1}) \leq \sum_{n=1}^{N-1} f_n(x_n, x_{n+1}) + f_N(x_N, x_{N+1}). \tag{8.28}$$

An infinite sequence $\mathbf{x} = (x_1, x_2, \ldots) \in \prod_{i=1}^{\infty} \mathfrak{S}_i$ is called a and a **minimizer**, if $\forall N \geq 3$, for all $y_i \in \mathfrak{S}_i$,

$$\sum_{n=1}^{N-1} f_n(y_n, y_{n+1}) + f_N(y_N, x_{N+1}) \geq \sum_{n=1}^{N-1} f_n(x_n, x_{n+1}) + f_N(x_N, x_{N+1}). \tag{8.29}$$

Theorem 8.33 *Suppose* (X, f) *is equicontinuous,* $V_N \geq cN$ *for some positive constant c, and all the state spaces* $\mathfrak{S}_i$ *are compact. Then maximizers* $\mathbf{x}^+ = (x_1^+, x_2^+, \ldots)$ *and minimizers* $\mathbf{x}^- = (x_1^-, x_2^-, \ldots)$ *exist, and satisfy*

$$\mathfrak{c}_- = \limsup_{N\to\infty} \frac{1}{V_N}\Big(\sum_{i=1}^{N} f_n(x_n^-, x_{n+1}^-) - \mathbb{E}(S_N)\Big), \qquad \mathfrak{c}_+ = \liminf_{N\to\infty} \frac{1}{V_N}\Big(\sum_{i=1}^{N} f_n(x_n^+, x_{n+1}^+) - \mathbb{E}(S_N)\Big).$$

Proof Let $\mathcal{M}_N^+$ (respectively, $\mathcal{M}_N^-$) denote the space of sequences satisfying (8.28) (respectively, (8.29)) for fixed N.

Step 1. *$\mathcal{M}_N^\pm$ are non-empty compact sets, and $\mathcal{M}_N^\pm \supset \mathcal{M}_{N+1}^\pm$ for all N. Thus, the sets* $\mathcal{M}^\pm := \bigcap_{N=1}^{\infty} \mathcal{M}_N^\pm \neq \varnothing$.

Proof of the Step. By Tychonoff's theorem, $\mathfrak{S} := \prod_{i=1}^{\infty} \mathfrak{S}_i$ is compact, and by equicontinuity, for every $x \in \mathfrak{S}_{N+1}$, the map $\varphi_{N,x}(y_1, y_2, \ldots) := \sum_{i=1}^{N-1} f_i(y_i, y_{i+1}) + f_N(y_N, x)$ is continuous. Therefore $\varphi_{N,x}$ attains its maximum and its minimum on $\mathfrak{S}$. It follows that $\mathcal{M}_N^\pm$ are non-empty.

Suppose $\mathbf{x}^j \in \mathcal{M}_N^+$ converge to $\mathbf{x} \in \mathfrak{S}$. We claim that $\mathbf{x} \in \mathcal{M}_N^+$. Otherwise there is $\overline{\mathbf{x}} = (\overline{x}_1, \overline{x}_2, \ldots)$ such that $\overline{x}_{N+1} = x_{N+1}$, but $\sum_{n=1}^{N} f_n(\overline{x}_n, \overline{x}_{n+1}) > \sum_{n=1}^{N} f_n(x_n, x_{n+1})$.

Let $\overline{\mathbf{x}}^j := (\overline{x}_1, \ldots, \overline{x}_N, x_{N+1}^j, \overline{x}_{N+2}, \overline{x}_{N+3}, \ldots)$. Since $x_{N+1}^j \to x_{N+1} = \overline{x}_{N+1}$, $\overline{\mathbf{x}}^j \to \overline{\mathbf{x}}$. By equicontinuity, for all j sufficiently large, $\sum_{n=1}^{N} f_n(\overline{x}_n^j, \overline{x}_{n+1}^j) > \sum_{n=1}^{N} f_n(x_n^j, x_{n+1}^j)$. But this contradicts the assumption that $\mathbf{x}^j \in \mathcal{M}_N^+$.

We see that $\mathcal{M}_N^+$ is closed, whence by the compactness of $\mathfrak{S}$, compact. Similarly one shows that $\mathcal{M}_N^-$ is compact.

Next we prove the monotonicity of $\mathcal{M}_N^\pm$. Suppose $\mathbf{x} \in \mathcal{M}_{N+1}^+$. Choose some arbitrary $\mathbf{y} \in \mathfrak{S}$. Looking at the sequence $\mathbf{y}' := (y_1, \ldots, y_N, x_{N+1}, x_{N+2}, y_{N+3}, y_{N+4}, \ldots)$ and using the defining property of $\mathcal{M}_{N+1}$, we see that

$$\sum_{n=1}^{N-1} f_n(y_n, y_{n+1}) + f_N(y_N, x_{N+1}) + f_{N+1}(x_{N+1}, x_{N+2}) \leq \sum_{n=1}^{N+1} f_n(x_n, x_{n+1}).$$

It follows that $\sum_{n=1}^{N-1} f_n(y_n, y_{n+1}) + f_N(y_N, x_{N+1}) \leq \sum_{n=1}^{N} f_n(x_n, x_{n+1})$. So $\mathbf{x} \in \mathcal{M}_N^+$.

Similarly, one shows that $\mathcal{M}_{N+1}^- \subset \mathcal{M}_N^-$.

Step 2. $\forall \mathbf{x}^\pm \in \mathcal{M}_N^\pm$, $\sum_{n=1}^{N} f_n(x_n^-, x_{n+1}^-) \leq \inf S_N + 2K$, $\quad \sum_{n=1}^{N} f_n(x_n^+, x_{n+1}^+) \geq \sup S_N - 2K$.

Proof of the Step. Let $(z_1, \ldots, z_N)$ be a point where $\sum_{n=1}^{N-1} f_n(z_n, z_{n+1}) = \max S_{N-1}$. Take $\mathbf{y} := (z_1, \ldots, z_N, x_{N+1}^+, x_{N+2}^+, \ldots)$, and recall that $\operatorname{ess\,sup} |\mathsf{f}| < K$. Then

$$\sum_{n=1}^{N} f_n(x_n^+, x_{n+1}^+) \geq \sum_{n=1}^{N} f_n(y_n, y_{n+1}) \geq \sum_{n=1}^{N-1} f_n(z_n, z_{n+1}) - K = \max S_{N-1} - K \geq \max S_N - 2K.$$

The statement for $\mathbf{x}^-$ has a similar proof, which we omit.

We can now prove the theorem. By Theorem 8.32, $\mathfrak{c}_+ = \liminf_{N\to\infty} \frac{\sup S_N - \mathbb{E}(S_N)}{V_N}$ (this is the only place where we are using the assumption that $V_N \geq \text{const.}N$).

By Step 2, $\sup S_N \leq \sum_{n=1}^{N} f_n(x_n^+, x_{n+1}^+) + O(1)$ for each $\mathbf{x}^+ \in \mathcal{M}^+$. So $\mathfrak{c}_+ = \liminf_{N\to\infty} \frac{1}{V_N}\left(\sum_{n=1}^{N} f_n(x_n^+, x_{n+1}^+) - \mathbb{E}(S_N)\right)$.

The proof of the identity for $\mathfrak{c}_-$ is similar, and we omit it. □

8.7 Notes and References

Sums of Independent Random Variables. The non-lattice LLT in Theorem 8.3 is due to Dolgopyat [56]. His proof also applies to unbounded vector-valued random variables, assuming only that $\sup \mathbb{E}(\|X_n\|^3) < \infty$.

The lattice LLT in Theorem 8.5 is due to Prokhorov. An extension to unbounded integer valued independent random variables is given by Rozanov [169].

Other conditions for the LLT for sums of independent random variables include the Mineka-Silverman condition [144], Statulevičius's condition [190], and conditions motivated by additive number theory such as those appearing in [146] and [147].

Mukhin [148] gave a unified discussion of some of these conditions, using the quantities $\mathfrak{D}(X, \xi)$.

Homogeneous Chains. The literature on homogeneous Markov chains is vast. Sufficient conditions for the CLT, LLT, and other limit theorems in non-Gaussian domains of attraction can be found in [46, 81, 82, 98, 114, 119, 139, 149, 98, 168, 88, 152, 10, 122, 123].

The LLT for *local* deviations holds under weaker assumptions than those in Theorems 8.9 and 8.13. The assumption that f has finite variance can be replaced by the assumption that the distribution of f is in the domain of attraction of the Gaussian distribution [4]; One can allow f to depend on infinitely many X_n assuming that the dependence of $f(x_1, x_2, \ldots)$ on $(x_n, x_{n+1}, \ldots)$ decays exponentially in n [88]; and the ellipticity assumption can be replaced by the assumption that the generator has a spectral gap [149, 98]. In particular, the LLT holds under the Doeblin condition saying that $\exists \varepsilon_0 > 0$ and a measure ζ on $\mathfrak{S}$ such that $\pi(x, dy) = \varepsilon_0 \zeta(dy) + (1-\varepsilon_0)\widetilde{\pi}(x, dy)$ where $\widetilde{\pi}$ is an arbitrary transition probability (cf. equation (2.10) in the proof of Lemma 2.12). There are also versions of this theorem for f in the domain of attraction of a stable law, see [5].

There are also generalizations of the LLT for *large* deviations to the case when $f(x_1, x_2, \ldots)$ has exponentially weak dependence on x_k with large k [123, 122, 10]. However, the unbounded case is still not understood. In fact, the large deviation probabilities could behave polynomially for unbounded functions, see [198, 141].

The characterization of coboundaries in terms of vanishing of the asymptotic variance σ^2 is due to Leonov [129]. A large number of papers discuss the regularity of the gradients in case an additive functional is a gradient, see [21, 37, 106, 134, 135, 153, 152] and the references wherein. Our approach is closest to [55, 106, 152].

We note that the condition $u(f) = 0$ (where $u^2(f)$ is the variance of the balance of a random hexagon), which is sufficient for f being a coboundary, is simpler than the equivalent condition $\sigma^2 = 0$. For example, for finite state Markov chains, to compute σ^2 one needs to compute infinitely many correlations $\mathbb{E}(f_0 f_n)$ while checking that $u = 0$ involves checking the balance of finitely many hexagons.

Nagaev's Theorem. Nagaev's Theorem and the idea to prove limit theorems for Markov chains using perturbation theory of linear operators first appeared in [149]. Nagaev only treated the lattice case. He did not assume the one-step ellipticity condition, he only assumed the weaker condition that for some k, the k-step transition kernel of X has contraction coefficient strictly smaller than one (see §2.2.2). This is sufficient to guarantee the spectral gap of $\mathcal{L}_0$, but it does not allow to characterize the cases when $\sigma^2 > 0$, $G_{ess}(\mathsf{X}, \mathsf{f}) = \mathbb{R}$, and $G_{ess}(\mathsf{X}, \mathsf{f}) = t\mathbb{Z}$ in a simple way as in Propositions 8.14 and 8.15. As a result, the LLT under Nagaev's condition is more complicated than in our case.

Nagaev's proof can be generalized even further, to the case when $\mathcal{L}_t$ all have spectral gap on some suitable Banach space. See [87, 88, 98].

A proof of the perturbation theorem (Theorem 8.17) can be found in [98], and we included it, for completeness in Appendix C. The expansion of the leading eigenvalue λ_t follows calculations in Guivarc'h & Hardy [88], see also [171, 152]. The identities for the derivatives of λ are often stated in the following form $(\ln\lambda)'_0 = i\mathbb{E}(f)$, $(\ln\lambda)''_0 = -\sigma^2/2$. We leave it to readers to check that these formulae are equivalent to Proposition 8.21(4). The proof of Proposition 8.14 uses ideas from [84].

The LLT for *stationary* homogeneous uniformly elliptic Markov chains in §8.3-§8.4 remain true if we remove the stationarity assumptions, with the one caveat: The characterization of the cases $\sigma^2 = 0$ and $G_{ess} \neq \mathbb{R}$ in terms of an a.s. functional equation for $f(X_1, X_2)$ must always be done using the stationary law of $\{X_n\}$.

To establish these LLT in the non-stationary case, one could either appeal to the more general results on asymptotically homogeneous chains in §8.5, or make minor modifications in the arguments of §8.3-§8.4. In the case of Theorem 8.9, it suffices to note that Theorem 5.1 holds for all initial distributions and that $G_{ess}(\mathsf{X}, \mathsf{f})$ does not depend on the initial distribution, because of Lemma 2.17 and Theorem 4.4. In the case of Theorem 8.13, one has to use the fact that Nagaev's perturbation theory allows to control $\mathbb{E}\left(e^{itS_N/\sqrt{V_N}}\right)$ for an arbitrary initial distribution.

It is interesting to note that the *higher order terms* in the asymptotic expansion of the CLT and LLT probabilities do depend on the initial distribution, see [99].

Asymptotically Homogeneous Chains. Asymptotically homogeneous Markov chains appear naturally in some stochastic optimization algorithms such as the Metropolis algorithm. For large deviations and other limit theorems for such examples, see [48, 47] and references therein.

Asymptotically homogenous systems are standard examples of inhomogeneous systems with linearly growing variance, cf. [31, 151].

We note that using the results of §2.5 it is possible to strengthen Lemma 8.27 to conclude that $\mathrm{Var}[S_N(\widetilde{\mathsf{X}}, \widetilde{\mathsf{f}})] = \mathrm{Var}[S_N(\mathsf{X}, \mathsf{f})] + o(N)$ as $N \to \infty$, but the present statement is sufficient for our purposes.

Equicontinuous Additive Functionals Minimizers play an important role in statistical mechanics, where they are called *ground states*. See e.g. [184, 164]. In the case the phase spaces $\mathfrak{S}_n$ are non-compact and/or the observable $f(x, y)$ is unbounded, the minimizers have an interesting geometry, see e.g. [29]. For finite state Markov chains, we have the following remarkable result of J. Brémont [18]: for each d there is a constant $p(d)$ such that for any homogeneous Markov chain with d states for any additive functional we have

$$\mathrm{r}_+ = \max_{q \le p} \frac{1}{q} \max_{x_1, \dots x_q} \left[f(x_1, x_2) + \cdots + f(x_{q-1}, x_q) + f(x_q, x_1) \right].$$

This result is false for more general homogenous chains, consider for example the case $\mathfrak{S} = \mathbb{N}$ and $f(x, y) = 1$ if $y = x + 1$ and $f(x, y) = 0$ otherwise.

Corollary 8.30 was proven in [56] for inhomogeneous sums of independent random variables. In the independent case, the assumption that $\lim\limits_{n\to\infty} \|g_n\|_\infty = 0$ can be removed, since the gradient obstruction does not appear.

Quantitative versions of Corollary 8.30 and Theorem 8.31 have been obtained in [59]. There it is shown that

$$\mathbb{P}[S_N - z\sqrt{V_N} \in (a_N, b_N)] = [1 + o(1)] \frac{e^{-z^2/2}}{\sqrt{2\pi V_N}} (b_N - a_N), \text{ as } N \to \infty$$

provided that $C \ge b_N - a_N \ge V_N^{-k/2}$ where k is integer such that

- either $\|f_n\| = O(n^{-\beta})$ and $k < \frac{1}{1-2\beta} - 1$;
- or $\mathfrak{S}_n = M$-a compact connected manifold, f_n are uniformly Hölder of order α and $k < \frac{1+\alpha}{1-\alpha} - 1$.

These results are consequences of so called *Edgeworth expansions*, which are precise asymptotic expansions for $\mathbb{P}(\frac{S_N - \mathbb{E}(S_N)}{\sqrt{V_N}} \le z)$. These results improve on Corollary 8.30 and Theorem 8.31, since the length of the target interval (a_N, b_N) is allowed to go to zero. The exponents k given above are optimal.

Chapter 9
Local Limit Theorems for Markov Chains in Random Environments

Abstract We prove quenched local limit theorems for Markov chains in random environments, with stationary ergodic noise processes.

9.1 Markov Chains in Random Environments

A Markov chain in a random environment (MCRE) is an inhomogeneous Markov chain whose transition probabilities depend on random external parameters Y_n which vary in time: $\pi_{n,n+1}(x, \mathrm{d}y) = \pi(Y_n, x, \mathrm{d}y)$.[1]

The noise $\{Y_n\}$ is a stochastic process, which we will always take to be stationary and ergodic. In this case it is possible and convenient to represent $\{Y_n\}$ as a random orbit of an ergodic measure preserving map, called the noise process. We proceed to give the formal definitions and some examples.

9.1.1 Formal Definitions

Noise Processes: These are quadruples $(\Omega, \mathscr{F}, m, T)$, where T is an ergodic measure preserving invertible Borel transformation on a standard measure space $(\Omega, \mathscr{F}, m)$.

- "**Ergodic**" means that for every $E \in \mathscr{F}$ s.t. $T^{-1}E = E$, $m(E) = 0$ or $m(\Omega \setminus E) = 0$.
- "**Measure preserving**" means that for every $E \in \mathscr{F}$, $m(T^{-1}E) = m(E)$.
- "**Invertible**" means that there exists $\Omega_1 \subset \Omega$ of full measure such that $T : \Omega_1 \to \Omega_1$ is injective, surjective, measurable, and with measurable inverse.[2]

The **noise at time *n*** is $Y_n := T^n\omega := (T \circ \cdots \circ T)(\omega)$ (n times), $\omega \in (\Omega, \mathscr{F}, m)$.

If $m(\Omega) < \infty$ then we will speak of a **finite noise process**, and we will always normalize m so that $m(\Omega) = 1$. Every stationary ergodic stochastic process taking values in a polish space can be modeled by a finite noise process, see Example 9.3. If $m(\Omega) = \infty$, then we will speak of an **infinite noise process**. In this case, we will always assume that $(\Omega, \mathscr{F}, m)$ is σ-finite and non-atomic. Such processes arise in the study of noise driven by null recurrent Markov chains, see Example 9.5.

Markov Chains in Random Environment (MCRE): These are quadruples

$$\mathsf{X}^\Omega := \big(\underbrace{(\Omega, \mathscr{F}, m, T)}_{\text{noise process}},\ \underbrace{(\mathfrak{S}, \mathscr{B})}_{\text{state space}},\ \underbrace{\{\pi(\omega, x, \mathrm{d}y)\}_{(\omega,x)\in\Omega\times\mathfrak{S}}}_{\text{transition kernel generator}},\ \underbrace{\{\mu_\omega\}_{\omega\in\Omega}}_{\text{initial distribution generator}}\big) \tag{9.1}$$

made of the following objects:

- **The Noise Process** $(\Omega, \mathscr{F}, m, T)$, see above.
- **The State Space** $(\mathfrak{S}, \mathscr{B})$ is a separable complete metric space $\mathfrak{S}$, with its Borel σ-algebra $\mathscr{B}$.

---

[1] MCRE should not be confused with "random walks in random environment." In the RWRE model, the transition kernel at time n depends on a random "environment" Y_0 independent of n, and on the *position* of the random walk at time n, i.e. $\pi_{n,n+1}(x, \mathrm{d}y) = \pi(Y_0, S_n, x, \mathrm{d}y)$.

[2] The invertibility assumption can be removed by replacing a non-invertible map by its natural extension, see [33, Ch. 10].

- **The Transition Kernel Generator** $\{\pi(\omega, x, \mathrm{d}y)\}_{(\omega,x)\in\Omega\times\mathfrak{S}}$ is a measurable family of Borel probability measures on $(\mathfrak{S}, \mathscr{B})$. Measurability means that $(\omega, x) \mapsto \int \varphi(y)\pi(x, \omega, \mathrm{d}y)$ is measurable for every bounded Borel $\varphi : \mathfrak{S} \to \mathbb{R}$.
- **The Initial Distribution Generator** $\{\mu_\omega\}_{\omega\in\Omega}$ is a measurable family of Borel probability measures μ_ω on $\mathfrak{S}$. Measurability means that for all bounded Borel $\varphi : \mathfrak{S} \to \mathbb{R}$, $\omega \mapsto \int \varphi(x)\mu_\omega(\mathrm{d}x)$ is measurable.

Fix $\omega \in \Omega$. The **quenched MCRE with noise parameter** $\boldsymbol{\omega}$ is the inhomogeneous Markov chain $\mathsf{X}^\omega = \{X_n^\omega\}_{n\geq 1}$ with

• state space $\mathfrak{S}$; • initial distribution $\mu_{T\omega}$; • transition kernels $\pi^\omega_{n,n+1}(x, \mathrm{d}y) = \pi(T^n\omega, x, \mathrm{d}y)$.

Remark. Note that the initial distribution of X^ω is $\mu_{T\omega}$, not μ_ω. Similarly, $\pi_{1,2}(x, \mathrm{d}y) = \pi(T\omega, x, \mathrm{d}y)$. These choices are consistent with the convention that whatever happens at time n depends on $T^n\omega$.

We denote by $\mathbb{P}^\omega$, $\mathbb{E}^\omega$ the probability and expectation associated with X^ω. These are sometimes called the **quenched probability and expectation** (for the noise value ω). By contrast, the **annealed probability and expectation** are given by $m(\mathrm{d}\omega)\mathbb{P}^\omega(\mathrm{d}\mathsf{X}^\omega)$ and $\int_\Omega \mathbb{E}^\omega m(\mathrm{d}\omega)$.

Additive Functional Generator: $\mathsf{f}^\Omega := f$, where $f : \Omega \times \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ is a measurable function. This generates the additive functional f^ω on X^ω given by

$$f_n^\omega(x, y) = f(T^n\omega, x, y). \tag{9.2}$$

We let $S_N^\omega := \sum_{n=1}^{N} f_n^\omega(X_n^\omega, X_{n+1}^\omega) \equiv \sum_{n=1}^{N} f(T^n\omega, X_n^\omega, X_{n+1}^\omega)$.

9.1.2 Examples

Let $(\mathfrak{S}, \mathscr{B})$ be a compact metric space, fix a countable set S, and let $\{\pi_i(x, \mathrm{d}y)\}_{i\in S}$ be some family of transition kernels on $\mathfrak{S}$.

Example 9.1 (Bernoulli Noise) Consider the noise process $(\Omega, \mathscr{F}, m, T)$ where

- $\Omega = S^{\mathbb{Z}} = \{(\cdots, \omega_{-1}, \omega_0, \omega_1, \cdots) : \omega_i \in S\}$.
- $\mathscr{F}$ is generated by the **cylinders** ${}_k[a_k, \ldots, a_n] := \{\omega \in \Omega : \omega_i = a_i, k \leq i \leq n\}$.
- $\{p_i\}_{i\in S}$ are non-negative numbers such that $\sum p_i = 1$, and m is the unique probability measure such that $m({}_k[a_k, \ldots, a_n]) = p_{a_k} \cdots p_{a_n}$ for all cylinders.
- $T : \Omega \to \Omega$ is the **left shift map**, $T[(\omega_i)_{i\in\mathbb{Z}}] = (\omega_{i+1})_{i\in\mathbb{Z}}$.

$(\Omega, \mathscr{F}, \mu, T)$ is ergodic and probability preserving, see [33].

Define $\pi(\omega, x, \mathrm{d}y) := \pi_{\omega_0}(x, \mathrm{d}y)$. Notice that $\pi(T^n\omega, x, \mathrm{d}y) = \pi_{\omega_n}(x, \mathrm{d}y)$, and ω_n are iid random variables taking the values $i \in S$ with probabilities p_i. So X^Ω represents a random Markov chain, whose transition probabilities vary randomly and independently in time.

Example 9.2 (Positive Recurrent Markovian Noise) Suppose $(Y_n)_{n\in\mathbb{Z}}$ is a stationary ergodic Markov chain with state space S and a stationary probability vector $(p_s)_{s\in S}$. In particular, $(Y_n)_{n\in\mathbb{Z}}$ is positive recurrent. Let:

• $\Omega := S^{\mathbb{Z}}$; • $\mathscr{F}$ is the σ-algebra generated by the cylinders (see above);

• m is the unique (probability) measure such that for all cylinders, $m({}_k[a_k, \ldots, a_n]) = \mathbb{P}[Y_k = a_k, \ldots, Y_n = a_n]$;

• T is the left shift map (see above).

Define as before, $\pi(\omega, x, \mathrm{d}y) := \pi_{\omega_0}(x, \mathrm{d}y)$. The resulting MCRE represents a Markov chain whose transition probabilities at time $n = 1, 2, 3, \ldots$ are $\pi_{Y_n}(x, \mathrm{d}y)$.

Example 9.3 (General Stationary Ergodic Noise Processes) The previous construction works verbatim with any stationary ergodic stochastic process $\{Y_n\}$ taking values in S. The assumption that S is countable can be replaced by the condition that S is a complete separable metric space, see e.g. [62].

Example 9.4 (Quasi-Periodic Noise) Let $(\Omega, \mathscr{F}, m, T)$ be the circle rotation: $\Omega = \mathbb{T}^1 := \mathbb{R}/\mathbb{Z}$; $\mathscr{F}$ is the Borel σ-algebra; m is the normalized Lebesgue measure; and $T : \Omega \to \Omega$ is the rotation by an angle α, $T(\omega) = \omega + \alpha \mod \mathbb{Z}$. T is probability preserving, and it is well-known that T is ergodic iff α is irrational, see [33].

Take a measurable (possibly continuous) 1-parameter family $\{\pi_\omega(x, \mathrm{d}y)\}_{\omega \in \mathbb{T}}$ of transition kernels on $\mathfrak{S}$, and form the generator $\pi(\omega, x, \mathrm{d}y) = \pi_\omega(x, \mathrm{d}y)$. Then X^ω are inhomogeneous Markov chains whose transition probabilities vary quasi-periodically: $\pi_{n,n+1}(x, \mathrm{d}y) = \pi_{\omega+n\alpha \bmod \mathbb{Z}}(x, \mathrm{d}y)$. More generally, let $\Omega := \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with the Haar measure m, and consider $T(\omega) := \omega + \alpha \bmod \mathbb{Z}^d$. This is a probability preserving map, and if $\{\alpha_1, \ldots, \alpha_d, 1\}$ are linearly independent over $\mathbb{Q}$, then T is ergodic [33]. Next, take a measurable (possibly continuous) d-parameter family of transition kernels $\{\pi_\omega(x, \mathrm{d}y)\}_{\omega \in \mathbb{T}^d}$ on $\mathfrak{S}$, and form the generator $\pi(\omega, x, \mathrm{d}y) = \pi_\omega(x, \mathrm{d}y)$. Then X^ω has transition probabilities which vary quasi-periodically:

$$\pi_{n,n+1}(x, \mathrm{d}y) = \pi_{(\omega_1+n\alpha_1, \ldots, \omega_d+n\alpha_d) \bmod \mathbb{Z}^d}(x, \mathrm{d}y).$$

Example 9.5 (Null Recurrent Markovian Noise) This is an example with an infinite noise process. Suppose $(Y_n)_{n\in\mathbb{Z}}$ is an ergodic null recurrent Markov chain with countable state space S, and stationary positive vector $(p_i)_{i\in S}$. Here $p_i > 0$ and (by null recurrence) $\sum p_i = \infty$. For example, $(Y_n)_{n\in\mathbb{Z}}$ could be the simple random walk on $\mathbb{Z}^d$ for $d = 1, 2$, with the stationary measure, $p_i = 1$ for all $i \in \mathbb{Z}^d$, (the counting measure). Let

- $\Omega = S^{\mathbb{Z}}$; • $\mathscr{F}$ is the σ-algebra generated by the cylinders;
- m is the unique (infinite) Borel measure s.t. $\forall$ cylinder $m({}_k[a_k, \ldots, a_n]) = p_{a_k}\mathbb{P}[Y_i = a_i \ (k \le i \le n) | Y_k = a_k]$;
- $T : \Omega \to \Omega$ is the left shift map $T[(\omega_i)_{i\in\mathbb{Z}}] = \omega_{i+1}$.

$(\Omega, \mathscr{F}, m, T)$ is an *infinite* ergodic measure preserving invertible map, see [1]. As in Example 9.2, one can construct MCRE with transition probabilities $\pi_{Y_n}(x, \mathrm{d}y)$ which vary randomly in time according to $(Y_n)_{n\in\mathbb{Z}}$. For each particular realization of $\omega = (Y_i)_{i\in\mathbb{Z}}$, X^ω is an ordinary inhomogeneous Markov chain (on a probability space). We shall see that some additive functionals on X^Ω may exhibit slower variance growth than in the case of finite noise processes (Example 9.16).

Example 9.6 (Transient Markovian Noise: A Non-Example) The previous construction fails for transient Markov chains such as the random walk on $\mathbb{Z}^d$ for $d \ge 3$, because in the transient case, $(\Omega, \mathscr{F}, m, T)$ is not ergodic, see [1].

We could try to work with the ergodic components of m, but this does not yield a new mathematical object, because of the following general fact [1]: Almost every ergodic component of an *invertible* totally dissipative infinite measure preserving map is concentrated on a single orbit $\{T^n\omega\}_{n\in\mathbb{Z}}$. Since MCRE with such noise processes have just one possible realization of noise up to time shift, their theory is the same as the theory of general inhomogeneous Markov chains.

9.1.3 Conditions and Assumptions

We present and discuss the conditions on X^Ω and f^Ω that will appear in some of the results in this chapter. In what follows, $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$ are as in (9.1) and (9.2). We begin with conditions on X^Ω:

(S) **Stationarity:** $\mu_{T\omega}(\mathrm{d}y) = \displaystyle\int_{\mathfrak{S}} \mu_\omega(\mathrm{d}x)\pi(\omega, x, \mathrm{d}y)$, i.e. for every $\varphi : \mathfrak{S} \to \mathbb{R}$ bounded and Borel, $\int \varphi(y)\mu_{T\omega}(\mathrm{d}y) = \int_{\mathfrak{S}} \left(\int_{\mathfrak{S}} \varphi(y)\pi(\omega, x, \mathrm{d}y)\right) \mu_\omega(\mathrm{d}x)$.

The following consequences of (S) can be easily proved by induction:

(S1) $\mathbb{P}^\omega(X_n^\omega \in E) = \mu_{T^n\omega}(E)$ for all $E \in \mathscr{B}$ and $n \ge 1$; (S2) $\forall k$ $\{X_{i+k}^\omega\}_{i\ge1}$ is equal in distribution to $\left\{X_i^{T^k\omega}\right\}_{i\ge1}$.

(E) **Uniform Ellipticity:** There is a constant $0 < \epsilon_0 < 1$ such that

(a) $\pi(\omega, x, \mathrm{d}y) = p(\omega, x, y)\mu_{T\omega}(\mathrm{d}y)$, with $p : \Omega \times \mathfrak{S} \times \mathfrak{S} \to [0, \infty)$ Borel; (b) $0 \le p \le 1/\epsilon_0$;

(c) $\int_{\mathfrak{S}} p(\omega, x, y)p(T\omega, y, z)\mu_{T\omega}(\mathrm{d}y) > \epsilon_0 \forall \omega, x, z$.

This implies that X^ω are all uniformly elliptic, and the ellipticity constant is uniformly bounded away from zero. In particular, X^ω *satisfies the exponential mixing estimates in Prop. 2.13 uniformly in* ω.

(G) **Global Support:** $\mu_\omega(\mathrm{d}y)$ is globally supported on $\mathfrak{S}$ for all ω.[3] The essence of this assumption is that *the support of* μ_ω *is independent of* ω. If $\mathrm{supp}(\mu_\omega) = \mathfrak{S}'$ for all ω, then we can get (G) by replacing $\mathfrak{S}$ by $\mathfrak{S}'$.

(C) **Continuity:** Ω and $\mathfrak{S}$ are separable complete metric spaces equipped with the Borel structure, and:
(C1) $T : \Omega \to \Omega$ is a homeomorphism; (C2) $(\omega, x, y) \mapsto p(\omega, x, y)$ is continuous;
(C3) $\omega \mapsto \int_{\mathfrak{S}} \varphi \mathrm{d}\mu_\omega$ is continuous $\forall$ bounded continuous $\varphi : \mathfrak{S} \to \mathbb{R}$; (C4) $(\omega, x, y) \mapsto f(\omega, x, y)$ are continuous.

(D) **Discreteness:** $\mathfrak{S}$ is finite or countable. This is an alternative to (C).

Example 9.7 It is fairly easy to get examples which satisfy (E),(G),(C) or (E),(G),(D). Take a noise process compatible with (C) (e.g. Examples 9.1–9.5). Next take a deterministic uniformly elliptic continuous globally supported transition kernel $p(x, y)\mu(\mathrm{d}y)$ on a compact metric (or finite) space. Let φ and ψ be bounded continuous functions on $\Omega \times \mathfrak{S}$ and $\Omega \times \mathfrak{S}^2$, respectively, and consider the generators

$$\mu_\omega(\mathrm{d}y) := \frac{\mathrm{e}^{\psi(\omega,y)}\mu(\mathrm{d}y)}{\int \mathrm{e}^{\psi(\omega,\eta)}\mu(\mathrm{d}\eta)}, \ p(\omega, x, y) := \frac{\mathrm{e}^{\varphi(\omega,x,y)}p(x, y)}{\int \mathrm{e}^{\varphi(\omega,x,\eta)}p(x, \eta)\mu_\omega(\mathrm{d}\eta)}.$$

Next we discuss how to obtain, in addition, stationarity (S). The following lemma will be proved in §9.3.1:

Lemma 9.8 *Let* X^Ω *be a uniformly elliptic MCRE with a compact metrizable state space* $(\mathfrak{S}, \mathscr{B})$ *and an initial distribution* $\{\mu_\omega\}_{\omega\in\Omega}$. *Suppose* $x \mapsto \pi(\omega, x, \mathrm{d}y)$ *is continuous in the weak-star topology for each* ω. *Then:*

(1) There exists an initial distribution generator $\{\mu'_\omega\}_{\omega\in\Omega}$ *which satisfies (S).*
(2) $\mu'_\omega \ll \mu_\omega$ *and there exists* $C > 0$ *such that* $|\log \frac{\mathrm{d}\mu'_\omega}{\mathrm{d}\mu_\omega}| < C$ *a.e. in* $\mathfrak{S}$ *for all* ω.
(3) Suppose in addition that Ω *is a metric space,* $T : \Omega \to \Omega$ *is continuous, and* $(\omega, x) \mapsto \pi(\omega, x, \mathrm{d}y)$ *is continuous. Then* $\omega \mapsto \mu'_\omega$ *is continuous.*

Corollary 9.9 *Suppose* X^Ω *satisfies conditions (E),(G) and (C) with a compact metrizable state space* $\mathfrak{S}$. *Then there is a MCRE* $\overline{\mathsf{X}}^\Omega$ *satisfying (E),(G),(C) and (S), so that for every* $\omega \in \Omega$ *and* $x \in \mathfrak{S}$, $\overline{\mathsf{X}}^\omega$ *conditioned on* $\overline{X}_1^\omega = x$ *is equal in distribution to* X^ω *conditioned on* $X_1^\omega = x$.

Proof Let $\overline{\mathsf{X}}^\Omega$ be the MCRE with the noise process, state space, and transition kernel generator of X^Ω, but with initial distribution generator μ'_ω from Lemma 9.8. □

Corollary 9.9 and Example 9.7 give many examples satisfying (E),(S),(G),(C). In the special case when $\mathfrak{S}$ is finite and discrete, we also obtain (D).

So far we only considered conditions on X^Ω. Next we discuss three conditions on f^Ω. We need the **annealed measure** $\mathbb{P}(\mathrm{d}\omega, \mathrm{d}x, \mathrm{d}y) := m(\mathrm{d}\omega)\mu_\omega(\mathrm{d}x)\pi(\omega, x, \mathrm{d}y)$. $\mathbb{P}$ represents the joint distribution of $(T\omega, X_1^\omega, X_2^\omega)$, because by the T-invariance of m,

$$\iiint \psi(\omega, x, y)\mathrm{d}\mathbb{P} = \int_{\mathfrak{S}} \int_{\mathfrak{S}} \int_{\Omega} \psi(T\omega, x, y)m(\mathrm{d}\omega)\mu_{T\omega}(\mathrm{d}x)\pi(T\omega, x, \mathrm{d}y). \tag{9.3}$$

(B) **Uniform Boundedness:** $|f| \le K$ where $K < \infty$ is a constant. (B) implies that f^ω is a uniformly bounded additive functional on X^ω, *and that the bound does not depend on* ω.

(RC1) f^Ω is called **relatively cohomologous to a constant** if there are bounded measurable functions $a : \Omega \times \mathfrak{S} \to \mathbb{R}$ and $c : \Omega \to \mathbb{R}$ such that $f(\omega, x, y) = a(\omega, x) - a(T\omega, y) + c(\omega)$ $\mathbb{P}$-a.e.

(RC2) Fix $t \neq 0$, then f^Ω is **relatively cohomologous to a coset of *tZ*** if there are measurable functions $a : \Omega \times \mathfrak{S} \to S^1$ and $\lambda : \Omega \to S^1$ such that

$$\mathrm{e}^{(2\pi \mathrm{i}/t) f(\omega,x,y)} = \lambda(\omega)\frac{a(\omega, x)}{a(T\omega, y)} \ \mathbb{P}\text{-a.e.}$$

We will use (RC1) and (RC2) to characterize a.e. center-tightness and a.e. reducibility for $(\mathsf{X}^\omega, \mathsf{f}^\omega)$, see Theorems 9.10 and 9.17, and Proposition 9.24.

[3] But $\pi(\omega, x, \mathrm{d}y)$ need not have global support, because $p(\omega, x, y)$ is allowed to vanish.

9.2 Main Results

Throughout this section, we assume that $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$ are as in (9.1) and (9.2), and we let V_N^ω denote the variance of $S_N^\omega := f(T\omega, X_1^\omega, X_2^\omega) + \cdots + f(T^N\omega, X_N^\omega, X_{N+1}^\omega)$, with respect to the (quenched) distribution of X^ω.

Theorem 9.10 *Suppose X^Ω has a finite noise process, and assume (B),(E) and (S).*

(1) If f^Ω is relatively cohomologous to a constant, then $\exists C = C(\epsilon_0, K)$ such that for a.e. ω, $V_N^\omega \le C$ for all N.
(2) If f^Ω is not relatively cohomologous to a constant, then there is a constant $\sigma^2 > 0$ such that for a.e. ω, $V_N^\omega \sim N\sigma^2$ as $N \to \infty$.

Thus, under the assumptions of Theorem 9.10, the limit $\sigma^2 := \lim\limits_{N\to\infty} V_N^\omega/N$ exists for a.e. ω, and is a.s. constant. We call σ^2 the **asymptotic variance** of $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$.

There is also an **asymptotic mean** μ such that $\mu := \lim \mathbb{E}^\omega(S_N^\omega)/N$ for a.e. ω: By (S), $\mathbb{E}^\omega[f(T^{n+1}\omega, X_{n+1}^\omega, X_{n+2}^\omega)] = \varphi(T^n\omega)$ where $\varphi(\omega) = \mathbb{E}^\omega(f(T\omega, X_1^\omega, X_2^\omega))$, and by the pointwise ergodic theorem, $\mathbb{E}(S_N^\omega) = \sum_{n=0}^{N-1}\varphi(T^n\omega) \sim N\int\varphi \mathrm{d}m$ m-a.e.

Theorem 9.11 *Assume that X^Ω has a finite noise process, and suppose $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$ satisfies (C) or (D), and each of (S),(E),(G),(B).*

(1) **Non-Lattice LLT:** *Suppose f^Ω is not relatively cohomologous to a coset of $t\mathbb{Z}$ for any $t \neq 0$. Then $\sigma^2 > 0$, and for a.e. ω, for every open interval (a, b), and for every $z_N, z \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{N}} \to z$,*

$$\mathbb{P}^\omega[S_N^\omega - z_N \in (a,b)] \sim \frac{1}{\sqrt{N}}\left(\frac{\mathrm{e}^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}\right)|a-b| \text{ as } N \to \infty.$$

(2) **Lattice LLT:** *Suppose all the values of f^Ω are integers, and f^Ω is not relatively cohomologous to a coset of $t\mathbb{Z}$ with some integer $t > 1$. Then $\sigma^2 > 0$, and for a.e. ω, for every $z_N \in \mathbb{Z}$ and $z \in \mathbb{R}$ such that $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{N}} \to z$,*

$$\mathbb{P}^\omega[S_N^\omega = z_N] \sim \frac{1}{\sqrt{N}}\left(\frac{\mathrm{e}^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}\right) \text{ as } N \to \infty.$$

Theorem 9.12 *Suppose X^Ω has a finite noise process, and assume (B),(E),(S). If f^Ω is not relatively cohomologous to a constant, then*

(1) There exists a continuously differentiable strictly convex function $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ such that for a.e. $\omega \in \Omega$,

$$\mathcal{F}(\xi) = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}^\omega(\mathrm{e}^{\xi S_N^\omega}) \quad (\xi \in \mathbb{R}).$$

(2) Let $\mathcal{F}'(\pm\infty) := \lim\limits_{\xi\to\pm\infty}\mathcal{F}'(\xi)$, and let $\mathcal{I}_N^\omega(\eta)$ and $\mathcal{I}(\eta)$ denote the Legendre transforms of $\mathcal{G}_N^\omega(\xi) := \frac{1}{N}\log\mathbb{E}^\omega(\mathrm{e}^{\xi S_N^\omega})$ and $\mathcal{F}(\xi)$. Then for a.e. ω, for every $\eta \in (\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$, $\mathcal{I}_N^\omega(\eta) \xrightarrow[N\to\infty]{} \mathcal{I}(\eta)$.

(3) $\mathcal{I}(\eta)$ is continuously differentiable, strictly convex, has compact level sets, is equal to zero at the asymptotic mean μ, and is strictly positive elsewhere.

(4) With probability one, we have full large deviations regime, and the large deviation thresholds and the positivity thresholds of $(\mathsf{X}^\omega, \mathsf{f}^\omega)$ (defined in §7.4) satisfy

$$\mathfrak{c}_- = \mathfrak{r}_- = \frac{\mathcal{F}'(-\infty) - \mu}{\sigma^2} = \lim_{N\to\infty}\frac{\text{ess inf}[S_N^\omega - \mathbb{E}^\omega(S_N^\omega)]}{\sigma^2 N}, \quad \mathfrak{c}_+ = \mathfrak{r}_+ = \frac{\mathcal{F}'(+\infty) - \mu}{\sigma^2} = \lim_{N\to\infty}\frac{\text{ess sup}[S_N^\omega - \mathbb{E}^\omega(S_N^\omega)]}{\sigma^2 N}.$$

Corollary 9.13 (Kifer) *Assume the conditions of the previous theorem, and let* $\overline{\mathcal{I}}(\eta) := \sup_{\xi\in\mathbb{R}}\{\xi\eta - \mathcal{F}(\xi)\}$*. Then* $\overline{\mathcal{I}} = \mathcal{I}$ *on* $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$*,* $\overline{\mathcal{I}} = +\infty$ *outside* $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$*, and for a.e.* ω*,* S_N^ω/N *satisfies the large deviations principle with the rate function* $\overline{\mathcal{I}}(\eta)$*:*

(1) $\limsup_{N\to\infty} \dfrac{1}{N}\log \mathbb{P}^\omega[S_N^\omega/N \in K] \leq -\inf_{\eta\in K} \overline{\mathcal{I}}(\eta)$ *for all closed sets* $K \subset \mathbb{R}$*.*

(2) $\liminf_{N\to\infty} \dfrac{1}{N}\log \mathbb{P}^\omega[S_N^\omega/N \in G] \geq -\inf_{\eta\in G} \overline{\mathcal{I}}(\eta)$ *for all open sets* $G \subset \mathbb{R}$*.*

Proof Use Lemma A.3 and the Gärtner-Ellis Theorem (Appendix A). □

Theorem 9.14 *Suppose* X^Ω *have a finite noise process, and* $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$ *satisfies (C) or (D), and each of (S), (E), (G), (B).*

(1) **Non-Lattice LLT for Large Deviations:** *Assume* f^Ω *is not relatively cohomologous to a coset of* $t\mathbb{Z}$ *for any* $t \neq 0$*. Then for a.e.* ω*, for every open interval* (a, b)*, and for every* $z_N \in \mathbb{R}$ *s.t.* $\frac{z_N}{N} \to z$*, if* $\frac{z-\mu}{\sigma^2} \in (\mathfrak{c}_-, \mathfrak{c}_+)$ *then*

$$\mathbb{P}^\omega[S_N^\omega - z_N \in (a,b)] = [1+o(1)]\cdot \frac{e^{-V_N^\omega \mathcal{I}_N^\omega(z_N/V_N^\omega)}}{\sqrt{2\pi\sigma^2 N}}|a-b|\rho_N^\omega\left(\tfrac{z_N - \mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}\right)\times \frac{1}{|a-b|}\int_a^b e^{-t\xi_N^\omega\left(\frac{z_N-\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}\right)}dt,$$

with $\rho_N^\omega, \xi_N^\omega$ *as in Theorem 7.8(4).*

(2) **Lattice LLT for Large Deviations:** *Suppose* f^Ω *is integer valued, and not relatively cohomologous to a coset of* $t\mathbb{Z}$ *for any integer* $t > 1$*. Then for a.e.* ω*, for every* $z_N \in \mathbb{Z}$ *such that* $\frac{z_N}{N} \to z$*, if* $\frac{z-\mu}{\sigma^2} \in (\mathfrak{c}_-, \mathfrak{c}_+)$ *then*

$$\mathbb{P}^\omega[S_N^\omega = z_N] = [1+o(1)]\cdot \frac{e^{-V_N^\omega \mathcal{I}_N^\omega(z_N/V_N^\omega)}}{\sqrt{2\pi\sigma^2 N}}\rho_N^\omega\left(\tfrac{z_N - \mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}\right),$$

with ρ_N^ω *as in Theorem 7.8(4).*

(3) For a.e. ω*, for each sequence* z_N *such that* $\lim_{N\to\infty} \dfrac{z_N - \mathbb{E}^\omega(S_N^\omega)}{N} = 0$*, it holds that*

$$V_N^\omega \mathcal{I}_N^\omega\left(z_N/V_N^\omega\right) = \frac{1+o(1)}{2\sigma^2}\left(\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{N}}\right)^2 \text{ as } N\to\infty.$$

So far we have focused on MCRE with finite noise spaces. We will now address the case of MCRE with infinite noise spaces. First, here are two examples of the new phenomena which may occur.

Example 9.15 (No Asymptotic Mean) For MCRE with infinite noise spaces, it is possible that $\mathbb{E}(S_N^\omega)/N$ oscillates for a.e. ω, without converging.

Proof. Let $\{Y_n\}_{n\in\mathbb{Z}}$ be the simple random walk on $\mathbb{Z}$, started from the stationary (infinite) distribution. In Example 9.5, we built a noise process $(\Omega, \mathscr{F}, m, T)$ such that $\Omega = \mathbb{Z}^\mathbb{Z}$, T is the left shift, and $\omega = (\omega_n)_{n\in\mathbb{Z}} \in \Omega$ is m-distributed like $(Y_n)_{n\in\mathbb{Z}}$.

Let $\mathsf{X} = \{X_n\}_{n\in\mathbb{Z}}$ be a sequence of bounded iid random variables with positive expectation c_0, and independent of $\{Y_n\}$. Let X^Ω be the MCRE with noise process $(\Omega, \mathscr{F}, m, T)$, such that $\mathsf{X}^\omega = \mathsf{X}$ for all ω (so the generators $\pi(\omega, x, dy)$ and μ_ω are independent of ω). Let f^Ω be the function $f(\omega, x, y) := 1_{[0,\infty)}(\omega)x$.

For the pair $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$, $\mathbb{E}^\omega[f_n^\omega(X_n^\omega, X_{n+1}^\omega)] = c_0 1_{[\omega_n \geq 0]}$, and so $\dfrac{1}{N}\mathbb{E}^\omega(S_N^\omega) = \dfrac{c_0}{N}\#\{1 \leq n \leq N : \omega_n \geq 0\}$.

We claim that the RHS oscillates a.e. without converging. The liminf and limsup of the RHS are T-invariant, whence by ergodicity, constant. If the claim were false, then $W_N := \frac{1}{N}\#\{1 \leq n \leq N : Y_n \geq 0\}$ would have converged a.e. to a constant. But this contradicts the arcsine law for the simple random walk.

Example 9.16 (Pathological Variance Growth) For MCRE with an infinite noise process, it is possible that $V_N^\omega \to \infty$ a.e., $V_N^\omega = o(N)$ a.e., and that there is no sequence of constants $a_N > 0$ (independent of ω) such that $V_N^\omega \sim a_N$ for a.e. ω.

Proof. Let X_n be iid bounded real random variables with variance one and distribution ζ. Let $f_n(x) = x$. Let $(\Omega, \mathscr{F}, m, T)$ be an infinite noise process, and fix $E \in \mathscr{F}$ of finite positive measure. Let $\pi(\omega, x, \mathrm{d}y) := \zeta(\mathrm{d}y)$, $f(\omega, x, y) := 1_E(\omega)x$.

Then $S_N^\omega = \sum_{n=1}^{N} 1_E(T^n\omega)X_n$, and $V_N^\omega = \sum_{n=1}^{N} 1_E(T^n\omega)$. We now appeal to general results from infinite ergodic theory. Let $(\Omega, \mathscr{F}, m, T)$ be an ergodic, invertible, measure preserving map on a non-atomic σ-finite infinite measure space. Let $L_+^1 := \{A \in L^1(\Omega, \mathscr{F}, m) : A \geq 0, \int A\mathrm{d}m > 0\}$. Then,

(1) $\sum_{n=1}^{N} A \circ T^n = \infty$ almost everywhere for all $A \in L_+^1$; (2) $\frac{1}{N}\sum_{n=1}^{N} A \circ T^n \xrightarrow[N\to\infty]{} 0$ almost everywhere for all $A \in L^1$;
(3)Let a_N be a sequence of positive real numbers, then
either $\liminf_{N\to\infty} \frac{1}{a_N}\sum_{n=1}^{N} A \circ T^n = 0$ a.e. for all $A \in L_+^1$; or $\limsup_{N\to\infty} \frac{1}{a_N}\sum_{n=1}^{N} A \circ T^n = \infty$ a.e. for all $A \in L_+^1$ (or both).

So $\nexists a_N > 0$ s.t. $\sum_{n=1}^{N} A(T^n\omega) \sim a_N$ for a.e. ω, even for a single $A \in L_+^1$. (See [1]: (1) is the Halmos recurrence theorem (see also Lemma 9.23); (2) follows from the ratio ergodic theorem; and (3) is a theorem of J. Aaronson.)

Specializing to the case $A = 1_E$ we find that $V_N^\omega \to \infty$ a.e.; $V_N^\omega = o(N)$ a.e. as $N \to \infty$; and $\nexists a_N$ so that $V_N^\omega \sim a_N$ for a.e. $\omega \in \Omega$. □

We continue to present our general results on MCRE with infinite noise spaces.

Theorem 9.17 *Suppose* X^Ω *has an infinite noise process, on a non-atomic* σ*-finite measure space. Assume (B),(E),(S).*

(1) If f^Ω *is relatively cohomologous to a constant, then* $V_N^\omega \leq C$ *for all N, for a.e.* ω*, where* $C = C(\epsilon_0, K)$ *is a constant.*
(2) If f^Ω *is not relatively cohomologous to a constant then* $V_N^\omega \to \infty$ *for a.e.* ω*.*

Theorem 9.18 *Suppose* X^Ω *has an infinite noise process, on a non-atomic* σ*-finite measure space. Assume* $(\mathsf{X}^\Omega, \mathsf{f}^\Omega)$ *satisfies (C) or (D), and each of (S),(E),(G),(B).*

(a) **Non-Lattice LLT:** *Suppose* f^Ω *is not relatively cohomologous to a coset of* $t\mathbb{Z}$ *for any* $t \neq 0$*. Then for a.e.* ω*, for every open interval* (a, b)*, and for every* $z_N, z \in \mathbb{R}$ *such that* $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{V_N^\omega}} \to z$*,*

$$\mathbb{P}^\omega[S_N^\omega - z_N \in (a, b)] \sim \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N^\omega}}|a - b| \text{ as } N \to \infty.$$

(b) **Lattice LLT:** *Suppose that all the values of* f^Ω *are integers, and* f^Ω *is not relatively cohomologous to a coset of* $t\mathbb{Z}$ *with an integer* $t > 1$*. Then for a.e.* ω*, for every* $z_N \in \mathbb{Z}$ *such that* $\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{\sqrt{V_N^\omega}} \to z$*,*

$$\mathbb{P}^\omega[S_N^\omega = z_N] \sim \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi V_N^\omega}} \text{ as } N \to \infty.$$

Theorem 9.19 *Under the assumptions of Theorem 9.18*

(1) Suppose f^Ω *is not relatively cohomologous to a coset of* $t\mathbb{Z}$ *for any* $t \neq 0$*. Then for a.e.* ω*,* $(\mathsf{X}^\omega, \mathsf{f}^\omega)$ *satisfies the non-lattice LLT for large deviations (Theorem 7.8 parts (1),(2), and (4)).*

(2) Suppose that all the values of f^Ω *are integers, and* f^Ω *is not relatively cohomologous to a coset of* $t\mathbb{Z}$ *with an integer* $t > 1$*. Then for a.e.* ω*,* $(\mathsf{X}^\omega, \mathsf{f}^\omega)$ *satisfies the lattice LLT for large deviations (Theorem 7.8 (1),(3), and (4)).*

9.3 Proofs

9.3.1 Existence of Stationary Measures

We prove Lemma 9.8. We need the following standard fact:

Lemma 9.20 *Let Ω be a measurable space, and $\mathfrak{S}$ be a compact metric space. Let ζ_ω ($\omega \in \Omega$) be Borel probability measures on $\mathfrak{S}$. If $\omega \mapsto \int \varphi d\zeta_\omega$ is measurable for all continuous $\varphi : \mathfrak{S} \to \mathbb{R}$, then $\omega \mapsto \int \varphi d\zeta_\omega$ is measurable for all bounded Borel $\varphi : \mathfrak{S} \to \mathbb{R}$.*

Proof We will show that $\mathscr{F} := \{\varphi : \mathfrak{S} \to \mathbb{R} : \ \omega \mapsto \int \varphi d\zeta_\omega \text{ is measurable}\}$ contains all bounded real-valued measurable functions on $\mathfrak{S}$.

(1) $\mathscr{F}$ contains $\mathscr{K} := \{1_K : K \subset \mathfrak{S} \text{ is compact}\}$, because each 1_K is the pointwise decreasing limit of a uniformly bounded sequence of continuous functions.
(2) The collection of compact subsets of $\mathfrak{S}$ generates the Borel σ-algebra, and is closed under finite intersections.
(3) $\mathscr{F}$ is closed under finite linear combinations, and under bounded increasing pointwise limits.

By the functional monotone class theorem, $\mathscr{F} \supset \{\text{bounded measurable functions}\}$. □

Proof of Lemma 9.8: $\mathfrak{S}$ is compact. Fix $\omega \in \Omega$. Let C denote the cone of non-negative continuous functions on $\mathfrak{S}$, with the supremum norm. The interior of C is C^+, the open cone of strictly positive continuous functions.

Let $L_k^\omega : C \to C$ denote the operator $(L_k^\omega \varphi)(x) = \int \varphi(y)\pi(T^k\omega, x, dy)$. The right-hand-side is in C, because of the assumptions on $\pi(\omega, x, dy)$. Observe that

$$L_k^\omega = L_{k-\ell}^{T^\ell \omega} \tag{9.4}$$

Since $\mathfrak{S}$ is compact and metrizable, the space of probability measures on $\mathfrak{S}$ is weak-star sequentially compact. Using a diagonal argument, we can construct $N_k = N_k(\omega) \to \infty$ and Borel probability measures μ_i' on $\mathfrak{S}$ as follows: $\frac{1}{N_k}\sum_{n=1}^{N_k}(L_{-n}^{T^i\omega} \cdots L_{-1}^{T^i\omega})^* \delta_x \xrightarrow[k\to\infty]{} \mu_i'$ weak star, for every $i \in \mathbb{Z}$. If $T^i\omega = T^j\omega$, then $\mu_i' = \mu_j'$, and we may define $\mu_{T^i\omega}' := \mu_i'$. By (9.4), for every continuous function φ,

$$\mu_i'(L_0^{T^i\omega}\varphi) = \mu_i'(L_{-1}^{T^{i+1}\omega}\varphi) = \lim_{k\to\infty}\frac{1}{N_k}\sum_{n=2}^{N_k+1}(L_{-n}^{T^{i+1}\omega}\cdots L_{-1}^{T^{i+1}\omega}\varphi)(x) = \mu_{i+1}'(\varphi). \tag{9.5}$$

Equivalently, $\int \pi(T^i\omega, x, dy)\mu_{T^i\omega}'(dx) = \mu_{T^{i+1}\omega}'$ for all $i \in \mathbb{Z}$. We obtain (S):

$$\mu_{T\omega}' = \int \pi(\omega, x, dy)\mu_\omega'(dx). \tag{9.6}$$

But it is not clear that $\omega \mapsto \mu_\omega'$ is measurable. We will address this now.

Let $T_n^\omega := L_n^\omega L_{n+1}^\omega$. By (E), for every $\varphi \in C$,

$$\begin{aligned}(T_n^\omega\varphi)(x) &= \iint \varphi(z)p(T^n\omega, x, y)p(T^{n+1}\omega, y, z)\mu_{T^{n+1}\omega}(dy)\mu_{T^{n+2}\omega}(dz)\\ &\in \left[\epsilon_0 \int \varphi(z)\mu_{T^{n+2}\omega}(dz), \epsilon_0^{-2}\int \varphi(z)\mu_{T^{n+2}\omega}(dz)\right].\end{aligned}$$

Therefore, the diameter of $T_n^\omega(C)$ with respect to the Hilbert projective metric of C is no larger than $6\log(1/\epsilon_0)$ (see Appendix B).

Call the projective metric d_C and let $\theta := \tanh(\frac{3}{2}\log(1/\epsilon_0))$ (a number in $(0,1)$). By Birkhoff's Theorem (Theorem B.6), for every $\varphi \in C^+$,

$$d_C(L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi, L^{\omega}_{-2n}\cdots L^{\omega}_{-1}1) \le \theta^{n-1} d_C(T^{\omega}_{-2}\varphi, T^{\omega}_{-2}1) \le 6\log(1/\epsilon_0)\theta^{n-1}.$$

Since $L^{\omega}_{-2n}\cdots L^{\omega}_{-1}1 \equiv 1$, this implies the existence of positive constants $M^{\omega}_n(\varphi)$ and $m^{\omega}_n(\varphi)$ such that $\log\left(\frac{M^{\omega}_n(\varphi)}{m^{\omega}_n(\varphi)}\right) \le 6\log(1/\epsilon_0)\theta^{n-1}$, and $m^{\omega}_n(\varphi) \le L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi \le M^{\omega}_n(\varphi)$.Therefore

$$\|m^{\omega}_n(\varphi)^{-1} L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi - 1\|_\infty \le \delta_n := \exp[6\log(1/\epsilon_0)\theta^{n-1}] - 1. \tag{9.7}$$

By (9.5), for every continuous function ψ, $\mu'_{T^{-k}\omega}(L_0^{T^{-k}\omega}\psi) = \mu'_{T^{-k+1}\omega}(\psi)$. So, $\mu'_{T^{-2n}\omega}(L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi) \equiv \mu'_{T^{-2n}\omega}(L_0^{T^{-2n}\omega}\cdots L_0^{T^{-1}\omega}\varphi) = \mu'_{T^{-2n+1}\omega}(L_0^{T^{-2n+1}\omega}\cdots L_0^{T^{-1}\omega}\varphi) = \cdots = \mu'_{T^{-1}\omega}(L_0^{T^{-1}\omega}\varphi) = \mu'_{\omega}(\varphi)$.

So after integrating $m^{\omega}_n(\varphi)^{-1}L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi$ with respect to $\mu'_{T^{-2n}\omega}$, (9.7) gives

$$|m^{\omega}_n(\varphi)^{-1}\mu'_{\omega}(\varphi) - 1| \le \delta_n \to 0. \tag{9.8}$$

In particular, $m^{\omega}_n(\varphi) \xrightarrow[n\to\infty]{} \int \varphi \mathrm{d}\mu'_{\omega}$, whence for every $\varphi \in C^+$ and $x \in \mathfrak{S}$,

$$(L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi)(x) \xrightarrow[n\to\infty]{} \int \varphi \mathrm{d}\mu'_{\omega}. \tag{9.9}$$

For fixed $x \in \mathfrak{S}$, $\omega \mapsto (L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi)(x)$ is measurable for all $\varphi \in C^+$. As $C^+ - C^+ = C(\mathfrak{S})$, $\omega \mapsto \mu'_{\omega}(\varphi)$ is measurable for all continuous φ. By Lemma 9.20, $\omega \mapsto \mu'_{\omega}(\omega)$ is measurable for all bounded Borel functions φ. We proved part (1).

To see part (2), fix ω, and let Y denote the inhomogeneous Markov chain with initial distribution $\mu'_{T^{-3}\omega}$ and transition kernels $\pi_n(x, \mathrm{d}y) := \pi(T^{n-3}\omega, x, \mathrm{d}y)$.

Since X^{Ω} is uniformly elliptic, Y is uniformly elliptic, with the same ellipticity constant, and with background measures $\mu_n := \mu'_{T^{n-3}\omega}$. By Proposition 2.8, $\epsilon_0 \le \frac{\mathbb{P}(Y_n \in E)}{\mu_{T^{n-3}\omega}(E)} \le \epsilon_0^{-1}$ for all $E \in \mathscr{B}(\mathfrak{S})$ and $n \ge 3$. By (9.6), Y satisfies (S), and the numerator equals $\mu'_{T^{n-3}\omega}(E)$. So $\mu'_{T^{n-3}\omega} \ll \mu_{T^{n-3}\omega}$ for all $n \ge 3$, and the Radon-Nikodym derivative is bounded between ϵ_0 and ϵ_0^{-1}. We proved part (2).

We proceed to part (3). Suppose Ω is a metric space, and $T : \Omega \to \Omega$ and $(\omega, x) \mapsto \pi(\omega, x, \mathrm{d}y)$ are continuous. By (9.8), for every $\varphi \in C^+$, $\sup_{\omega\in\Omega} |m^{\omega}_n(\varphi) - \mu'_{\omega}(\varphi)| \le \frac{\delta_n\|\varphi\|_\infty}{1-\delta_n} \xrightarrow[n\to\infty]{} 0$.

By (9.7), for each fixed $x \in \mathfrak{S}$, $(L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi)(x) \xrightarrow[n\to\infty]{} \mu'_{\omega}(\varphi)$ uniformly on Ω. By our assumptions, $(L^{\omega}_{-2n}\cdots L^{\omega}_{-1}\varphi)(x)$ are continuous in ω. Since the uniform limit of continuous functions is continuous, $\omega \mapsto \mu'_{\omega}(\varphi)$ is continuous. □

9.3.2 The Essential Range is Almost Surely Constant

From this point, and until the end of section 9.3, we assume that $(\mathsf{X}^{\Omega}, \mathsf{f}^{\Omega})$ *are as in* (9.1) *and* (9.2), *and that (B),(E),(S) hold. Unless stated otherwise, we allow the noise process to be infinite.*

The purpose of this section is to prove the following result:

Proposition 9.21 *There exist closed subgroups* $H, G_{ess} \le \mathbb{R}$ *s.t. for m–a.e.* ω, *the co-range of* $(\mathsf{X}^{\omega}, \mathsf{f}^{\omega})$ *equals* H, *the essential range of* $(\mathsf{X}^{\omega}, \mathsf{f}^{\omega})$ *equals* G_{ess}, *and*

$$G_{ess} = \begin{cases} \mathbb{R} & H = \{0\}, \\ \frac{2\pi}{t}\mathbb{Z} & H = t\mathbb{Z}, \quad t \ne 0, \\ \{0\} & H = \mathbb{R}. \end{cases}$$

We call H and G_{ess} the **a.s. co-range** and **a.s. essential range**.

We need a few preliminary comments on the structure constants of $(\mathsf{X}^\omega, \mathsf{f}^\omega)$. Fix an element ω in the noise space, and let $\mathrm{Hex}(\omega)$ denote the probability space of position 3 hexagons for X^ω. Let m_ω denote the hexagon measure, as defined in §2.3.1. Recall the definition of the balance $\Gamma(P)$ of a hexagon P, and define

$$u(\omega) := \mathbb{E}(|\Gamma(P)|^2)^{1/2}, \quad d(\omega, \xi) := \mathbb{E}(|e^{i\xi\Gamma(P)} - 1|^2)^{1/2}, \quad \text{(expectation on } P \in \mathrm{Hex}(\omega) \text{ w.r.t. } m_\omega).$$

By (S), the probability space of position $n+3$ hexagons for X^ω is $(\mathrm{Hex}(T^n\omega), m_{T^n\omega})$. Therefore the structure constants of $(\mathsf{X}^\omega, \mathsf{f}^\omega)$ are given by

$$d_{n+3}(\xi, \mathsf{f}^\omega) = d(T^n\omega, \xi) \quad \text{and} \quad u_{n+3}(\mathsf{f}^\omega) = u(T^n\omega) \ \ (n \geq 0). \tag{9.10}$$

Lemma 9.22 *$u(\cdot), d(\cdot,\cdot)$ are Borel measurable, and for every ω, $d(\omega,\cdot)$ is continuous. In addition, if X^Ω satisfies (C), then $u(\cdot), d(\cdot,\cdot)$ are continuous.*

Proof The lemma follows from the explicit formulas for the hexagon measure and the function $\Gamma : \mathfrak{S}^6 \to \mathbb{R}$. We omit the details, which are routine. □

Proof of Proposition 9.21. Let $H_\omega := H(\mathsf{X}^\omega, \mathsf{f}^\omega)$ be the essential range of $(\mathsf{X}^\omega, \mathsf{f}^\omega)$. By Theorem 4.3, H_ω is either $\mathbb{R}$ or $t\mathbb{Z}$ for some $t = t(\omega) \geq 0$. By (9.10) $D_N(\xi, \omega) := \sum_{n=3}^{N} d_n(\xi, \mathsf{f}^\omega)^2 \equiv \sum_{n=0}^{N-3} d(T^n\omega, \xi)^2$.

Step 1: *$U(a,b) := \{\omega \in \Omega : D_N(\cdot,\omega) \xrightarrow[N\to\infty]{} \infty$ uniformly on $(a,b)\}$ is measurable and T-invariant $\forall a < b$.*

Proof. Observe that $d^2 \leq 4$, therefore $|D_N(\xi, T\omega) - D_N(\xi,\omega)| \leq 8$. It follows that $U(a,b)$ is T-invariant.

Measurability is because of the identity $U(a,b) = \left\{\omega \in \Omega : \begin{array}{l} \forall M \in \mathbb{Q}\ \exists N \in \mathbb{N} \text{ s.t.} \\ \text{for all } \xi \in (a,b) \cap \mathbb{Q},\ D_N(\omega,\xi) > M \end{array}\right\}$. The inclusion $\subset$ is obvious. The inclusion $\supset$ is because if $\omega \notin U(a,b)$ then for some $M \in \mathbb{Q}$, for all $N \in \mathbb{N}$ there exists some $\eta_N \in (a,b)$ such that $D_N(\omega, \eta_N) < M$, whence by the continuity of $\eta \mapsto D_N(\omega,\eta)$ there is some $\xi_N \in (a,b) \cap \mathbb{Q}$ such that $D_N(\omega, \xi_N) < M$. So $\omega \notin U(a,b) \Rightarrow \omega \notin$ RHS.

Step 2: *The sets $\Omega_1 := \{\omega \in \Omega : H_\omega = \{0\}\}$, $\Omega_2 := \{\omega \in \Omega : H_\omega = \mathbb{R}\}$, and $\Omega_3 := \{\omega \in \Omega : \exists t \neq 0 \text{ s.t. } H_\omega = t\mathbb{Z}\}$ are measurable and T-invariant. Therefore by ergodicity, for each i, either $m(\Omega_i) = 0$ or $m(\Omega_i^c) = 0$.*

Proof. Recall that for Markov chains, $D_N \to \infty$ uniformly on compact subsets of the complement of the co-range (Theorem 4.9). So $\Omega_1 = \bigcap_{n=1}^{\infty} U\left(\frac{1}{n}, n\right)$, $\Omega_2 = \bigcap_{0<a<b \text{ rational}} U(a,b)^c$, $\Omega_3 = \Omega_1^c \cap \Omega_2^c$.

By Step 1, Ω_i are T-invariant and measurable. Since T is ergodic, these sets are either of measure zero or of full measure.

By Theorem 4.4, if Ω_1 has full measure, then the essential range is a.e. $\mathbb{R}$, and if Ω_2 has full measure, then the essential range is a.e. $\{0\}$. It remains to consider the case when Ω_3 has full measure.

Step 3: *If Ω_3 has full measure, then there exist $t \neq 0$ such that $\Omega_3(t) := \{\omega \in \Omega : H_\omega = t\mathbb{Z}\}$ has full measure, and then the essential range is a.e. $(2\pi/t)\mathbb{Z}$.*

Proof. For every $\omega \in \Omega_3$ there exists $t(\omega) > 0$ such that $H_\omega = t(\omega)\mathbb{Z}$. We can characterize $t(\omega)$ as follows:

$$t(\omega) = \sup\left\{t \in \mathbb{Q} \cap (0,\infty) : \begin{array}{c} D_N(\omega,\cdot) \to \infty \text{ uniformly} \\ \text{on compact subsets of } (0,t) \end{array}\right\}.$$

It is clear from this expression that $t(T\omega) = t(\omega)$, and that for every $A > 0$, $[t(\omega) \geq A] = \bigcap_{0<a<b<A \text{ rational}} U(a,b)$.

So $t(\cdot)$ is a measurable T-invariant function.

By ergodicity, there is a constant t such that $t(\omega) = t$ for a.e. ω. So $H_\omega = t\mathbb{Z}$ a.e. By Theorem 4.4, $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = (2\pi/t)\mathbb{Z}$ a.e. □

9.3.3 Variance Growth

In this section we prove Theorems 9.10 and 9.17 on the behavior of V_N^ω, as $N \to \infty$.

Lemma 9.23 *Suppose $(\Omega, \mathscr{F}, m, T)$ is an invertible, ergodic, measure preserving map on a probability space or a non-atomic infinite measure space. Let $A : \Omega \to \mathbb{R}$ be a non-negative measurable function. Either $A = 0$ a.e., or $\sum_{n\geq 0} A \circ T^n = \infty$ a.e.*

Proof If $m(\Omega) < \infty$, then the lemma follows from the pointwise ergodic theorem. If $m(\Omega) \leq \infty$, then we can use the following well-known argument [1].

If A is not equal to 0 a.e., then there is an $\varepsilon > 0$ such that $E := \{\omega \in \Omega : A(\omega) \geq \varepsilon\}$ has positive measure. We claim that

$$\sum_{n\geq 0} 1_E(T^n\omega) = \infty \text{ a.e. on } E. \tag{9.11}$$

Since $A \geq \varepsilon 1_E$, (9.11) implies that $\sum_{n\geq 0} A(T^n\omega) = \infty$ almost everywhere on E, whence (by ergodicity) almost everywhere on Ω.

Suppose by way of contradiction that (9.11) fails, then there exists N s. t. $W := \{\omega \in E : \sum_{n=0}^{\infty} 1_E(T^n\omega) = N\}$ has positive measure.

The invertibility and measurability of T imply that $T^n(W)$ are measurable and pairwise disjoint. By non-atomicity, we can break $W = W_1 \cup W_2$ where W_i are measurable, disjoint, and with positive measure. By invertibility, $\widehat{W}_i := \bigcup_{n\in\mathbb{Z}} T^n W_i$ are disjoint T-invariant sets with positive measure. But this contradicts ergodicity. □

Part 1: Either $\mathbf{V_N^\omega}$ is Bounded Almost Surely , or $\mathbf{V_N^\omega \to \infty}$ Almost Surely. Recall that $|f| < K$, and ϵ_0 is an ellipticity constant for X^ω. By Theorem 3.7 and (9.10) there are positive constants $C_i = C_i(\epsilon_0, K)$ $(i = 1, 2)$ such that for all N, $C_1^{-1} \sum_{n=3}^{N} u(T^n\omega)^2 - C_2 \leq V_N^\omega \leq C_1 \sum_{n=3}^{N} u(T^n\omega)^2 + C_2$.

If $u(\omega) = 0$ m-a.e., then for a.e. ω, $V_N^\omega \leq C_2$ for all N. Otherwise, by Lemma 9.23, $\sum_{n=3}^{N} u(T^n\omega)^2 \xrightarrow[N\to\infty]{} \infty$, whence $V_N^\omega \to \infty$ almost everywhere.

Part 2: Linear Growth of Variance when $\mathbf{V_N^\omega \to \infty}$ a.e. and $\mathbf{m(\Omega) = 1}$. Suppose $m(\Omega) = 1$ and $V_N^\omega \to \infty$ almost surely. We claim that

$$\exists \sigma^2 > 0 \text{ s.t. } V_N^\omega \sim N\sigma^2 \text{ a.s.} \tag{9.12}$$

Let $\sigma_0^2 := \int_\Omega u^2 dm$. This is a finite number, because $\|u\|_\infty \leq 6K$ by (B), and $m(\Omega) = 1$. This is a positive number, because as we saw in part 1, if $u = 0$ a.e., then $V_N^\omega = O(1)$ a.e. contrary to our assumptions.

By the pointwise ergodic theorem, $\sum_{n=3}^{N} u(T^n\omega)^2 = [1 + o(1)]\sigma_0^2 N$. Hence

$$V_N^\omega \geq [1 + o(1)]C_1(\epsilon_0, K)^{-1} N\sigma_0^2 \to \infty. \tag{9.13}$$

$V_N^\omega = \sum_{n=1}^{N} \mathrm{Var}^\omega(F_n) + 2\sum_{n=1}^{N}\sum_{k=1}^{N-n} \mathrm{Cov}^\omega(F_n, F_{n+k})$, where $F_n := f(T^n\omega, X_n^\omega, X_{n+1}^\omega)$. By (S), $\{X_i^{T^{n-1}\omega}\}_{i\geq 1}$ is equal in distribution to $\{X_i^\omega\}_{i\geq n}$. So $\mathrm{Cov}^\omega(F_n, F_{n+k}) = \psi_k(T^{n-1}\omega)$, where $\psi_k(\omega) := \mathrm{Cov}^\omega(f(T\omega, X_1^\omega, X_2^\omega), f(T^{k+1}\omega, X_{k+1}^\omega, X_{k+2}^\omega))$. Thus $V_N^\omega = \sum_{n=0}^{N-1} \psi_0(T^n\omega) + 2\sum_{n=0}^{N-1}\sum_{k=1}^{N-n} \psi_k(T^n\omega)$. The next step is to find the limit of $(1/N)\times$ RHS as $N \to \infty$.

By the pointwise ergodic theorem, for a.e. ω, $\lim\limits_{N\to\infty}\frac{1}{N}\sum\limits_{n=0}^{N-1}\psi_0(T^n\omega)=\int\psi_0\mathrm{d}m$. Next we recall that $\|\psi_k\|_\infty\le C_{mix}\|f\|_\infty^2\theta^k$, with $C_{mix}>0$ and $0<\theta<1$ which depend only on ϵ_0 (Proposition 2.13). Therefore for every M,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=1}^{N-n}\psi_k(T^n\omega)=\Big[\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=1}^{M-1}\psi_k(T^n\omega)\Big]+O(\theta^M),$$

whence by the ergodic theorem $\lim\limits_{N\to\infty}\dfrac{1}{N}\sum\limits_{n=0}^{N-1}\sum\limits_{k=1}^{N-n}\psi_k(T^n\omega)=\sum\limits_{k=1}^{\infty}\int\psi_k\mathrm{d}m$, with the last sum converging exponentially fast. In summary, $\dfrac{1}{N}V_N^\omega\xrightarrow[N\to\infty]{}\sigma^2:=\int\Big(\psi_0+2\sum\limits_{k=1}^{\infty}\psi_k\Big)\mathrm{d}m$. By (9.13), $\frac{1}{N}V_N^\omega\not\to 0$ a.e, so $\sigma^2>0$, and (9.12) is proved.

We now show that the following properties are equivalent:

(a) f^Ω is relatively cohomologous to a constant; (b) $\operatorname{ess\,sup}_{\omega\in\Omega}(\sup_{N\in\mathbb{N}}V_N^\omega)<\infty$; (c) $m\{\omega: V_N^\omega \text{ is bounded}\}>0$.

Part 3: (a)⇒(b)⇒(c): Suppose f^Ω is relatively cohomologous to a constant. Then there are uniformly bounded measurable functions $a:\Omega\times\mathfrak{S}\to\mathbb{R}$ and $c:\Omega\to\mathbb{R}$ such that for m-a.e. ω, for every n, and with full probability with respect to the distribution of $\{X_k^\omega\}$, $f_n^\omega(X_n^\omega,X_{n+1}^\omega)=a(T^n\omega,X_n^\omega)-a(T^{n+1}\omega,X_{n+1}^\omega)+c(T^n\omega)$.

Summing over n, we see that for a.e. ω, $\forall N$, $|S_N^\omega-\sum\limits_{n=1}^{N}c(T^n\omega)|=|a(T\omega,X_1^\omega)-a(T^{N+1}\omega,X_{N+1}^\omega)|\le 2\sup|a|$.

Recalling that the variance of a random variable S is $\inf\limits_{c\in\mathbb{R}}\|S-c\|_2^2$, we deduce that for a.e. ω, $V_N^\omega\le 4\sup a^2$, whence (b). Clearly (b)⇒(c).

Part 4: (c)⇒(a): We saw in the proof of part (1) that if (c) holds, then $V_N^\omega=O(1)$ a.e. and $u(\omega)=0$ a.e. Since m is T-invariant, $\sum u_n^2(\mathsf{X}^\omega,\mathsf{f}^\omega)=0$ a.e. Applying the gradient lemma to X^ω, we find bounded functions g_n^ω and constants c_n^ω such that $f_n^\omega(X_n^\omega,X_{n+1}^\omega)=g_n^\omega(X_n^\omega)-g_{n+1}^\omega(X_{n+1}^\omega)+c_n^\omega$ $\mathbb{P}^\omega$–a.s. Moreover, the proof of the gradient lemma shows that we can take

$$c_n^\omega=\mathbb{E}^\omega[f_{n-2}^\omega(X_{n-2}^\omega,X_{n-1}^\omega)],\quad g_n^\omega(z)=\mathbb{E}^\omega\Big(f_{n-2}^\omega(Z_{n-2}^\omega,Y_{n-1}^\omega)+f_{n-1}^\omega(Y_{n-1}^\omega,X_n^\omega)\Big|X_n^\omega=z\Big),$$

where $\underline{L}_n^\omega:=(Z_{n-2}^\omega,Y_{n-1}^\omega,X_n^\omega)$ is the n-th element of the ladder process of X^ω. By (S), $g_n^\omega(x)=a(T^n\omega,x)$ and $c_n^\omega=c(T^n\omega)$, where $a(\cdot,\cdot),c(\cdot)$ are bounded measurable functions. So f^Ω is relatively cohomologous to a constant. □

9.3.4 Irreducibility and the LLT

In this section we prove Theorems 9.11 and 9.18. The main ingredient in the proof is the following criterion for irreducibility:

Proposition 9.24 *Suppose $(\mathsf{X}^\Omega,\mathsf{f}^\Omega)$ satisfies (C) or (D), and each of (S),(E),(G).*

(1) f^ω is irreducible with essential range $\mathbb{R}$ for a.e. ω iff f^Ω is not relatively cohomologous to a coset $t\mathbb{Z}$ for any $t\neq 0$.

(2) Suppose f^Ω is integer valued, then f^ω is irreducible with essential range $\mathbb{Z}$ for a.e. ω iff f^Ω is not relatively cohomologous to a coset $t\mathbb{Z}$ for any $t>1$.

(In this proposition, we allow the noise space to be infinite.)

Lemma 9.25 *Suppose W_1,W_2 are two independent random variables such that for some $a,t\in\mathbb{R}$, $W_1+W_2\in a+t\mathbb{Z}$ with full probability. Then we can decompose $a=a_1+a_2$ so that $W_1\in a_1+t\mathbb{Z}$ a.s., and $W_2\in a_2+t\mathbb{Z}$ a.s.*

Proof Without loss of generality $a = 0$, $t = 2\pi$. Then $|\mathbb{E}(e^{iW_1})| \cdot |\mathbb{E}(e^{iW_2})| = |\mathbb{E}(e^{i(W_1+W_2)})| = 1$. Necessarily, $|\mathbb{E}(e^{iW_k})| = 1 \quad (k = 1, 2)$. Choose a_k such that $\mathbb{E}(e^{i(W_k - a_k)}) = 1$, then $\mathbb{E}(\cos(W_k - a_k)) = 1$, whence $W_k - a_k \in 2\pi\mathbb{Z}$ almost surely. Thus $a_1 + a_2 \in 2\pi\mathbb{Z}$, and there is no problem to adjust a_1 to get $a_1 + a_2 = 0$. □

Lemma 9.26 *Let Ω be a measurable space, $\mathfrak{S}$ be a separable metric space, and $\psi : \Omega \times \mathfrak{S} \to \mathbb{R}$ be Borel. If, for each ω, $\psi(\omega, \cdot)$ is continuous on $\mathfrak{S}$ and positive somewhere, then there exists a measurable $x : \Omega \to \mathfrak{S}$ such that $\psi(\omega, x(\omega)) > 0$.*

Proof Fix a countable dense set $\{x_i\} \subset \mathfrak{S}$. For every ω there exists an i such that $\psi(\omega, x_i) > 0$. So $i(\omega) := \min\{i \in \mathbb{N} : \psi(\omega, x_i) > 0\}$ is well-defined and Borel measurable. Take $x(\omega) := x_{i(\omega)}$. □

Proof of Proposition 9.24 We begin with part 1 of the proposition.

Proof of (⇒): Suppose $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{R}$ for a.e. ω, and assume by way of contradiction that f^Ω *is* relatively cohomologous to a coset $t\mathbb{Z}$ for some t.

It is easy to see that in this case there are measurable functions $g(\omega, x, y)$ and $c(\omega)$ so that for m-a.e. ω,

$$f(T\omega, X_1^\omega, X_2^\omega) + g(T\omega, X_1^\omega) - g(T^2\omega, X_2^\omega) + c(T\omega) \in t\mathbb{Z} \quad \mathbb{P}^\omega\text{-a.s.}$$

By the T-invariance of m, we may replace ω by $T^{n-1}\omega$ and obtain that for m-a.e. ω,

$$f(T^n\omega, X_1^{T^{n-1}\omega}, X_2^{T^{n-1}\omega}) + g(T^n\omega, X_1^{T^{n-1}\omega}) - g(T^{n+1}\omega, X_2^{T^{n-1}\omega}) + c(T^n\omega) \in t\mathbb{Z} \text{ a.s.}$$

By (S), $(X_1^{T^{n-1}\omega}, X_2^{T^{n-1}\omega})$ is equal in distribution to $(X_n^\omega, X_{n+1}^\omega)$. So

$$f(T^n\omega, X_n^\omega, X_{n+1}^\omega) + g(T^n\omega, X_n^\omega) - g(T^{n+1}\omega, X_{n+1}^\omega) + c(T^n\omega) \in t\mathbb{Z} \text{ a.s. for a.e. } \omega.$$

Let $\mathsf{g}^\omega := \{g(T^n\omega, \cdot)\}_{n\geq 1}$ and $\mathsf{c}^\omega := \{c(T^n\omega)\}_{n\geq 1}$, then $\mathsf{f}^\omega - \nabla \mathsf{g}^\omega + \mathsf{c}^\omega$ is a reduction of f^ω to an additive functional with algebraic range inside $t\mathbb{Z}$, a contradiction.

Proof of (⇐): Suppose f^Ω is not relatively cohomologous to a coset $t\mathbb{Z}$ for any $t \neq 0$. Necessarily f^Ω is not relatively cohomologous to a constant, and by Theorems 9.10 and 9.17, $V_N^\omega \to \infty$ for a.e. ω.

Assume by way of contradiction that $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) \neq \mathbb{R}$ on a set of positive measure of ω. By Proposition 9.21, $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = G_{ess}$ a.e., where $G_{ess} = \{0\}$ or $\frac{2\pi}{t}\mathbb{Z}$ with $t \neq 0$. The first possibility cannot happen, because it implies that f^ω is center-tight, so by Theorem 3.8, $V_N^\omega = O(1)$, whereas $V_N^\omega \to \infty$ a.e. So there exists $t \neq 0$ such that $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = (2\pi/t)\mathbb{Z}$ a.e., and $H_\omega := H(\mathsf{X}^\omega, \mathsf{f}^\omega) = t\mathbb{Z}$ a.e.

Fix ω such that $H_\omega = t\mathbb{Z}$. By the reduction lemma, there are measurable functions $g_n^\omega(x)$, $h_n^\omega(x, y)$ with $\sum \mathrm{Var}^\omega[h_n^\omega] < \infty$, and there are constants c_n^ω, such that

$$\exp\left[it(f_n^\omega(x, y) - g_n^\omega(x) + g_{n+1}^\omega(y) + h_n^\omega(x, y) - c_n^\omega)\right] = 1$$

almost surely with respect to the distribution of $(X_n^\omega, X_{n+1}^\omega)$. Let $\lambda_n^\omega = e^{itc_n^\omega}$, and $a_n^\omega(x) = e^{itg_n^\omega(x)}$. Then for a.e. ω,

$$e^{it(f(T^n\omega, x, y) + h_n^\omega(x, y))} = \lambda_n^\omega a_n^\omega(x)/a_{n+1}^\omega(y) \text{ a.e. w.r.t } \mu_{T^n\omega}(dx)\pi(T^n\omega, x, dy).$$

This seems close to a contradiction to the assumption that f^Ω is not relatively cohomologous to a coset, but we are not quite there yet. Firstly, our proof of the reduction lemma does not provide g_n^ω and c_n^ω of the form $c_n^\omega = c(T^n\omega)$, $a_n^\omega = a(T^n\omega, x)$ with $c(\cdot), a(\cdot, \cdot)$ measurable. Secondly, we need to get rid of h_n^ω.

To resolve these issues we look closer at the structure of the hexagon spaces for MCRE (see §2.3.1). For a.e. ω, $H_\omega = t\mathbb{Z}$ so $\sum d(T^n\omega, t)^2 < \infty$ μ-almost everywhere. By Lemma 9.23, this can only happen if $d(\omega, t) := d(\mathsf{f}^\omega, t) = 0$ a.e. Hence

$$\Gamma(P) \in (2\pi/t)\mathbb{Z} \text{ for a.e. hexagon } P \in \mathrm{Hex}(\omega), \text{ for } m\text{-a.e. } \omega. \tag{9.14}$$

(Γ is the balance, defined in §2.3.2.)

Let $\underline{L}_n^\omega = (Z_{n-2}^\omega, Y_{n-1}^\omega, X_n^\omega)$ denote the ladder process of X^ω (see §2.3.3), and $H^\omega(\underline{L}_n^\omega, \underline{L}_{n+1}^\omega) := \Gamma\left(Z_{n-2}^\omega, \begin{smallmatrix} Z_{n-1}^\omega \\ Y_{n-1}^\omega \end{smallmatrix}, \begin{smallmatrix} Y_n^\omega \\ X_n^\omega \end{smallmatrix}, X_{n+1}^\omega\right)$.

It is a property of the ladder process that the hexagon in the RHS is distributed exactly like a random hexagon in Hex (ω). So

$$H^\omega \in (2\pi/t)\,\mathbb{Z} \quad \mathbb{P}^\omega\text{-a.s., } m\text{-a.e.} \tag{9.15}$$

Next we define the *octagon* balance

$$\Gamma\left(Z_1^\omega, \begin{matrix} Z_2^\omega \\ Y_2^\omega \end{matrix}, \begin{matrix} Z_3^\omega \\ X_3^\omega \end{matrix}, \begin{matrix} Y_4^\omega \\ X_4^\omega \end{matrix}, X_5^\omega\right) \overset{!}{:=} H^\omega(\underline{L}_3^\omega, \underline{L}_4^\omega) + H^\omega(\underline{L}_4^\omega, \underline{L}_5^\omega). \tag{9.16}$$

(The definition requires clarification, because the right-hand-side seems to depend through $\underline{L}_4^\omega$ also on Y_3^ω. In fact, there is no such dependence: The octagon is obtained by stacking $\left(Z_2^\omega, \begin{smallmatrix} Z_3^\omega \\ Y_3^\omega \end{smallmatrix}, \begin{smallmatrix} Y_4^\omega \\ X_4^\omega \end{smallmatrix}, X_5^\omega\right)$ on top of $\left(Z_1^\omega, \begin{smallmatrix} Z_2^\omega \\ Y_2^\omega \end{smallmatrix}, \begin{smallmatrix} Y_3^\omega \\ X_3^\omega \end{smallmatrix}, X_4^\omega\right)$ and removing the common edge $\underline{L}_4^\omega = (Z_2^\omega, Y_3^\omega, X_4^\omega)$. When we add the balances of these hexagons, this edge appears twice with opposite signs, and cancels out.)

Claim 1. *Let $\mathbb{P}^\omega$ denote the distribution of $\{\underline{L}_n^\omega\}$. For each $\zeta^* \in \mathfrak{S}$, there is a measurable function $\widetilde{\zeta}(\omega) \in \mathfrak{S}$ such that for a.e. ω,* $\Gamma\left(\zeta^*, \begin{matrix} \widetilde{\zeta}(\omega) \\ Y_2^\omega \end{matrix}, \begin{matrix} \zeta^* \\ X_3^\omega \end{matrix}, \begin{matrix} Y_4^\omega \\ X_4^\omega \end{matrix}, X_5^\omega\right) \in \frac{2\pi}{t}\mathbb{Z} \qquad \mathbb{P}^\omega\left(\cdot \;\middle|\; \begin{matrix} Z_3^\omega = \zeta^* \\ Z_2^\omega = \widetilde{\zeta}(\omega) \\ Z_1^\omega = \zeta^* \end{matrix}\right)$–a.e.

Remark. This is the only point in the proof where we need conditions (G), (C), (D).

Proof of the Claim. By (9.15) and (9.16), $\Gamma \in \frac{2\pi}{t}\mathbb{Z}$ with full $\mathbb{P}^\omega$–probability, for a.e. ω. The point it to obtain this a.s. with respect to the conditional measures.

By the assumptions of the proposition, at least one of (C) and (D) is true. Assume (D). Then $\mathfrak{S}$ is countable or finite, and for fixed ω, the $\mathbb{P}^\omega$-distribution of $(\underline{L}_3^\omega, \underline{L}_4^\omega, \underline{L}_5^\omega)$ is purely atomic.

Necessarily, $\Gamma \in \frac{2\pi}{t}\mathbb{Z}$ for every octagon with positive $\mathbb{P}^\omega$–probability. So the claim holds for any pair $(\zeta^*, \widetilde{\zeta}) \in \mathfrak{S}$ such that $\mathbb{P}^\omega\big[(Z_1^\omega, Z_2^\omega, Z_3^\omega) = (\zeta^*, \widetilde{\zeta}(\omega), \zeta^*)\big] > 0$. For the ladder process, $\{Z_i^\omega\}$ is equal in distribution to $\{X_i^\omega\}$, therefore such pairs exist by assumptions (E) and (G). Since $\mathfrak{S}$ is countable there is no problem to choose $\widetilde{\zeta}(\omega)$ measurably, and the claim is proved, under assumption (D).

Now suppose (D) fails. Then (C) must hold. There is no loss of generality in assuming that m, the measure on the noise space Ω, is globally supported. Otherwise we replace Ω by $\mathrm{supp}(m)$.

By (9.15), (9.16) and Fubini's Theorem, for a.e. $\omega \in \Omega$, for a.e. $(\zeta_1, \zeta_2, \zeta_3)$ with respect to the distribution $(\zeta_1, \zeta_2, \zeta_3) \sim (Z_1^\omega, Z_2^\omega, Z_3^\omega)$,

$$\mathbb{E}_{\mathbb{P}^\omega}\left(\left|e^{it\Gamma\left(Z_1^\omega, \begin{smallmatrix} Z_2^\omega \\ Y_2^\omega \end{smallmatrix}, \begin{smallmatrix} Z_3^\omega \\ X_3^\omega \end{smallmatrix}, \begin{smallmatrix} Y_4^\omega \\ X_4^\omega \end{smallmatrix}, X_5^\omega\right)} - 1\right|^2 \middle| \begin{smallmatrix} Z_1^\omega = \zeta_1 \\ Z_2^\omega = \zeta_2 \\ Z_3^\omega = \zeta_3 \end{smallmatrix}\right) = 0. \tag{9.17}$$

Let $\mathbb{P}'$ denote the (annealed) joint distribution of $(\omega, X_1^\omega, X_2^\omega, X_3^\omega)$. By the Markov property, (G), (E) and (C), the LHS has a continuous $\mathbb{P}'$-version on

$$A = \{(\omega, \zeta_1, \zeta_2, \zeta_3) \in \Omega \times \mathfrak{S}^3 : p(T\omega, \zeta_1, \zeta_2)p(T^2\omega, \zeta_2, \zeta_3) > 0\}.$$

Henceforth we replace the LHS of (9.17) by this continuous version.

By (C1) and (C2), A is open, and by (G) and the assumption that $\mathrm{supp}(m) = \Omega$, $A \subset \mathrm{supp}(\mathbb{P}')$. So every open subset of A has positive $\mathbb{P}'$-measure, and (9.17) holds on a dense subset of A, whence *everywhere* in A.

To prove the claim it remains to construct a measurable function $\widetilde{\zeta}(\omega)$ such that $(\omega, \zeta^*, \widetilde{\zeta}(\omega), \zeta^*) \in A$ for all ω. By (E) and (G) $\int_{\mathfrak{S}} p(T\omega, \zeta^*, \zeta)p(T^2\omega, \zeta, \zeta^*)\mu_{T^2\omega}(d\zeta)$ is strictly positive, so for every ω there is $\widetilde{\zeta}$ such that

$$\psi(\omega, \widetilde{\zeta}) := p(T\omega, \zeta^*, \widetilde{\zeta})p(T^2\omega, \widetilde{\zeta}, \zeta^*) > 0.$$

We now apply Lemma 9.26, and deduce Claim 1.

Henceforth we fix some $\zeta^* \in \mathfrak{S}$ and will apply the claim for that particular point.

Given $\omega \in \Omega$ and $a, b \in \mathfrak{S}$, construct the bridge distribution $\mathbb{P}^\omega_{ab}(E) = \mathbb{P}^\omega(Y_2^\omega \in E | Z_1^\omega = a, X_3^\omega = b)$ as in §2.2.3.

Claim 2. *For a.e. ω, for a.e. (ξ_3, ξ_4, ξ_5) sampled from the joint distribution of $(X_3^\omega, X_4^\omega, X_5^\omega)$, the random variables*

$$\begin{aligned} W_3^\omega &:= f(T\omega, \zeta^*, Y_2) + f(T^2\omega, Y_2, \xi_3), \quad Y_2 \sim \mathbb{P}^\omega_{\zeta^*, \xi_3} \\ W_5^{T^2\omega} &:= f(T^3\omega, \zeta^*, Y_4) + f(T^4\omega, Y_4, \xi_5), \quad Y_4 \sim \mathbb{P}^{T^2\omega}_{\zeta^*, \xi_5} \end{aligned}$$

are purely atomic, and there are $c_i = c_i(\omega, \xi_i)$ such that $W_3^\omega \in c_3 + \frac{2\pi}{t}\mathbb{Z}$ and $W_5^\omega \in c_5 + \frac{2\pi}{t}\mathbb{Z}$ with full probability.

Proof of the Claim. By the choice of $\widetilde{\zeta}(\omega)$ and Fubini's theorem, for a.e. $(\xi_3, \xi_4, \xi_5) \sim (X_3^\omega, X_4^\omega, X_5^\omega)$,

$$\Gamma\left(\zeta^*, \begin{matrix}\widetilde{\zeta}(\omega) \\ Y_2^\omega\end{matrix}, \begin{matrix}\zeta^* \\ \xi_3\end{matrix}, \begin{matrix}Y_4^\omega \\ \xi_4\end{matrix}, \xi_5\right) \in \frac{2\pi}{t}\mathbb{Z} \qquad \mathbb{P}^\omega\left(\cdot \left| \begin{matrix} Z_3^\omega = \zeta^* & X_3^\omega = \xi_3 \\ Z_2^\omega = \widetilde{\zeta}(\omega) & X_4^\omega = \xi_4 \\ Z_1^\omega = \zeta^* & X_5^\omega = \xi_5 \end{matrix}\right.\right)\text{–a.e.}$$

Notice that $\Gamma\left(\zeta^*, \begin{matrix}\widetilde{\zeta}(\omega) \\ Y_2^\omega\end{matrix}, \begin{matrix}\zeta^* \\ \xi_3\end{matrix}, \begin{matrix}Y_4^\omega \\ \xi_4\end{matrix}, \xi_5\right)$ is equal to the *independent* difference of $W_5^{T^2\omega}$ and W_3^ω, plus a constant which only depends on $(\omega, \xi_3, \xi_4, \xi_5)$. Now Lemma 9.25 gives the claim, except that the lemma gives c_i depending on both ξ_3, ξ_4 and ξ_5. However for fixed ω and ξ_3 the distribution of W_3^ω is independent of ξ_4 and ξ_5, whence c_3 is a function of ξ_3 only. Likewise, c_5 depends only on ξ_5.

Claim 3. *Given ω and (ξ_3, ξ_4, ξ_5) as in Claim 2, let*

$$\begin{aligned} g(\omega, \xi_3) &:= \begin{pmatrix} \text{the smallest positive atom of } W_3^\omega \text{ if there are positive atoms,} \\ \text{otherwise, the largest non-positive atom of } W_3^\omega \end{pmatrix} \\ c(\omega) &:= -f(T\omega, \zeta^*, \widetilde{\zeta}(\omega)) - f(T^2\omega, \widetilde{\zeta}(\omega), \zeta^*). \end{aligned}$$

These functions are well-defined, measurable, and for m-a.e. ω, for a.e. $(\xi_3, \xi_4, \xi_5) \sim (X_3^\omega, X_4^\omega, X_5^\omega)$,

$$[f(T^3\omega, \xi_3, \xi_4) + f(T^4\omega, \xi_4, \xi_5)] + g(\omega, \xi_3) - g(T^2\omega, \xi_5) + c(\omega) \in \frac{2\pi}{t}\mathbb{Z}. \tag{9.18}$$

Proof of the Claim. The function $g(\omega, \xi_3)$ is well-defined for a.e. ω because of claim 2. It is measurable, because $(\omega, \xi_3) \mapsto \mathbb{P}(W_3^\omega \in (a, b))$ are measurable, and

$$[g(\omega, \xi_3) > a] = \begin{cases} \{(\omega, \xi_3) : \mathbb{P}^\omega(0 < W_3^\omega \leq a) = 0\,, \mathbb{P}^\omega(W_3^\omega > a) \neq 0\} & (a > 0) \\ \{(\omega, \xi_3) : \mathbb{P}(W_3^\omega > a) \neq 0\} & (a \leq 0). \end{cases}$$

The measurability of $c(\omega)$ is clear.

The LHS of equation (9.18) is an atom of

$$-\Gamma\left(\zeta^*, \begin{matrix}\widetilde{\zeta}(\omega) \\ Y_2^\omega\end{matrix}, \begin{matrix}\zeta^* \\ \xi_3\end{matrix}, \begin{matrix}Y_4^\omega \\ \xi_4\end{matrix}, \xi_5\right), \quad (\underline{L}_3^\omega, \underline{L}_4^\omega) \sim \mathbb{P}^\omega\left(\cdot \left| \begin{matrix} Z_3^\omega = \zeta^* & X_3^\omega = \xi_3 \\ Z_2^\omega = \widetilde{\zeta}(\omega) & X_4^\omega = \xi_4 \\ Z_1^\omega = \zeta^* & X_5^\omega = \xi_5 \end{matrix}\right.\right).$$

By Claim 1 and Fubini's theorem, it takes values in $\frac{2\pi}{t}\mathbb{Z}$ for a.e. $(\omega, \xi_3, \xi_4, \xi_5)$ distributed like the annealed distribution of $(\omega, X_3^\omega, X_4^\omega, X_5^\omega)$.

By (S), $\{X_i^{T^3\omega}\}_{i \geq 3}$ is equal in distribution to $\{X_i^{T^n\omega}\}_{i \geq n}$, so (9.18) gives a bounded measurable function $\alpha(\omega, x)$ such that for all n,

$$f_n^\omega(X_n^\omega, X_{n+1}^\omega) + f_{n+1}^\omega(X_{n+1}^\omega, X_{n+2}^\omega) + \alpha(T^n\omega, X_n^\omega) - \alpha(T^{n+2}\omega, X_{n+2}^\omega) + c(T^n\omega) \in \frac{2\pi}{t}\mathbb{Z}.$$

Fix ω, and let

$$f_n^* := f(T^n\omega, X_n^\omega, X_{n+1}^\omega) + \alpha(T^n\omega, X_n^\omega) - \alpha(T^{n+1}\omega, X_{n+1}^\omega).$$

Then $f_n^* + f_{n+1}^* + c(T^n\omega) \in \frac{2\pi}{t}\mathbb{Z}$ $\mathbb{P}^\omega$-a.s. So

$$e^{itf_n^*} = e^{-itc(T^n\omega)}e^{-itf_{n+1}^*} = e^{it[c(T^{n+1}\omega - c(T^n\omega)]}e^{itf_{n+2}^*}.$$

By induction, there are constants $\lambda_n = \lambda_n(\omega)$ such that $e^{itf_1^*} = \lambda_{2n}e^{itf_{2n}^*}$. Then

$$\begin{aligned} e^{itf_1^*(X_1^\omega, X_2^\omega)} &= \lambda_{2n}\mathbb{E}(e^{itf_{2n}^*(X_{2n}^\omega, X_{2n+1}^\omega)}|X_1^\omega, X_2^\omega) = \lambda_{2n}\mathbb{E}(e^{itf_{2n}^*(X_{2n}^\omega, X_{2n+1}^\omega)}|X_2^\omega) \\ &= \lambda_{2n}\left[\mathbb{E}(e^{itf_{2n}^*}) + O(\theta^n)\right], \text{ where } 0<\theta<1, \text{ see (2.11).} \end{aligned}$$

Choose $n_k \to \infty$ so that $\lambda_{2n_k}\mathbb{E}(e^{itf_{2n_k}^*}) \to \lambda(\omega)$. Necessarily $|\lambda(\omega)| = 1$ and

$$e^{itf_1^*} = \lambda \quad \mathbb{P}^\omega \text{a.s.}$$

This argument works for a.e. ω. Since the left-hand-side is measurable in ω, $\lambda(\omega)$ equals a.e. to a measurable function. Without loss of generality $\lambda(\omega)$ is measurable. Recalling the definition of f_n^* and setting $a(\omega, x) := \exp(-it\alpha(\omega, x))$, we obtain

$$e^{itf(T\omega, X_1^\omega, X_2^\omega)} = \lambda(\omega)\frac{a(T\omega, X_1^\omega)}{a(T^2\omega, X_2^\omega)}.$$

Thus f^Ω is relatively cohomologous to $t\mathbb{Z}$, in contradiction to our assumptions. This contradiction shows that $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{R}$ for a.e. ω, and proves part 1 of the proposition.

To prove part 2 ($\Leftarrow$), we assume that f^Ω is integer valued, but not relatively cohomologous to a coset of $n\mathbb{Z}$ with $n > 1$, and show that $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{Z}$ a.s. Equivalently, we must show that $H_\omega := H(\mathsf{X}^\omega, \mathsf{f}^\omega) = 2\pi\mathbb{Z}$ for a.e. ω. Since f^Ω is integer valued, $2\pi \in H_\omega$, so if $H_\omega \neq 2\pi\mathbb{Z}$, then $H_\omega = t\mathbb{Z}$ for $t = \frac{2\pi}{n}$ and $n \in \{2, 3, 4, \ldots\}$. We can now repeat the proof of part 1 verbatim, and obtain a relative cohomology to a coset of $n\mathbb{Z}$, a contradiction to our assumptions. The proof of the implication ($\Rightarrow$) is similar to that in the non-lattice case, and we omit it. □

Proof of Theorem 9.11. In the non-lattice case, f^Ω is not relatively cohomologous to a coset of $t\mathbb{Z}$ with $t \neq 0$. In particular, f^Ω is not relatively cohomologous to a constant, and by Theorem 9.10, $\exists\sigma^2 > 0$ such that $V_N^\omega \sim N\sigma^2$.

By Proposition 9.24 (1), for a.e. ω, $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{R}$ and f^ω is irreducible. The non-lattice LLT in Theorem 9.11 now follows from Theorems 5.1.

The lattice case has a similar proof, except that now we use Proposition 9.24 (2) to check irreducibility. □

Proof of Theorem 9.18 The proof is identical to the proof of Theorem 9.11, except that now, since the noise process is infinite, we do not know that $V_N^\omega \sim \text{const.}N$, we only know that $V_N^\omega \to \infty$ (Theorem 9.17). □

9.3.5 LLT for Large Deviations

We prove Theorems 9.12, 9.14 and 9.19.

Proof of Theorem 9.12. Let $\mathcal{G}_N^\omega(\xi) := \frac{1}{N}\log\mathbb{E}^\omega(e^{\xi S_N^\omega})$.

Part (1): We show that for a.e. ω, $\mathcal{G}_N^\omega(\xi)$ converges pointwise on $\mathbb{R}$ to a continuously differentiable and strictly convex function $\mathcal{F}(\xi)$, which does not depend on ω.

Step 1: *Given $\xi \in \mathbb{R}$, for every $\omega \in \Omega$ and for all $n \geq 1$, there are unique numbers $\overline{p}_n(\xi, \omega) \in \mathbb{R}$ and unique non-negative functions $\overline{h}_n(\cdot, \xi, \omega) \in L^\infty(\mathfrak{S}, \mathscr{B}(\mathfrak{S}), \mu_{T^n\omega})$ such that $\int_{\mathfrak{S}} \overline{h}_n(x, \xi, \omega)\mu_{T^n\omega}(dx) = 1$, and*

$$\int_{\mathfrak{S}} \mathrm{e}^{\xi f(T^n\omega,x,y)} \frac{\overline{h}_{n+1}(y,\xi,\omega)}{\mathrm{e}^{\overline{p}_n(\xi,\omega)}\overline{h}_n(x,\xi,\omega)} \pi(T^n\omega, x, \mathrm{d}y) = 1. \tag{9.19}$$

Furthermore, there is a measurable function $\overline{p}(\xi,\omega)$ *such that* $\overline{p}_n(\xi,\omega) = \overline{p}(\xi, T^n\omega)$.

Proof of the Step. The existence and uniqueness of $\overline{h}_n, \overline{p}_n$ follow from Lemma 7.13, applied to $(\mathsf{X}^\omega, \mathsf{f}^\omega)$ with $a_n = 0$. Writing (9.19) first for (n,ω) and then for $(n-1, T\omega)$, and then invoking uniqueness, we find that $\overline{p}_n(\xi,\omega) = \overline{p}_{n-1}(\xi, T\omega)$. So

$$\overline{p}_n(\xi,\omega) = \overline{p}_{n-1}(\xi,T\omega) = \cdots = \overline{p}_1(\xi, T^{n-1}\omega) =: \overline{p}(\xi, T^n\omega).$$

The proof of Lemma 7.13 represents $h_n(x,\xi,\omega)$ as a limit of expressions which are measurable in (x,ξ,ω), so $(x,\xi,\omega) \mapsto \overline{h}_n(x,\xi,\omega)$ is measurable. By (9.19), $(\omega,\xi) \mapsto \overline{p}(\xi,\omega)$ is measurable.

Step 2: *Let* $K := \operatorname{ess\,sup} |\mathsf{f}|$ *and let* ϵ_0 *be the ellipticity constant from (E). For every* $R > 0$ *there exists a constant* $C(\epsilon_0, K, R)$ *such that* $|\overline{p}(\xi,\omega)| \le C(\epsilon_0, K, R)$ *for all* $\omega \in \Omega$ *and* $|\xi| \le R$.

Proof of the Step. See the proof of Lemma 7.14.

Step 3: *Let* $\overline{P}_N(\xi,\omega) := \sum_{k=1}^{N} \overline{p}(\xi, T^k\omega)$*, then for every* $\omega \in \Omega$ *such that* $V_N^\omega \to \infty$, $\mathcal{G}_N^\omega(\xi) = (V_N^\omega/N)[\overline{P}_N(\xi,\omega)/V_N^\omega + O(1/V_N^\omega)]$ *uniformly on compact sets of* ξ *in* $\mathbb{R}$.

Proof of the Step. It is convenient to work with

$$\mathcal{F}_N^\omega(\xi) := \frac{1}{V_N^\omega} \log \mathbb{E}^\omega(\mathrm{e}^{\xi S_N^\omega}) \equiv (N/V_N^\omega)\mathcal{G}_N^\omega(\xi).$$

Let $P_N(\xi,\omega) := \overline{P}_N(\xi,\omega) + (\mathbb{E}^\omega(S_N^\omega) - \frac{\mathrm{d}}{\mathrm{d}\eta}|_{\eta=0}\overline{P}_N(\eta,\omega))\xi$. By Lemmas 7.16–7.18,

(1) $\frac{\mathrm{d}}{\mathrm{d}\eta}|_{\eta=0} P_N(\eta,\omega)$ exists,
(2) $|P_N(\xi,\omega) - \overline{P}_N(\xi,\omega)| = O(1)$ uniformly on compact sets of ξ,
(3) $|\mathcal{F}_N^\omega(\xi) - P_N(\xi)/V_N^\omega| = O(1/V_N^\omega)$ uniformly on compact sets of ξ.

Step 3 follows.

We can now prove the a.s. convergence of $\mathcal{G}_N^\omega(\xi)$. By the assumptions of the theorem, f^Ω is not relatively cohomologous to a constant. Therefore, by Theorem 9.10, there exists $\sigma^2 > 0$ such that $V_N^\omega \sim \sigma^2 N$ as $N \to \infty$ for a.e. ω.

Fix a countable dense set $\{\xi_1, \xi_2, \ldots\} \subset \mathbb{R}$. For each i, $\omega \mapsto \overline{p}(\xi_i,\omega)$ is bounded and measurable. By Fact 3, for a.e. ω,

$$\lim_{N\to\infty} \mathcal{G}_N^\omega(\xi_i) = \sigma^2 \lim_{N\to\infty} \frac{1}{V_N^\omega} \sum_{k=1}^{N} \overline{p}(\xi_i, T^k\omega) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \overline{p}(\xi_i, T^k\omega)$$
$$= \int_\Omega \overline{p}(\xi_i,\omega) m(\mathrm{d}\omega), \text{ by the pointwise ergodic theorem.}$$

This shows that for all i there exists $\mathcal{G}(\xi_i) \in \mathbb{R}$ such that $\lim_{N\to\infty} \mathcal{G}_N^\omega(\xi_i) = \mathcal{G}(\xi_i)$ for a.e. ω, *with* $\mathcal{G}(\xi_i)$ *independent of* ω. Let Ω' denote the set of full measure of ω where this holds for all $i \in \mathbb{N}$.

If $K := \operatorname{ess\,sup} |\mathsf{f}^\Omega|$, then $|(\mathcal{G}_N^\omega)'(\xi)| \le \left| \frac{|\xi| \mathbb{E}^\omega(|S_N^\omega| \mathrm{e}^{\xi S_N^\omega})}{N\mathbb{E}^\omega(\mathrm{e}^{\xi S_N^\omega})} \right| \le \frac{|\xi| K N}{N} = K|\xi|$. Therefore, the functions $\xi \mapsto \mathcal{G}_N^\omega(\xi)$ are equicontinuous on compacts.

If a sequence of functions on $\mathbb{R}$ is equicontinuous on compacts and converges on a dense subset of $\mathbb{R}$, then it converges on all of $\mathbb{R}$ to a continuous limit. So there is a continuous function $\mathcal{F}^\omega(\xi)$ such that

$$\lim_{N\to\infty} \mathcal{G}_N^\omega(\xi) = \mathcal{F}^\omega(\xi) \text{ for all } \xi \in \mathbb{R} \text{ and } \omega \in \Omega'.$$

In fact $\mathcal{F}^\omega(\xi)$ does not depend on ω, because by virtue of continuity,

$$\mathcal{F}^{\omega}(\xi) = \lim_{k\to\infty} \mathcal{F}^{\omega}(\xi_{i_k}) = \lim_{k\to\infty} \mathcal{G}(\xi_{i_k}), \text{ whenever } \xi_{i_k} \xrightarrow[k\to\infty]{} \xi_i,$$

and the RHS is independent of ω. We are therefore free to write $\mathcal{F}^{\omega}(\xi) = \mathcal{F}(\xi)$.

It remains to show that $\mathcal{F}(\xi)$ is continuously differentiable and strictly convex on $\mathbb{R}$. Fix $\omega \in \Omega'$. Applying Theorem 7.3 to $(\mathsf{X}^{\omega}, \mathsf{f}^{\omega})$ we find that for every $R > 0$ there is a $C = C(R)$ such that $C^{-1} \le (\mathcal{F}_N^{\omega})'' \le C$ on $[-R, R]$. So $\mathcal{F}$ is continuously differentiable and strictly convex on $(-R, R)$, because of the following fact:

Lemma 9.27 *Suppose $\varphi_n : \mathbb{R} \to \mathbb{R}$ are twice differentiable convex functions such that $C^{-1} \le \varphi_n'' \le C$ with $C > 0$, on $(-R, R)$. If $\varphi_n \xrightarrow[N\to\infty]{} \varphi$ pointwise on $(-R, R)$, then φ is continuously differentiable and strictly convex on $(-R, R)$.*

Proof Recall that a pointwise limit of convex functions is convex, and convex functions have one-sided derivatives $\varphi_{\pm}'(\xi) := \lim_{h\to 0^{\pm}} \frac{\varphi(\xi+h)-\varphi(\xi)}{h}$.
Differentiability: For all $|\xi| < R$,

$$|\varphi_+'(\xi) - \varphi_-'(\xi)| = \lim_{h\to 0^+} \left| \frac{\varphi(\xi+h)-\varphi(\xi)}{h} - \frac{\varphi(\xi-h)-\varphi(\xi)}{h} \right|$$

$$\begin{aligned}
&= \lim_{h\to 0^+} \lim_{n\to\infty} \left| \frac{\varphi_n(\xi+h)-\varphi_n(\xi)}{h} - \frac{\varphi_n(\xi-h)-\varphi_n(\xi)}{h} \right| \\
&= \lim_{h\to 0^+} \lim_{n\to\infty} |\varphi_n'(\xi_n) - \varphi_n'(\eta_n)| \text{ for some } \xi_n, \eta_n \in (\xi - h, \xi + h) \\
&\le \lim_{h\to 0^+} \lim_{n\to\infty} 2Ch = 0, \text{ because } |\varphi_n''| \le C \text{ on a neighborhood of } \xi.
\end{aligned}$$

We find that $\varphi_+'(\xi) = \varphi_-'(\xi)$, whence φ is differentiable at ξ.
Strict Convexity: Suppose $-R < \xi < \eta < R$, then

$$\begin{aligned}
&\varphi'(\eta) - \varphi'(\xi) = \varphi_+'(\eta) - \varphi_-'(\xi) = \lim_{h\to 0^+} \frac{\varphi(\eta+h)-\varphi(\eta)}{h} - \frac{\varphi(\xi-h)-\varphi(\xi)}{h} \\
&= \lim_{h\to 0^+} \lim_{n\to\infty} \frac{\varphi_n(\eta+h)-\varphi_n(\eta)}{h} - \frac{\varphi_n(\xi-h)-\varphi_n(\xi)}{h} \\
&= \lim_{h\to 0^+} \lim_{n\to\infty} \varphi_n'(\eta_n) - \varphi_n'(\xi_n) \text{ for some } \xi_n \in [\xi - h, \xi],\ \eta_n \in [\eta, \eta + h] \\
&\ge \lim_{h\to 0^+} \liminf_{n\to\infty} C^{-1}|\eta_n - \xi_n| = C^{-1}(\eta - \xi), \text{ because } \varphi_n'' > C^{-1} \text{ on } (-R, R).
\end{aligned}$$

It follows that φ' is strictly increasing on $(-R, R)$, whence the strict convexity of φ.
The Derivative is Lipschitz Continuous: The same calculation as before shows that if $-R < \xi < \eta < R$, then $|\varphi'(\eta) - \varphi'(\xi)| \le C|\xi - \eta|$. □

Part (2): We show that for a.e. ω, the Legendre transforms of $\mathcal{G}_N^{\omega}$ converge to the Legendre transform of $\mathcal{F}$. Again, the proof is based on general properties of convex functions.

Lemma 9.28 *Suppose $\varphi_n(\xi), \varphi(\xi)$ are finite, convex, and differentiable on $(-R, R)$. If $\varphi_n(\xi) \xrightarrow[n\to\infty]{} \varphi(\xi)$ on $(-R, R)$, then $\varphi_n'(\xi) \xrightarrow[n\to\infty]{} \varphi'(\xi)$ on $(-R, R)$.*

Proof Fix $\xi \in (-R, R)$. By convexity, for every $h > 0$ sufficiently small,

$$\frac{\varphi_n(\xi)-\varphi_n(\xi-h)}{h} \le \varphi_n'(\xi) \le \frac{\varphi_n(\xi+h)-\varphi_n(\xi)}{h}. \tag{9.20}$$

This is because the LHS is at most $(\varphi_n)_-'(\xi)$, the RHS is at least $(\varphi_n)_+'(\xi)$, and by differentiability, $(\varphi_n)_{\pm}'(\xi) = \varphi_n'(\xi)$. Passing to the limit $n \to \infty$, we find that

$$\limsup \varphi_n'(\xi), \liminf \varphi_n'(\xi) \in \left[\frac{\varphi(\xi)-\varphi(\xi-h)}{h}, \frac{\varphi(\xi+h)-\varphi(\xi)}{h}\right] \text{ for all } h>0.$$

We now invoke the differentiability of φ, pass to the limit $h \to 0^+$, and discover that $\limsup \varphi_n'(\xi)$ and $\liminf \varphi_n'(\xi)$ are both equal to $\varphi'(\xi)$. □

Lemma 9.29 *Let $\varphi_n(\xi), \varphi(\xi)$ be finite, strictly convex, C^1 functions on $\mathbb{R}$, s.t. $\varphi_n(\xi) \to \varphi(\xi)$ for all $\xi \in \mathbb{R}$. Let $\varphi'(\pm\infty) := \lim_{\xi\to\pm\infty} \varphi'(\xi)$. Let φ_n^*, φ^* denote the Legendre transforms of φ_n, φ. Then for all $\eta \in (\varphi'(-\infty), \varphi'(+\infty))$, for n sufficiently large, φ_n^* is well-defined on a neighborhood of η, and $\varphi_n^*(\eta) \to \varphi^*(\eta)$.*

Proof Fix $\eta \in (\varphi'(-\infty), \varphi'(+\infty))$. By assumption, φ' is continuous and strictly increasing. Therefore, there exists some ξ such that $\varphi'(\xi) = \eta$.

Fix two constants $\varepsilon_0, M_0 > 0$ such that $|\varphi'| \le M_0$ on $[\xi-\varepsilon_0, \xi+\varepsilon_0]$, and choose $0 < \varepsilon < \varepsilon_0$ arbitrarily small. Choose $\xi_1 < \xi < \xi_2$ such that $|\xi_1 - \xi_2| < \varepsilon$. Then $\varphi'(\xi_1) < \eta < \varphi'(\xi_2)$. Choose $\delta > 0$ such that $\varphi'(\xi_1) < \eta - \delta < \eta + \delta < \varphi'(\xi_2)$. By Lemma 9.28, $\varphi_n'(\xi_i) \to \varphi'(\xi_i)$, and therefore there exists N such that for all $n > N$,

$$-M_0 - 1 < \varphi'(\xi_1) - 1 < \varphi_n'(\xi_1) < \eta - \delta < \eta + \delta < \varphi_n'(\xi_2) < \varphi'(\xi_2) + 1 < M_0 + 1. \quad (9.21)$$

$I_\eta := (\eta-\delta, \eta+\delta) \subset (\varphi_n'(\xi_1), \varphi_n'(\xi_2))$ and φ_n' is continuous and strictly increasing. Therefore, for every $\eta' \in I_\eta$ there exists a unique $\xi_n' \in (\xi_1, \xi_2)$ so that $\varphi_n'(\xi_n') = \eta'$. So φ_n^* is well-defined on I_η.

Let ξ_n be the solution to $\varphi_n'(\xi_n) = \eta$. Then $\varphi_n^*(\eta) = \xi_n\eta - \varphi_n(\xi_n)$. Similarly, $\varphi^*(\eta) = \xi\eta - \varphi(\xi)$. So

$$\begin{aligned}
&|\varphi_n^*(\eta) - \varphi^*(\eta)| \le |\xi_n - \xi| \cdot |\eta| + |\varphi_n(\xi_n) - \varphi(\xi)| \\
&\le |\xi_1 - \xi_2| \cdot |\eta| + |\varphi_n(\xi_n) - \varphi_n(\xi)| + |\varphi_n(\xi) - \varphi(\xi)| \\
&\overset{!}{\le} \varepsilon|\eta| + (M_0+1)|\xi_n - \xi| + |\varphi_n(\xi) - \varphi(\xi)|, \ \because |\varphi_n'| \le M_0 + 1 \text{ on } (\xi_1, \xi_2) \text{ by (9.21)} \\
&\le \varepsilon(M_0 + 1 + |\eta|) + o(1), \text{ as } n \to \infty.
\end{aligned}$$

Since ε is arbitrary, $\varphi_n^*(\eta) \to \varphi^*(\eta)$. □

Let $\mathcal{I}_N^\omega$ be the Legendre transforms of $\mathcal{G}_N^\omega$, and let $\mathcal{I}$ be the Legendre transform of $\mathcal{F}$. By the last lemma, for a.e. ω, $\mathcal{I}_N^\omega$ is eventually defined on each compact subset of $(-\mathcal{F}'(-\infty), \mathcal{F}(\infty))$, and converges there to $\mathcal{I}$.

Part (3): We analyze the function $\mathcal{I}(\eta)$. Fix ω such that $\varphi_N(\xi) = \frac{1}{N}\log \mathbb{E}^\omega(e^{\xi S_N^\omega})$ converges pointwise to $\mathcal{F}$. By Lemma 9.29, φ_N^* converges pointwise to $\mathcal{I}$. Since φ_N'' is uniformly bounded away from zero and infinity on compacts (see the first part of the proof), $(\varphi_N^*)''$ is uniformly bounded away from zero and infinity on compacts, see Lemma 7.23. By Lemma 9.27, $\mathcal{I} = \lim \varphi_N^*$ is strictly convex and continuously differentiable.

By Lemma 9.28, $(\varphi_N^*)'(\eta) \xrightarrow[N\to\infty]{} \mathcal{I}'(\eta)$ for all η in the interior of the range of φ', and $\varphi_N'(\xi) \xrightarrow[N\to\infty]{} \mathcal{F}'(\xi)$ for all $\xi \in \mathbb{R}$. The convergence is uniform on compacts, because $(\varphi_N^*)'', \varphi_N''$ are bounded on compacts.

It is easy to verify that φ_N is twice differentiable. Therefore by Lemma 7.23, φ_N^* is twice differentiable and $(\varphi_N^*)'(\varphi_N'(\xi)) = \xi$ for all ξ. Passing to the limit as $N \to \infty$ we obtain the identity

$$\mathcal{I}'(\mathcal{F}'(\xi)) = \xi \text{ for all } \xi \in \mathbb{R}. \quad (9.22)$$

$\mathcal{F}'$ is strictly increasing, because it is the derivative of a strictly convex function. Since $\mathcal{I}'(\mathcal{F}'(\xi)) = \xi$, $\mathcal{I}'(\eta) \xrightarrow[\eta\to\mathcal{F}'(\pm\infty)]{} \pm\infty$. By convexity, $\mathcal{I}$ is continuous where it is finite, and therefore has compact level sets.

Substituting $\xi = 0$ in (9.22), we obtain $\mathcal{I}'(\mathcal{F}'(0)) = 0$, so $\eta = \mathcal{F}'(0)$ is a critical point of $\mathcal{I}(\cdot)$. By strict convexity, $\mathcal{I}$ attains its global minimum at $\mathcal{F}'(0)$. In addition,

$$\mathcal{I}(\mathcal{F}'(0)) = 0 \cdot \mathcal{F}'(0) - \mathcal{F}(0) = 0.$$

We conclude that $\mathcal{I}(\eta) = 0$ when $\eta = \mathcal{F}'(0)$, and $\mathcal{I}(\eta) > 0$ for $\eta \ne \mathcal{F}'(0)$.

It remains to see that $\mathcal{F}'(0)$ is equal to the asymptotic mean μ. In part 1 we saw that for a.e. ω, $\mathcal{F}(\xi) = \lim(\mathcal{G}_N^\omega)(\xi)$ on $\mathbb{R}$, and $\mathcal{G}_N^\omega$ and $\mathcal{F}$ are differentiable. By Lemma 9.28, for a.e. ω, $\mathcal{F}'(0) = \lim(\mathcal{G}_N^\omega)'(0) = \lim \mathbb{E}^\omega(S_N^\omega)/N = \mu$.

Part (4): We calculate the large deviations thresholds.

Without loss of generality, $\sigma^2 = 1$, otherwise we work with $\widetilde{\mathsf{f}}^\Omega := \mathsf{f}^\Omega/\sigma$ and notice that the thresholds, asymptotic mean, and the asymptotic log-moment generating function associated to $\widetilde{\mathsf{f}}^\Omega$ are related to those of f by $\widetilde{\mathfrak{c}}_\pm = \sigma\mathfrak{c}_\pm$, $\widetilde{\mathfrak{r}}_\pm = \sigma\mathfrak{r}_\pm$, $\widetilde{\mu} = \mu/\sigma$, and $\widetilde{\mathcal{F}}(\xi) = \mathcal{F}(\xi/\sigma)$. Since $\sigma^2 = 1$, for a.e. ω, $V_N^\omega \sim N$, and for all $\xi \in \mathbb{R}$,

$$\mathcal{F}(\xi) = \lim_{N\to\infty} \mathcal{F}_N^\omega(\xi), \text{ where } \mathcal{F}_N^\omega(\xi) = \frac{1}{V_N^\omega} \log \mathbb{E}^\omega(e^{\xi S_N^\omega}).$$

Step 1. *For a.e. ω, $\mathfrak{c}_+(\mathsf{X}^\omega, \mathsf{f}^\omega) \geq \mathcal{F}'(+\infty) - \mu$ and $\mathfrak{c}_-(\mathsf{X}^\omega, \mathsf{f}^\omega) \leq \mathcal{F}'(-\infty) - \mu$.*

Proof of the Step. Fix $\eta \in (\mathcal{F}'(-\infty) - \mu, \mathcal{F}'(+\infty) - \mu)$, and choose $\eta^\pm$ such that

$$\mathcal{F}'(-\infty) - \mu < \eta^- < \eta < \eta^+ < \mathcal{F}'(+\infty) - \mu.$$

Take $\xi^\pm$ such that $\mathcal{F}'(\xi^\pm) - \mu = \eta^\pm$. By Lemma 9.28 and the definition of μ,

$$\lim_{N\to\infty} (\mathcal{F}_N^\omega)'(\xi^\pm) - \frac{\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega} = \eta^\pm \text{ a.s.}$$

In particular, if $\frac{z_N - \mathbb{E}(S_N^\omega)}{V_N^\omega} \to \eta$, then for all large N,

$$\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{V_N^\omega} \in \left[(\mathcal{F}_N^\omega)'(\xi^-) - \frac{\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}, (\mathcal{F}_N^\omega)'(\xi^+) - \frac{\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}\right],$$

and $\{z_N\}$ is admissible. So for a.e. ω, every $\eta \in (\mathcal{F}'(-\infty) - \mu, \mathcal{F}'(+\infty) - \mu)$ is reachable (recall that by our assumption $\lim_{N\to\infty} V_N/N = 1$). Since $\mathfrak{c}_+$ is the supremum of reachable points, $\mathfrak{c}_+ \geq \mathcal{F}'(+\infty) - \mu$. Similarly, one shows that $\mathfrak{c}_- \leq \mathcal{F}'(-\infty) - \mu$.

Step 2. *For a.e. ω, $\mathfrak{c}_+(\mathsf{X}^\omega, \mathsf{f}^\omega) \leq \mathcal{F}'(+\infty) - \mu$ and $\mathfrak{c}_-(\mathsf{X}^\omega, \mathsf{f}^\omega) \geq \mathcal{F}'(-\infty) - \mu$.*

Proof of the Step. Fix ω such that $V_N^\omega/N \to \sigma^2{=}1$, $\mathbb{E}(S_N^\omega)/N \to \mu$, and $(\mathcal{F}_N^\omega)' \to \mathcal{F}'$ on $\mathbb{R}$ (a.e. ω is like that). Take $\eta > \mathcal{F}'(\infty) - \mu$, and assume by way of contradiction that η is reachable. Let z_N be an admissible sequence such that $\frac{z_N - \mathbb{E}(S_N^\omega)}{V_N^\omega} \to \eta$.

By admissibility, for some R, for all N large,

$$\frac{z_N - \mathbb{E}^\omega(S_N^\omega)}{V_N^\omega} \in \left[(\mathcal{F}_N^\omega)'(-R) - \frac{\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}, (\mathcal{F}_N^\omega)'(R) - \frac{\mathbb{E}^\omega(S_N^\omega)}{V_N^\omega}\right].$$

Necessarily, $\eta \leq \mathcal{F}'(R) - \mu$, whence by convexity, $\eta \leq \mathcal{F}'(\infty) - \mu$. But this contradicts the choice of η.

Step 3. $\mathfrak{r}_+(\omega) := \lim_{N\to\infty} \dfrac{\text{ess sup}[S_N^\omega - \mathbb{E}(S_N^\omega)]}{V_N^\omega}$, $\mathfrak{r}_-(\omega) := \lim_{N\to\infty} \dfrac{\text{ess inf}[S_N^\omega - \mathbb{E}(S_N^\omega)]}{V_N^\omega}$ *exist a.e., and are a.e. constant.*

Proof of the Step. It is enough to prove the statement for $\mathfrak{r}_+$; the statement for $\mathfrak{r}_-$ follows by considering $-\mathsf{f}^\Omega$.

Let $\mathcal{S}_N(\omega) = \text{ess sup}\, S_N^\omega - \mathbb{E}^\omega(S_N^\omega)$. By (S), $\{X_i^{T^N\omega}\}_{i\geq 1}$ is equal in distribution to $\{X_i^\omega\}_{i\geq N+1}$. So $\mathbb{E}^\omega(S_{N+M}^\omega) = \mathbb{E}^\omega(S_N^\omega) + \mathbb{E}^{T^N\omega}(S_M^{T^N\omega})$. It is not difficult to see using (E) and (B) that

$$\text{ess sup}\, S_{N+M}^\omega \leq \text{ess sup}\, S_N^\omega + \text{ess sup}\, S_M^{T^N\omega} - 4K.$$

Thus the sequence $\mathcal{T}_N(\omega) = \mathcal{S}_N(\omega) - 4K$ is sub-additive, with respect to the noise process $(\Omega, \mathscr{F}, m, T)$. Since $\mathcal{S}_N(\omega) \geq -KN$, the subadditive ergodic theorem implies that the limit

$$\lim_{N\to\infty} \frac{\mathcal{S}_N(\omega)}{N} = \lim_{N\to\infty} \frac{\mathcal{T}_N(\omega)}{N}$$

exists, and is independent of ω with probability one. The step follows, since

$$V_N^\omega/N \to \sigma^2 = 1 \text{ a.s.}$$

Step 4. $\mathfrak{c}_+ = \mathfrak{r}_+$ *and* $\mathfrak{c}_- = \mathfrak{r}_-$.

Proof of the Step. By §7.4, $\mathfrak{c}_+ \le \mathfrak{r}_+$ and $\mathfrak{c}_- \ge \mathfrak{r}_-$, so it is enough to show that $\mathfrak{c}_+ \ge \mathfrak{r}_+$ and $\mathfrak{c}_- \le \mathfrak{r}_-$.

Fix $\varepsilon > 0$. By Step 3, for each sufficiently large N_0, there exists $\gamma'_{\varepsilon,N_0} > 0$ and a set $\Omega'_{\varepsilon,N_0}$ with measure bigger than $1 - \frac{1}{3}\varepsilon^2$, such that all $\omega \in \Omega'_{\varepsilon,N_0}$,

$$\mathbb{P}^\omega\left(S_{N_0}^\omega - \mathbb{E}(S_{N_0}^\omega) \ge (\mathfrak{r}_+ - \frac{1}{2}\varepsilon)V_{N_0}^\omega + 8K\right) \ge \gamma'_{\varepsilon,N_0}. \tag{9.23}$$

By (E), (B) and (S), for all sufficiently large N_0, we can find $\gamma''_{\varepsilon,N_0} > 0$ and $\Omega''_{\varepsilon,N_0}$ with measure bigger than $1 - \frac{1}{2}\varepsilon^2$ such that for all $\omega \in \Omega''_{\varepsilon,N_0}$,

$$\mathbb{P}^\omega\left(S_{N_0}^\omega - \mathbb{E}(S_{N_0}^\omega) \ge (\mathfrak{r}_+ - \frac{2}{3}\varepsilon)V_{N_0}^\omega \middle| X_1^\omega, X_{N_0+1}^\omega\right) \ge \gamma''_{\varepsilon,N_0}.$$

Since $V_N^\omega/N \to \sigma^2 = 1$, by choosing N_0 sufficiently large, we can find $\gamma_{\varepsilon,N_0} > 0$ and Ω_{ε,N_0} with measure bigger than $1 - \varepsilon^2$ such that for all $\omega \in \Omega_{\varepsilon,N_0}$,

$$\mathbb{P}^\omega\left(S_{N_0}^\omega - \mathbb{E}(S_{N_0}^\omega) \ge (\mathfrak{r}_+ - \varepsilon)N_0 \;\middle|\; X_1^\omega, X_{N_0+1}^\omega\right) \ge \gamma_{\varepsilon,N_0}, \tag{9.24}$$

and $V_N^\omega \ge N/2$ for $N \ge N_0$.

Given M, let $j_1(\omega) < j_2(\omega) < \cdots < j_{n_M(\omega)}(\omega)$ be all the times $1 \le j < M$ when $T^{jN_0}(\omega) \in \Omega_{\varepsilon,N_0}$. Then

$$\mathbb{P}^\omega\left(S_{N_0M}^\omega - \mathbb{E}^\omega(S_{N_0M}^\omega) \ge n_M(\mathfrak{r}_+ - \varepsilon)N_0 - (M - n_M)N_0K\right) \overset{!}{\ge} \gamma_{\varepsilon,N_0}^{n_M} \ge \gamma_{\varepsilon,N_0}^M.$$

To see this, condition on $X_1^\omega, X_{N_0+1}^\omega, \ldots, X_{MN_0+1}^\omega$ to make the partial sums of terms involving $X_{\ell N_0+1}^\omega, \ldots, X_{(\ell+1)N_0+1}^\omega$ independent for different ℓ; then use (9.24) or (B) to control the partial sums; finally take the expectation over $X_1^\omega, X_{N_0+1}^\omega, \ldots, X_{MN_0+1}^\omega$.

The pointwise ergodic theorem for $T^{N_0} : \Omega \to \Omega$ says that there is a T^{N_0}-invariant function $\beta(\omega)$ such that for a.e. ω,

$$\beta(\omega) = \lim_{M\to\infty} \frac{n_M(\omega)}{M} \quad \text{and} \quad \int \beta(\omega)\mathrm{d}m = m(\Omega_{\varepsilon,N_0}) > 1 - \varepsilon^2$$

($\beta(\omega)$ is not necessarily constant, because T^{N_0} is not necessarily ergodic). Clearly $\beta(\omega) \le 1$. Therefore $m[\beta > 1 - \varepsilon] > 1 - \varepsilon$. So for large M, and on a set $\overline{\Omega}_\varepsilon$ of measure bigger than $1 - \varepsilon$, $\dfrac{n_M}{M} > 1 - \varepsilon$. On $\overline{\Omega}_\varepsilon$,

$$n_M(\mathfrak{r}_+ - \varepsilon)N_0 - (M - n_M)N_0K \ge [(1 - \varepsilon)(\mathfrak{r}_+ - \varepsilon) - \varepsilon K]\, N_0M.$$

Thus, for all $\varepsilon > 0$ small, on a set of ω with probability at least $1 - \varepsilon - \varepsilon^2$, there is an $\eta_\varepsilon > 0$ and two constants $C_1, C_2 > 0$ independent of ε such that for all N large,

$$\mathbb{P}^\omega[S_N^\omega \ge \mathbb{E}^\omega(S_N^\omega) + (\mathfrak{r}_+ - C_1\varepsilon)N + C_2\varepsilon N] \ge \eta_\varepsilon^{V_N^\omega}. \tag{9.25}$$

(In (9.25) we used the inequality $V_N^\omega > N/2$, which is valid on Ω_{ε,N_0}.)

Next we claim that for all $\varepsilon > 0$ small, on a set of ω with probability at least $1 - \varepsilon$, there is a $\theta_\varepsilon > 0$ such that for all N large,

$$\mathbb{P}^\omega[S_N^\omega \le \mathbb{E}^\omega(S_N^\omega) + (\mathfrak{r}_+ - C_1\varepsilon)N - C_2\varepsilon N] \ge \theta_\varepsilon^{V_N^\omega}. \quad (9.26)$$

Indeed, By Theorem 7.26, $\mathfrak{r}_+ \ge \mathfrak{c}_+ > 0$. Choose $\varepsilon > 0$ such that $(C_1 + C_2)\varepsilon < \mathfrak{r}_+$. Then $\mathbb{P}^\omega[S_N^\omega \le \mathbb{E}^\omega(S_N^\omega) + (\mathfrak{r}_+ - C_1\varepsilon)N - C_2\varepsilon N] \ge \mathbb{P}^\omega[S_N^\omega \le \mathbb{E}^\omega(S_N^\omega)] \to 1/2$, by Dobrushin's CLT.

By (9.25) and (9.26), $z_N{:=}\mathbb{E}^\omega(S_N^\omega) + (\mathfrak{r}_+ - C_1\varepsilon)N$ is admissible for ω in a set of measure at least $1 - 2\varepsilon$, and therefore $\mathfrak{c}_+ \ge (1 - O(\varepsilon))\mathfrak{r}_+$ on a set of ω with probability bigger than $1 - 2\varepsilon$. Taking $\varepsilon \to 0$, we obtain $\mathfrak{c}_+ \ge \mathfrak{r}_+$ as required.

By symmetry, $\mathfrak{c}_- \le \mathfrak{r}_-$. □

Proof of Theorems 9.14 and 9.19. By Proposition 9.24, $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{R}$ a.e. in case (1), and $G_{ess}(\mathsf{X}^\omega, \mathsf{f}^\omega) = \mathbb{Z}$ a.e. in case (2). Parts (1) and (2) of Theorem 9.14 then follow from Theorem 7.26, using $V_N^\omega \sim N\sigma^2$, $\mathbb{E}(S_N^\omega) \sim N\mu$. Part (3) follows from Theorem 7.4(4). Theorem 9.19 has a similar proof, using Theorem 7.8. □

9.4 Notes and References

Markov chains in random environment were introduced by Cogburn [27]. Probabilistic limit theorems for MCRE are given in Cogburn [28], Seppäläinen [180], Kifer [111], [112] and Hafouta & Kifer [93, Ch. 6,7],[92].

In dynamical systems, one studies a setup similar to MCRE, called a "random dynamical system." In this setup, one iterates a map T_ω with ω varying randomly from iterate to iterate. For a fixed realization of noise, a random dynamical system reduces to a "sequential" (aka "time-dependent" or "non-autonomous") dynamical system. Limit theorems for random dynamical systems can be found in Kifer [112] Conze, Le Borgne & Roger [30], Denker & Gordin [42], Aimino, Nicol & Vaienti [9], Nicol, Török & Vaienti [151], and Dragičević, Froyland & González-Tokman [64] (this is a partial list). For limit theorems for sequential dynamical systems, see Bakhtin [12], Conze & Raugi [31], Haydn, Nicol, Török & Vaienti [97], Korepanov, Kosloff & Melbourne [120], and Hafouta [90, 91].

If we set the noise process to be the identity on the one point space, then the LLT in this chapter reduce to LLT for homogeneous stationary Markov chains, see Chapter 8 and references therein.

The results of this chapter are all essentially known in the case when T preserves a finite measure. Theorem 9.10 was proved in the more general setup of random dynamical systems by Kifer [112],[110]. Corollary 9.13 and the first two parts of Lemma 9.8 are close to results in [110],[111], and the third part is due to Hafouta (private communication). Theorem 9.11 is close to the results of Dragičević, Froyland & González-Tokman [64], and Hafouta & Kifer [93, chapter 7, Theorem 7.1.5]. The main difference is in the irreducibility assumptions. Our condition of not being relatively cohomologous to a coset is replaced in [93] by what these authors call the "lattice" and "non-lattice" cases (this is not the same as our terminology). In the paper [64], the non-cohomology condition is replaced by a condition on the decay of the norms of certain Nagaev perturbation operators, and a connection to a non-cohomology condition is made under additional assumptions.

The results for infinite measure noise processes seem to be new. The reason we can also treat this case, is that the LLT we provide in this work do not require any assumptions on the rate of growth of V_N, and they also work when it grows sub-linearly. It would be interesting to obtain similar results for more general stochastic processes (or deterministic systems) in random environment with infinite invariant measure.

Appendix A
The Gärtner-Ellis Theorem in One Dimension

A.1 The Statement

The **Legendre-Fenchel transform** of a convex function $\varphi : \mathbb{R} \to \mathbb{R}$ is the function $\varphi^* : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ given by

$$\varphi^*(\eta) := \sup_{\xi \in \mathbb{R}} \{\xi\eta - \varphi(\xi)\}.$$

This is closely related to the Legendre transform, defined by (7.40), see Lemma A.3 below. Our purpose is to show the following special case of the Gärtner-Ellis theorem:

Theorem A.1 *Suppose $a_n \to \infty$, and let W_n be a sequence of random variables such that $\mathbb{E}(e^{\xi W_n}) < \infty$ for all $\xi \in \mathbb{R}$. Assume that the limit*

$$\mathcal{F}(\xi) := \lim_{n\to\infty} \frac{1}{a_n} \log \mathbb{E}(e^{\xi W_n})$$

exists for all $\xi \in \mathbb{R}$, and is differentiable and strictly convex on $\mathbb{R}$. Let $\mathcal{I}(\eta)$ be the Legendre-Fenchel transform of $\mathcal{F}(\xi)$. Then:

(1) For every closed set $F \subset \mathbb{R}$, $\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in F] \leq -\inf_{\eta \in F} I(\eta)$.

(2) For every open set $G \subset \mathbb{R}$, $\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in G] \geq -\inf_{\eta \in G} I(\eta)$.

A.2 Background from Convex Analysis

To prove Theorem A.1 we need to recall some standard facts from convex analysis.

Lemma A.1 *Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function which is differentiable on $\mathbb{R}$. Then φ is continuously differentiable on $\mathbb{R}$.*

Proof By convexity, for every $h > 0$, $\frac{\varphi(z)-\varphi(z-h)}{h} \leq \varphi'(z) \leq \frac{\varphi(z+h)-\varphi(z)}{h}$. So $\limsup_{y\to x} \varphi'(y) - \varphi'(x) \leq \limsup_{y\to x} \left[\frac{\varphi(y+h)-\varphi(y)}{h} - \frac{\varphi(x)-\varphi(x-h)}{h}\right] = \frac{\varphi(x+h)-\varphi(x)}{h} - \frac{\varphi(x)-\varphi(x-h)}{h} \xrightarrow[h\to 0]{} \varphi'(x) - \varphi'(x) = 0$. Similarly, $\liminf_{y\to x} \varphi'(y) - \varphi'(x) \geq 0$. Thus $\lim_{y\to x} \varphi'(y) = \varphi'(x)$. □

The derivative of a differentiable convex function is monotone increasing. For such functions we can safely define $\varphi'(\infty) := \lim_{\xi\to\infty} \varphi'(\xi), \quad \varphi'(-\infty) := \lim_{\xi\to-\infty} \varphi'(\xi)$.

Lemma A.2 *If $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly convex and differentiable on $\mathbb{R}$, then the Legendre transform ψ of φ is continuously differentiable and strictly convex on $(\varphi'(-\infty), \varphi'(\infty))$. In addition, $\psi' \circ \varphi' = id$ there.*

Proof By strict convexity, φ' is strictly increasing. By the previous lemma, φ' is continuous. So the Legendre transform ψ of φ is well-defined on $(\varphi'(-\infty), \varphi'(\infty))$. Fix $h \neq 0$ and η so that $\eta, \eta + h \in (\varphi'(-\infty), \varphi'(\infty))$. Then $\exists! \xi, \xi_h$ such that $\varphi'(\xi) = \eta, \quad \psi(\eta) = \xi\eta - \varphi(\xi), \quad \varphi'(\xi_h) = \eta + h, \quad \psi(\eta + h) = \xi_h(\eta + h) - \varphi(\xi_h)$.

The following identities hold:

$$\frac{\psi(\eta + h) - \psi(\eta)}{h} = \frac{[\xi_h(\eta + h) - \varphi(\xi_h)] - [\xi\eta - \varphi(\xi)]}{h}$$

$$= \frac{(\xi_h - \xi)\eta + \varphi(\xi) - \varphi(\xi_h)}{h} + \xi_h = (\xi_h - \xi)\left[\frac{1}{h}\left(\eta - \frac{\varphi(\xi_h) - \varphi(\xi)}{\xi_h - \xi}\right)\right] + \xi_h.$$

By convexity, the term in the square brackets lies between $\frac{1}{h}(\eta - \varphi'(\xi)) = 0$ and $\frac{1}{h}(\eta - \varphi'(\xi_h))$=−1. Therefore it is bounded, and $\dfrac{\psi(\eta+h)-\psi(\eta)}{h} = \xi_h + O(|\xi_h - \xi|)$.

By strict convexity, φ' is increasing, and by the previous lemma, φ' is continuous. It follows that the inverse function of φ' is well-defined and continuous. Consequently, $\xi_h = (\varphi')^{-1}(\eta + h) \xrightarrow[h\to 0]{} (\varphi')^{-1}(\eta) = \xi$. It follows that $\dfrac{\psi(\eta+h)-\psi(\eta)}{h} = \xi_h + O(|\xi_h - \xi|) \xrightarrow[h\to 0]{} \xi = (\varphi')^{-1}(\eta)$.

Thus ψ is differentiable on $(\varphi'(-\infty), \varphi(\infty))$, and $\psi' = (\varphi')^{-1}$ there. Looking at this formula, we recognize that ψ' is continuous and increasing, which shows that ψ is C^1 and strictly convex on $(\varphi'(-\infty), \varphi'(\infty))$. □

Lemma A.3 *Let $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ be a finite, differentiable, and strictly convex function. Let $\mathcal{I}$ denote the Legendre-Fenchel transform of $\mathcal{F}$. Then:*

(1) $\mathcal{I}$ is convex on $\mathbb{R}$. (2) $\mathcal{I} = +\infty$ outside of $[\mathcal{F}'(-\infty), \mathcal{F}'(\infty)]$.

(3) $\mathcal{I}$ agrees with the Legendre transform of $\mathcal{F}$ on $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$. On this interval $\mathcal{I}$ is strictly convex, differentiable, and $\mathcal{I}' \circ \mathcal{F}' = id$.

(4) $\mathcal{I}$ is increasing to the right of $\mathcal{F}'(0)$ and decreasing to the left of $\mathcal{F}'(0)$.

Proof The first statement follows from the sub-additivity of the supremum.

To see the second statement, let $\varphi_\eta(\xi) := \xi\eta - \mathcal{F}(\xi)$.

- If $\eta > \mathcal{F}'(\infty)$, then $\varphi'_\eta(\xi) \xrightarrow[\xi\to\infty]{} \eta - \mathcal{F}'(\infty) > 0$, so $\varphi_\eta(\xi) \xrightarrow[\xi\to\infty]{} +\infty$.
- If $\eta < \mathcal{F}'(-\infty)$, then $\varphi'_\eta(\xi) \xrightarrow[\xi\to-\infty]{} \eta - \mathcal{F}'(-\infty) < 0$, and $\varphi_\eta(\xi) \xrightarrow[\xi\to-\infty]{} +\infty$.

In both cases, $\mathcal{I}(\eta) = \sup_\xi \varphi_\eta(\xi) = +\infty$.

Now suppose $\eta \in (\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$. By Lemma A.1, $\mathcal{F}'$ is continuous, and by strict convexity, $\mathcal{F}'$ is strictly increasing. So there is exactly one ξ_0, where $\mathcal{F}'(\xi_0) = \eta$. As $\varphi_\eta(\xi)$ is concave, this is the point where $\varphi_\eta(\xi)$ attains its global maximum, and we find that $\mathcal{I}(\eta) = \varphi_\eta(\xi_0) = \xi_0\eta - \mathcal{F}(\xi_0)$. It follows that $\mathcal{I}$ agrees with the Legendre transform of $\mathcal{F}$ at η. The remaining parts of part 3 follows from Lemma A.2.

By part 3, $\mathcal{I}'(\mathcal{F}'(0)) = 0$, therefore $\mathcal{I}$ attains its global minimum at $\mathcal{F}'(0)$, and is decreasing on $(\mathcal{F}'(-\infty), \mathcal{F}'(0))$ and increasing on $(\mathcal{F}'(0), \mathcal{F}'(\infty))$. At this point it is already clear that $\mathcal{I}$ satisfies the conclusion of part 4 on $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$. Since $\mathcal{I}$ is finite on $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$ and equal to $+\infty$ outside $[\mathcal{F}'(-\infty), \mathcal{F}'(\infty)]$, we just need to check that the values of $\mathcal{I}$ at $\mathcal{F}'(\pm\infty)$ do not spoil the monotonicity. Indeed they do not. For example, for every $\mathcal{F}'(0) \le \xi < \mathcal{F}'(\infty)$,

$$\mathcal{I}(\mathcal{F}'(\infty)) = \lim_{t\to 0^+}(1-t)\mathcal{I}(\mathcal{F}'(\infty)) + t\mathcal{I}(\mathcal{F}'(0)) \ge \lim_{t\to 0^+}\mathcal{I}\Big((1-t)\mathcal{F}'(\infty) + t\mathcal{F}'(0)\Big) \ge \mathcal{I}(\xi),$$

where the first inequality holds because $\mathcal{I}$ is convex on $\mathbb{R}$ and the second holds because $\mathcal{I}$ is increasing on $[\mathcal{F}'(0), \mathcal{F}'(\infty))$.

Similarly, for every $\mathcal{F}'(-\infty) \le \xi < \mathcal{F}'(0)$, we have $\mathcal{I}(\mathcal{F}'(-\infty)) \ge \mathcal{I}(\xi)$. □

Lemma A.4 *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function with Legendre-Fenchel transform φ^*, and suppose $\varphi^*(\eta_0) = \xi_0\eta_0 - \varphi(\xi_0)$. Then the Legendre-Fenchel transform of $\varphi(\xi + \xi_0) - \varphi(\xi_0)$ is $\varphi^*(\eta) - \varphi^*(\eta_0) + \xi_0(\eta_0 - \eta)$.*

Proof At η, this Legendre transform is equal to

$$\begin{aligned}&\sup_\xi\left[\xi\eta - \varphi(\xi+\xi_0) + \varphi(\xi_0)\right] = \sup_\xi\left[(\xi+\xi_0)\eta - \varphi(\xi+\xi_0)\right] + \varphi(\xi_0) - \xi_0\eta\\&= \varphi^*(\eta) - \left[\xi_0\eta - \varphi(\xi_0)\right] = \varphi^*(\eta) - \left[\xi_0\eta_0 - \varphi(\xi_0)\right] + \xi_0(\eta_0 - \eta).\end{aligned}$$

The lemma follows from the identity $\varphi^*(\eta_0) = \xi_0\eta_0 - \varphi(\xi_0)$. □

A.3 Proof of the Gärtner-Ellis Theorem

Proof of the Upper Bound: Suppose $\eta \geq \mathcal{F}'(0)$. For every $\xi \geq 0$,

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \geq \eta] \leq \lim_{n\to\infty} \frac{1}{a_n} \log \mathbb{E}(\mathrm{e}^{\xi(W_n - a_n\eta)}) = \mathcal{F}(\xi) - \xi\eta.$$

This is also true for $\xi < 0$, because for such ξ, since $\mathcal{F}(0) = 0$,

$$\mathcal{F}(\xi) - \xi\eta = |\xi| \left(\eta - \frac{\mathcal{F}(\xi) - \mathcal{F}(0)}{\xi}\right) \geq |\xi|(\eta - \mathcal{F}'(0)) \geq 0 \geq \limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \geq \eta].$$

In summary, $\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \geq \eta] \leq \mathcal{F}(\xi) - \xi\eta$ for all $\xi \in \mathbb{R}$. Passing to the infimum on ξ, we obtain that $\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \geq \eta] \leq -\mathcal{I}(\eta)$ for all $\eta \geq \mathcal{F}'(0)$. Similarly, one shows that

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \leq \eta] \leq -\mathcal{I}(\eta) \text{ for all } \eta \leq \mathcal{F}'(0).$$

Every closed set $F \subset \mathbb{R}$ can be covered by at most two sets of the form $(-\infty, \eta_1]$ and $[\eta_2, \infty)$ where $\eta_i \in F$ and $\eta_1 \leq \mathcal{F}'(0) \leq \eta_2$. By Lemma A.3(4), $\inf_F \mathcal{I} = \min\{\mathcal{I}(\eta_i)\}$.

So $\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in F] \overset{!}{\leq} -\inf_F \mathcal{I}$ ($\because \log(A + B) \leq 2 + \max(\log A, \log B)$).

Proof of the Lower Bound: We begin with the special case $G = (\alpha, \beta)$, where $\mathcal{F}'(0) \leq \alpha < \beta \leq \mathcal{F}'(\infty)$. Fix $0 < \delta < \frac{\beta-\alpha}{2}$ arbitrarily close to zero.

Define ξ_δ, η_δ by

$$\eta_\delta := \alpha + \delta, \quad \mathcal{F}'(\xi_\delta) = \eta_\delta.$$

Since $\mathcal{I}' \circ \mathcal{F}' = id$ on $(\mathcal{F}'(-\infty), \mathcal{F}'(\infty))$, $\xi_\delta = \mathcal{I}'(\eta_\delta)$, and since $\mathcal{I}$ is increasing to the right of $\mathcal{F}'(0)$, $\xi_\delta > 0$.

Let $\mu_n(\mathrm{d}t)$ denote the probability measures on $\mathbb{R}$ given by $\mu_n(E) := \mathbb{P}[W_n \in E]$. Construct the change of measure

$$\widetilde{\mu}_n(\mathrm{d}t) = \frac{\mathrm{e}^{\xi_\delta t} \mu_n(\mathrm{d}t)}{\mathbb{E}(\mathrm{e}^{\xi_\delta W_n})},$$

and let $\widetilde{W}_n$ denote the random variables such that $\mathbb{P}[\widetilde{W}_n \in E] = \widetilde{\mu}_n(E)$.

Claim. $\mathbb{P}[\widetilde{W}_n/a_n \in (\alpha, \alpha + 2\delta)] \xrightarrow[n\to\infty]{} 1$.

Proof of the Claim. Clearly, $\log \mathbb{E}(\mathrm{e}^{t\widetilde{W}_n}) = \log \mathbb{E}[\mathrm{e}^{(t+\xi_\delta)W_n}/\mathbb{E}(\mathrm{e}^{\xi_\delta W_n})]$. Therefore,

$$\lim_{n\to\infty} \frac{1}{a_n} \log \mathbb{E}(\mathrm{e}^{t\widetilde{W}_n}) = \widetilde{\mathcal{F}}(t) := \mathcal{F}(t + \xi_\delta) - \mathcal{F}(\xi_\delta).$$

By Lemma A.4, the Legendre-Fenchel transform of $\widetilde{\mathcal{F}}$ is

$$\widetilde{\mathcal{I}}(s) = \mathcal{I}(s) - \mathcal{I}(\eta_\delta) + \xi_\delta(\eta_\delta - s).$$

Note that $\widetilde{\mathcal{F}}'(0) = \eta_\delta = \alpha + \delta$. Therefore, by the upper bound we just proved,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[\widetilde{W}_n/a_n \geq \alpha + 2\delta] \leq -\widetilde{\mathcal{I}}(\alpha + 2\delta) \\
&= -\left[\mathcal{I}(\alpha + 2\delta) - \mathcal{I}(\eta_\delta) - \xi_\delta\delta\right] = -\delta \left[\frac{\mathcal{I}(\eta_\delta + \delta) - \mathcal{I}(\eta_\delta)}{\delta} - \mathcal{I}'(\eta_\delta)\right] \overset{!}{<} 0,
\end{aligned}$$

where the last inequality is because $\mathcal{I}$ is strictly convex. It follows that

$$\mathbb{P}[\widetilde{W}_n/a_n \geq \alpha + 2\delta] \to 0.$$

Similarly, working with the random variables $-\widetilde{W}_n$, one shows that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[\widetilde{W}_n/a_n \leq \alpha] \leq -\widetilde{\mathcal{I}}(\alpha) \\
&= -\left[\mathcal{I}(\alpha) - \mathcal{I}(\eta_\delta) + \xi_\delta \delta\right] = -\delta \left[\mathcal{I}'(\eta_\delta) - \frac{\mathcal{I}(\eta_\delta) - \mathcal{I}(\alpha)}{\delta}\right] < 0,
\end{aligned}$$

whence, again, $\mathbb{P}[\widetilde{W}_n/a_n \leq \alpha] \to 0$. The claim follows.

We now return to the problem of bounding $\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in G]$. Since $(\alpha, \beta) \supset (\alpha, \alpha + 2\delta)$ and $\xi_\delta > 0$,

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in (\alpha,\beta)] \geq \liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in (\alpha, \alpha+2\delta)] \\
&\geq \liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{E}\left(1_{(\alpha,\alpha+2\delta)}(W_n/a_n) \frac{e^{\xi_\delta (W_n - (\alpha+2\delta)a_n)}}{\mathbb{E}(e^{\xi_\delta W_n})} \mathbb{E}(e^{\xi_\delta W_n})\right) \\
&= \liminf_{n\to\infty} \frac{1}{a_n} \log \left[\mathbb{E}\left(1_{(\alpha,\alpha+2\delta)}(\widetilde{W}_n/a_n)\right) e^{-\xi_\delta(\alpha+2\delta)a_n} \mathbb{E}(e^{\xi_\delta W_n})\right] \\
&= \lim_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}\left(\widetilde{W}_n/a_n \in (\alpha, \alpha+2\delta)\right) - \xi_\delta(a + 2\delta) + \mathcal{F}(\xi_\delta) \\
&= 0 - (\xi_\delta \eta_\delta - \mathcal{F}(\xi_\delta)) - \xi_\delta \delta = -\mathcal{I}(\eta_\delta) - \xi_\delta \delta \xrightarrow[\delta\to 0]{} -\mathcal{I}(\alpha) \stackrel{!}{=} -\inf_{(\alpha,\beta)} \mathcal{I},
\end{aligned}$$

because $\mathcal{I}$ is increasing on (α, β), by the assumptions on α and β.

Similarly, one shows that whenever $\mathcal{F}'(-\infty) \leq \alpha < \beta \leq \mathcal{F}'(0)$, then

$$\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in (\alpha,\beta)] \geq -\inf_{(\alpha,\beta)} \mathcal{I}.$$

Since $\mathcal{I} = +\infty$ outside $[\mathcal{F}'(-\infty), \mathcal{F}'(\infty)]$, and every open set is a union of intervals, this implies the lower bound for every open set G which does not contain $\mathcal{F}'(0)$.

Now suppose that G does contain $\mathcal{F}'(0)$. Observe that $\inf_G \mathcal{I} = 0$, because zero is the global minimum of $\mathcal{I}$, and by Lemma A.3(3), $\mathcal{I}(\mathcal{F}'(0)) = 0 \times \mathcal{F}'(0) - \mathcal{F}(0) = 0$. Since G is open, $G \supset (\mathcal{F}'(0) - \alpha, \mathcal{F}'(0) + \alpha)$ with α positive, and then by the upper bound and the positivity of $\mathcal{I}(\mathcal{F}'(0) \pm \alpha)$ (which follows from Lemma A.3(4))

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \notin (\mathcal{F}'(0) - \alpha, \mathcal{F}'(0) + \alpha)] < 0.$$

So $\mathbb{P}[W_n/a_n \notin G] \to 0$, and $\lim_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}[W_n/a_n \in G] = 0$. Since, as noted above, $\inf_G \mathcal{I} = 0$, the result follows. □

A.4 Notes and References

Theorem A.1 is a special case of results by Gärtner [77] and Ellis [69], which apply to vector valued random variables, and which assume less on $\mathcal{F}(\xi)$. The special case discussed above follows (up to minor details) from the work of Plachky & Steinbach [158]. The proof we gave here is based on Ellis's book [70, §VII.3]. See [40, §2.3] and [164, §12.2] for additional discussions of Gärtner–Ellis Theorem and its applications.

Appendix B
Hilbert's Projective Metric and Birkhoff's Theorem

B.1 Hilbert's Projective Metric

Suppose $(V, \|\cdot\|)$ is a normed vector space over $\mathbb{R}$. A **cone** is a set $K \subset V$ such that

(1) $K + K \subset K$;
(2) $\lambda K \subset K$ for all $\lambda > 0$;
(3) $K \cap (-K) = \{0\}$;
(4) K is closed, and the interior of K, $\text{int}(K)$, is non-empty.

Necessarily $K = \overline{\text{int}(K)}$: Suppose $x \in K$ and $y \in \text{int}(A)$; Then $x = \lim x_n$, where $x_n := x + \frac{1}{n}y$, and $x_n \in \text{int}(K)$. Indeed, by the assumption on y there is an open ball $B \subset K$ such that $y \in B$, and therefore $x_n \in B' := x + \frac{1}{n}B \subset K + K \subset K$.

Every cone determines a partial order on V by $x \le y \Leftrightarrow y - x \in K$. Sometimes we will write $\le_K$ instead of $\le$.

Note that $x \ge 0 \Leftrightarrow x \in K$, and $x \le y \Rightarrow \lambda x \le \lambda y$ for all $\lambda > 0$.

Two $x, y \in K \setminus \{0\}$ are **comparable** if $my \le x \le My$ for some $M, m > 0$. Let

$$M(x|y) := \inf\{M > 0 : x \le My\};$$
$$m(x|y) := \sup\{m > 0 : my \le x\}.$$

Clearly, $m(x|y) = M(y|x)^{-1}$.

Lemma B.1 *If $x, y \in K \setminus \{0\}$ are comparable, then $M(x|y), m(x|y)$ are finite positive numbers. They are the best constants in the inequality $m(x|y)y \le x \le M(x|y)y$.*

Proof Choose $M_n \downarrow M(x|y)$ such that $x \le M_n y$. So $M_n y - x \in K$ for all n. Passing to the limit and recalling that K is closed, we obtain $M(x|y)y - x \in K$. So $x \le M(x|y)y$. Necessarily, $M(x|y) > 0$: Otherwise $M(x|y) = 0$, and $x \le 0$, whence $-x \in K$. But this is impossible, since $K \cap (-K) = \{0\}$ and $x \ne 0$.

By the symmetry $m(x|y) = M(y|x)^{-1}$, $x \ge m(x|y)y$, and $m(x|y) < \infty$. □

Hilbert's projective metric (of K) is $d_K(x, y) := \log\left(\dfrac{M(x|y)}{m(x|y)}\right)$, $(x, y$ comparable$)$.

Proposition B.2 *Any two $x, y \in \text{int}(K)$ are comparable. Hilbert's projective metric is a pseudo-metric on $\text{int}(K)$, and $d_K(x, y) = 0$ iff x, y are collinear. If x, x' are collinear and y, y' are collinear, then $d_K(x', y') = d_K(x, y)$.*

Proof Let $B(z, r) := \{x \in V : \|x - z\| < r\}$.

Comparability of $x, y \in \text{int}(K)$: Choose $r > 0$ such that $\overline{B(x, r)}, \overline{B(y, r)} \subseteq \text{int}(K)$, then $x - ry/\|y\|, y - rx/\|x\| \in K$, whence $\frac{r}{\|y\|}y \le x \le \frac{\|x\|}{r}y$.

Positivity: Fix two comparable x, y and let $M := M(x|y)$, $m := m(x|y)$. We saw that $my \le x \le My$, so $x - my, My - x \in K$, whence $\left(\frac{M}{m} - 1\right)x = \frac{M}{m}(x - my) + (My - x) \in K$. Necessarily, $M/m \ge 1$, otherwise $\left(1 - \frac{M}{m}\right)x \in K$, and $K \cap (-K) \ni \left(\frac{M}{m} - 1\right)x \ne 0$. So $d_K(x, y) = \log(M/m) \ge 0$. In addition, if $d_K(x, y) = 0$ iff $M = m$. In this case $\pm(My - x) \in K$, so $My = x$ and x, y are collinear.

Symmetry: $d_K(\cdot, \cdot)$ is symmetric, because $M(x|y) = m(y|x)^{-1}$.

Triangle Inequality: If $x, y, z \in \text{int}(K)$, then $x \le M(x|y)y \le M(x|y)M(y|z)z$ and $x \ge m(x|y)y \ge m(x|y)m(y|z)z$. Since $K + K \subset K$, $\le$ is transitive. Therefore $M(x|z) \le M(x|y)M(y|z)$ and $m(x|z) \ge m(x|y)m(y|z)$. The result follows.

Projective Property: If $x' = \lambda x$, then $M(x'|y) = \lambda M(x|y)$ and $m(x'|y) = \lambda m(x|y)$, so $d_K(x', y) = d_K(x, y)$. Similarly, if $y' = \lambda' y$, then $d_K(x', y') = d_K(x', y)$. So $d_K(x', y') = d_K(x, y)$. □

Corollary B.3 *Hilbert's projective metric is a proper metric on the "projectivization" of the interior of K, $\mathbb{P}K := \mathrm{int}(K)/\sim$, where $x \sim y$ iff x, y are collinear.*

Proposition B.4 *If $x \in K \setminus \mathrm{int}(K)$ and $y \in \mathrm{int}(K)$, then x, y are not comparable.*

Proof Suppose by contradiction that x, y are comparable; then $x - my \in K$ for some $m > 0$. Let B denote the open unit ball in V. Since $y \in \mathrm{int}(K)$, $y + \varepsilon B \subset K$ for some $\varepsilon > 0$. Therefore $x + m\varepsilon B = (x - my) + m(y + \varepsilon B) \subset K + mK \subset K$, and $x \in \mathrm{int}(K)$. But this contradicts the assumption on x. □

B.2 Contraction Properties

Let V_i be two normed vector spaces, and let $K_i \subset V_i$ be cones. A linear map $A : V_1 \to V_2$ is called **non-negative (with respect to K_1, K_2)**, if $A(K_1) \subset K_2$, and **positive (with respect to K_1, K_2)** if $A(\mathrm{int}(K_1)) \subset \mathrm{int}(K_2)$. Positivity implies non-negativity, because $K_i = \overline{\mathrm{int}(K_i)}$.

Let $\le_i := \le_{K_i}$. Every non-negative linear map satisfies $Ax \ge_2 0$ on K_1, and

$$x \le_1 y \Rightarrow Ax \le_2 Ay.$$

Proposition B.5 *If A is non-negative, then $d_{K_2}(Ax, Ay) \le_2 d_{K_1}(x, y)$ on $\mathrm{int}(K_1)$.*

Proof If $x, y \in \mathrm{int}(K_1)$ then x, y are comparable, so $m(x|y)y \le_1 x \le_1 M(x|y)y$. Since A is non-negative, $m(x|y)Ay \le_2 Ax \le_2 M(x|y)Ay$, so $M(Ax|Ay) \le M(x|y)$ and $m(Ax|Ay) \ge m(x|y)$. It follows that $d_{K_2}(Ax, Ay) \le d_{K_1}(x, y)$. □

The **projective diameter** of a positive linear map $A : V_1 \to V_2$ (with respect to cones K_1, K_2) is

$$\Delta_{K_1,K_2}(A) := \sup\{d_{K_2}(Ax, Ay) : x, y \in \mathrm{int}(K_1)\} \le \infty.$$

The **hyperbolic tangent function** is $\tanh(t) = \dfrac{e^t - e^{-t}}{e^t + e^{-t}}$. We let $\tanh(\infty) := 1$.

Theorem B.6 (G. Birkhoff) *Suppose K_i are cones in normed vector spaces V_i $(i = 1, 2)$, and let $A : V_1 \to V_2$ be a linear mapping which is positive with respect to K_1, K_2. Then for all $x, y \in \mathrm{int}(K_1)$*

$$d_{K_2}(Ax, Ay) \le \tanh\left(\tfrac{1}{4}\Delta_{K_1,K_2}(A)\right) d_{K_1}(x, y). \tag{B.1}$$

In particular, if $\Delta_{K_1,K_2}(A) < \infty$, then A is a strict contraction.

Proof in a Special Case. Suppose $V_1 = V_2 = V$ and $K_1 = K_2 = K$, where $V = \mathbb{R}^2$, and $K := \{(x, y) : x, y \ge 0\}$.

Let $A : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map such that $A(\mathrm{int}(K)) \subset \mathrm{int}(K)$. The theorem is trivial when $\det(A) = 0$, because in this case Ax, Ay are collinear for all x, y, and $d_K(Ax, Ay) = 0$. Henceforth we assume that $\det(A) \neq 0$.

Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since positive maps are non-negative, $A(K) \subset K$, and this implies that $a, b, c, d \ge 0$ (calculate A on the standard basis).

Two vectors $x = \binom{x_1}{x_2}$, $y = \binom{y_1}{y_2}$ belong to $\mathrm{int}(K)$ iff $x_i, y_i > 0$. In this case $M(x|y) = \max\limits_{i=1,2}(x_i/y_i)$ and $m(x|y) = \min\limits_{i=1,2}(x_i/y_i)$. It follows that

$$d_K(x, y) = \left|\log\left(\frac{x_1}{x_2}\Big/\frac{y_1}{y_2}\right)\right|, \text{ and therefore}$$

$$d_K(Ax, Ay) = \left|\log\left(\frac{ax_1 + bx_2}{cx_1 + dx_2}\Big/\frac{ay_1 + by_2}{cy_1 + dy_2}\right)\right| = \left|\log\frac{ax_1 + bx_2}{cx_1 + dx_2} - \log\frac{ay_1 + by_2}{cy_1 + dy_2}\right|$$

$$= |\log\varphi_A(t) - \log\varphi_A(s)|, \quad \text{where } \varphi_A(\xi) := \frac{a\xi + b}{c\xi + d},\ t := \frac{x_1}{x_2}, \text{ and } s := \frac{y_1}{y_2}.$$

As x, y range over $\mathrm{int}(K)$, t, s range over $(0, \infty)$. Since $\varphi_A'(\xi)=\frac{\det(A)}{(c\xi+d)^2}$ and $\det(A)\neq 0$, φ_A is monotonic, and the image of $\log \varphi_A$ is an interval with endpoints $\varphi_A(0) = \frac{b}{d}$ and $\varphi_A(\infty) = \frac{a}{c}$. It follows that

$$\Delta_{K,K}(A) = \left|\log \frac{ad}{bc}\right|, \tag{B.2}$$

with the understanding that $|\log(\text{zero}/\text{non-zero})| = |\log(\text{non-zero}/\text{zero})| = \infty$. (We do not need to worry about $\log(\text{zero}/\text{zero})$, because $ad - bc \neq 0$.)

By (B.2), $\Delta_{K,K}(A) = \infty$ whenever some of a, b, c, d are zero. In this case $\tanh(\frac{1}{4}\Delta_{K,K}(A)) = 1$, and Birkhoff's Theorem follows from Proposition B.5. Henceforth we assume that $a, b, c, d > 0$.

If $x, y \in \mathrm{int}(K)$ are not collinear, then $d_K(x, y) = \log|\frac{x_1}{x_2}/\frac{y_1}{y_2}| = |\log t - \log s| \neq 0$, with t, s as above. This leads to

$$\frac{d_K(Ax, Ay)}{d_K(x, y)} = \left|\frac{\log \varphi_A(t) - \log \varphi_A(s)}{\log t - \log s}\right|.$$

By Cauchy's mean value theorem, there is some ξ between t and s such that

$$\frac{d_K(Ax, Ay)}{d_K(x, y)} = \left|\frac{(\log \varphi_A)'(\xi)}{(\log)'(\xi)}\right| = \left|\frac{\det(A)\xi}{(a\xi + b)(c\xi + d)}\right| \equiv e^{\psi(\xi)}|\det(A)|, \tag{B.3}$$

where $\psi(\xi) := \log \xi - \log(a\xi + b) - \log(c\xi + d)$. Note that $\xi \in (0, \infty)$.

The task now is to find the global maximum of $e^{\psi(\xi)}|\det(A)|$ on $(0, \infty)$. Since $a, b, c, d \neq 0$, $\psi(\xi) \xrightarrow[\xi\to 0,\infty]{} -\infty$. So ψ has a global maximum inside $(0, \infty)$. Since there is exactly one critical point, $\xi_0 = \sqrt{(bd)/(ac)}$, $\psi(\xi)$ and $e^{\psi(\xi)}|\det(A)|$ attain their global maximum at ξ_0. Substituting ξ_0 in (B.3), we find, after some algebraic work, that the global maximum of the RHS of (B.3) is $\theta(A) := \left|\frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}\right|$. So

$$d_K(Ax, Ay) \leq \theta(A) d_K(x, y) \text{ for all } x, y \in \mathrm{int}(K).$$

(B.2) and some elementary algebra shows that $\theta(A) = \tanh(\Delta_{K,K}(A)/4)$, and Theorem B.6 follows.

Proof for Maps Between General Two-Dimensional Spaces: Suppose V_1, V_2 are two-dimensional.

By finite dimensionality, the topological properties of K_i do not change if we change the norm of V_i. We choose for V_i norms coming from inner products. Euclidean geometry tells us that the intersection of K_i with $S^1 := \{x \in V : \|x\| = 1\}$ is a circle arc $A_i \subset S^1$. Let $\partial A_i := \{\xi_i, \eta_i\}$. If $\xi_1 = \eta_1$ then $\mathrm{int}(K_1) = \varnothing$. If $\xi_2 = \eta_2$ then K_2 is one-dimensional. In both cases, (B.1) holds in a trivial way.

Henceforth we assume that $\xi_i \neq \eta_i$ $(i = 1, 2)$. Let

$$P_i : V_i \to \mathbb{R}^2$$

denote the linear map such that $P_i : \xi_i \mapsto \binom{1}{0}$ and $P_i : \eta_i \mapsto \binom{0}{1}$. Clearly $P_i(\mathrm{int}(K_i)) = \mathrm{int}(K)$, where K is the positive quadrant in $\mathbb{R}^2$.

Denote the partial orders of K and K_i by $\leq$ and $\leq_i$. Then for all $x, y \in \mathrm{int}(K_i)$ and $m > 0$, $x \leq_i my \Leftrightarrow my - x \in K_i \Leftrightarrow mP_i y - P_i x \in K \Leftrightarrow P_i x \leq mP_i y$. It follows that $M(x|y) = M(P_i x|P_i y)$. Similarly, $m(x|y) = m(P_i x|P_i y)$, and we conclude that

$$d_{K_i}(x, y) = d_K(P_i x, P_i y) \text{ for all } x, y \in \mathrm{int}(K_i). \tag{B.4}$$

Let $A' := P_2 A P_1^{-1}$, then A' is a linear map of $\mathbb{R}^2$ such that $A'(\mathrm{int}(K)) \subset \mathrm{int}(K)$, and (B.4) implies that for all $x, y \in \mathrm{int}(K_1)$ we have $d_{K_2}(Ax, Ay) = d_K(P_2 Ax, P_2 Ay) = d_K(A'P_1 x, A'P_1 y)$. So

$$\begin{aligned} d_{K_2}(Ax, Ay) &= d_K(A'P_1 x, A'P_1 y) \leq \tanh(\Delta_{K,K}(A')/4) d_K(P_1 x, P_1 y) \\ &= \tanh(\Delta_{K_1,K_2}(A)/4) d_{K_1}(x, y) = \tanh(\Delta_{K_1,K_2}(A)/4) d_{K_1}(x, y), \text{ whence (B.1).} \end{aligned}$$

Proof of Theorem B.6 in the General Case: Suppose K_i are cones in a general normed vector spaces V_i, and let $A : V_1 \to V_2$ be a linear map such that $A(\mathrm{int}(K_1)) \subset \mathrm{int}(K_2)$. We estimate $d_{K_2}(Ax, Ay)$ for $x, y \in \mathrm{int}(K_1)$.

If Ax, Ay are collinear, then $d_{K_2}(Ax, Ay) = 0$. Henceforth, assume Ax, Ay are not collinear. Necessarily, x, y are not collinear.

Let $V_1' := \mathrm{span}\{x, y\}$ and $K_1' := K_1 \cap V_1'$. We claim that the interior of K_1' as a subset of V_1' is equal to $\mathrm{int}(K_1) \cap V_1'$:

- Suppose $z \in \mathrm{int}(K_1) \cap V_1'$. Let B_1 denote the open unit ball with center zero in V_1. Then $z + \varepsilon B_1 \subset K_1$ for some $\varepsilon > 0$. Let B_1' denote the open unit ball with center zero in V_1'. Clearly, $B_1' \subset B_1 \cap V_1'$, so $z + \varepsilon B_1' \subset (z + \varepsilon B_1) \cap V_1' \subset K_1 \cap V_1' = K_1'$. Since $z + \varepsilon B_1'$ is a neighborhood of z, z is in the interior of K_1' as a subset of V_1'.
- Suppose z is in the interior of K_1' as a subset of V_1'. Since $x \in V_1'$, $z - \alpha x \in K_1$ for some $\alpha > 0$. Since $x \in \mathrm{int}(K_1)$, $x + \varepsilon B_1 \subset K_1$ for some $\varepsilon > 0$. So $z + \alpha\varepsilon B_1 = (z - \alpha x) + \alpha(x + \varepsilon B_1) \subset K_1 + \alpha K_1 \subset K_1$, whence $z \in \mathrm{int}(K_1) \cap V_1'$.

Consequently, x, y are in the interior of K_1' as a subset of V_1'. As this interior is non-empty, K_1' is a cone in V_1'.

Similarly, if $V_2' := \mathrm{span}\{Ax, Ay\}$ and $K_2' := K_2 \cap V_2'$, then the interior of K_2' as a subset of V_2' is equal to $\mathrm{int}(K_2) \cap V_2'$, which contains Ax, Ay. So K_2' is a cone in V_2'.

Next we claim that $d_{K_i'} = d_{K_i}$ on $\mathrm{int}(K_i) \cap V_i'$ $(i = 1, 2)$. This is because if $x', y' \in K_i'$, then the condition $my' \le x' \le My'$ only involves the vectors $x' - my'$, $My' - x'$ $(m, M > 0)$ which all lie in K_i'.

Clearly, $A(\mathrm{int}(K_1) \cap V_1') \subset \mathrm{int}(K_2) \cap V_2'$, therefore $A : V_1' \to V_2'$ is positive with respect to K_1', K_2'. As V_1', V_2' are two-dimensional, $A : V_1' \to V_2'$ satisfies (B.1), and

$$\begin{aligned} d_{K_2}(Ax, Ay) &= d_{K_2'}(Ax, Ay) \le \tanh(\tfrac{1}{4}\Delta_{K_1',K_2'}(A))d_{K_1'}(x, y) \\ &= \tanh(\tfrac{1}{4}\Delta_{K_1',K_2'}(A))d_{K_1}(x, y). \end{aligned}$$

We claim that $\Delta_{K_1',K_2'}(A) \le \Delta_{K_1,K_2}(A)$:

$$\begin{aligned} \Delta_{K_1',K_2'}(A) &= \sup\{d_{K_2'}(Ax', Ay') : x', y' \in \mathrm{int}(K_1) \cap V_1'\}, \ \because \mathrm{int}(K_1') = \mathrm{int}(K_1) \cap V_1' \\ &\overset{!}{=} \sup\{d_{K_2}(Ax', Ay') : x', y' \in \mathrm{int}(K_1) \cap V_1'\} \le \Delta_{K_1,K_2}(A), \end{aligned}$$

because $A(\mathrm{int}(K_1) \cap V_1') \subset \mathrm{int}(K_2) \cap V_2'$ and $d_{K_2'} = d_{K_2}$ on $\mathrm{int}(K_2) \cap V_2'$.

It follows that $d_{K_2}(Ax, Ay) \le \tanh\left(\frac{1}{4}\Delta_{K_1,K_2}(A)\right) d_{K_1}(x, y)$. □

B.3 Notes and References

Hilbert's projective metric was introduced by David Hilbert in [100]. Theorems B.5 and B.6 and the proofs given here are due to Garrett Birkhoff [13]. For other nice proofs, see [22] and references therein. Rugh [173] extended Hilbert's projective metric to complex cones.

Appendix C
Perturbations of Operators with Spectral Gap

C.1 The Perturbation Theorem

Let $\mathfrak{X}$ be a Banach space over $\mathbb{C}$, and suppose $\mathcal{L}_0 : \mathfrak{X} \to \mathfrak{X}$ is a bounded linear operator. Recall that we say that $\mathcal{L}_0$ has **spectral gap**, with simple **leading eigenvalue $\boldsymbol{\lambda_0}$** and associated **eigenprojection $\boldsymbol{P_0}$**, when

$$\mathcal{L}_0 = \lambda_0 P_0 + N_0,$$

where $\lambda_0 \in \mathbb{C}$, and P_0, N_0 are bounded linear operators with the following properties:

(1) $\mathcal{L}_0 P_0 = P_0 \mathcal{L}_0 = \lambda_0 P_0$; (2) $P_0^2 = P_0$, and $\dim\{P_0 u : u \in \mathfrak{X}\} = 1$;

(3) $P_0 N_0 = N_0 P_0 = 0$, and $\rho(N_0) < |\lambda_0|$, where $\rho(N_0) := \lim_{n\to\infty} \sqrt[n]{\|N_0^n\|}$, the spectral radius of N_0. (Necessarily, $\lambda_0 \neq 0$.)

By Lemma 8.16, in this case the spectrum of $\mathcal{L}_0$ consists of a simple eigenvalue λ_0, and a compact subset of some disk centered at zero with radius $r < |\lambda_0|$.

The purpose of this appendix is to prove the following result, from Chapter 8:

Theorem C.1 (Perturbation Theorem) *Fix $r \geq 1$ and $a > 0$. Suppose $\mathcal{L}_t : \mathfrak{X} \to \mathfrak{X}$ is a bounded linear operator for each $|t| < a$, and $t \mapsto \mathcal{L}_t$ is C^r-smooth. If $\mathcal{L}_0$ has spectral gap with simple leading eigenvalue λ_0 and eigenprojection P_0, then there exists a number $0 < \kappa < a$ such that:*

(1) For each $|t| < \kappa$, $\mathcal{L}_t$ has spectral gap with simple leading eigenvalue λ_t, and associated eigenprojection P_t;
(2) $t \mapsto \lambda_t$ and $t \mapsto P_t$ are C^r-smooth on $(-\kappa, \kappa)$;
(3) There exists $\gamma > 0$ such that $\rho(\mathcal{L}_t - \lambda_t P_t) < |\lambda_t| - \gamma$ for all $|t| < \kappa$.

C.2 Some Facts from Analysis

Let $\mathfrak{X}^*$ denote the dual of $\mathfrak{X}$. Every bounded linear operator $A : \mathfrak{X} \to \mathfrak{X}$ determines a bounded linear operator A^* on $\mathfrak{X}^*$ via $(A^*\varphi)(x) = \varphi(Ax)$. It holds that $\|A^*\| = \|A\|$. Recall the definition of C^r-smoothness from §8.4.

Lemma C.2 *Suppose that for each $|t| < a$ we have a scalar $c(t) \in \mathbb{C}$, a vector $h_t \in \mathfrak{X}$, a bounded linear functional $\varphi_t \in \mathfrak{X}^*$, and bounded linear operators $A_t, B_t : \mathfrak{X} \to \mathfrak{X}$. If $t \mapsto c(t), h_t, \varphi_t, A_t, B_t$ are C^r-smooth on $(-a, a)$, then the following objects are C^r-smooth on $(-a, a)$:*

(1) A_t^*, $c(t)A_t$, $A_t + B_t$, $A_t B_t$; (2) the operator $\varphi_t(\cdot)h_t$; (3) the scalar $\varphi_t(A_t h_t)$.

Proof We prove (1) and leave (2) and (3) to the reader.

Suppose $r = 1$. Since A_t is C^1, $A_{t+\theta}x = A_t x + \theta A_t' x + \varepsilon_\theta(x)$ where $\|\varepsilon_\theta(x)\| = o(|\theta|)\|x\|$. So for every $\varphi \in \mathfrak{X}^*$,

$$\begin{aligned}((A_{t+\theta}^* - A_t^*)\varphi)(x) &= \varphi[A_{t+\theta}x - A_t x] = \varphi[A_t x + \theta A_t' x + \varepsilon_\theta(x) - A_t x] \\ &= \theta(A_t'^*\varphi)(x) + o(|\theta|)\|x\|\|\varphi\|.\end{aligned}$$

So $\left\|\frac{A_{t+\theta}^* - A_t^*}{\theta} - (A_t')^*\right\| \xrightarrow[\theta\to 0]{} 0$, and A_t^* is differentiable, with derivative $(A_t')^*$. Next $\|(A_{t+\theta}')^* - (A_t')^*\| = \|A_{t+\theta}' - A_t'\| \xrightarrow[\theta\to 0]{} 0$, because A_t is C^1. This proves that A_t^* is C^1 with derivative $(A_t')^*$.

Similarly, one shows that $A_t + B_t$, $c(t)A_t$ and $A_t B_t$ are C^1 with derivatives $A_t' + B_t'$, $c'(t)A_t + c(t)A_t'$, and $A_t' B_t + A_t B_t'$ (the operators need not commute, and the order matters). Part (1) follows, in case $r = 1$. For higher r, we argue by induction. □

Lemma C.3 *If $\rho(L) < |\lambda|$, then $\lambda^{-1}L - I$ has a bounded inverse. (I :=identity.)*

Proof Let $A := -\sum_{n\geq 0} \lambda^{-n}L^n$. The sum converges in norm because $\rho(\lambda^{-1}L) < 1$, and a straightforward calculation shows that $A(\lambda^{-1}L - I) = (\lambda^{-1}L - I)A = I$. □

Theorem C.4 (Implicit Function Theorem) *Suppose $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}$ are Banach spaces, $A \subset \mathfrak{X} \times \mathfrak{Y}$ is open, and $F : A \to \mathfrak{Z}$ is C^1. Suppose $(x_0, y_0) \in A$ is a point where*

(1) $F(x_0, y_0) = 0$, and
(2) $(\partial_y F)(x_0, y_0) : \mathfrak{Y} \to \mathfrak{Z}$, the partial derivative of F with respect to the second variable at (x_0, y_0), is an invertible bounded operator with a bounded inverse.

Then there exists an open neighborhood $U \subset \mathfrak{X}$ of x_0 and a unique C^1 map $u : U \to \mathfrak{Y}$ such that $(x, u(x)) \in A$ for all $x \in U$, $u(x_0) = y_0$, and $F(x, u(x)) = 0$. If F is C^r-smooth on A, then u is C^r-smooth on U.

The proof is a fixed-point argument, see [49, Ch. 10].

C.3 Proof of the Perturbation Theorem

Since $\dim\{P_0 u : u \in \mathfrak{X}\} = 1$, P_0 must take the form $P_0 u = \varphi_0(u)h_0$, where $\varphi_0 \in \mathfrak{X}^*$, h_0 is a non-zero vector in $\mathfrak{X}$, and $\varphi_0(h_0) = 1$. Since $\mathcal{L}_0 P_0 = P_0 \mathcal{L}_0 = \lambda_0 P_0$,

$$\mathcal{L}_0 h_0 = \lambda_0 h_0 \quad \text{and} \quad \mathcal{L}_0^* \varphi_0 = \lambda_0 \varphi_0.$$

To prove the perturbation theorem, we will construct for small $|t|$ a scalar $\lambda_t \in \mathbb{C}$, a bounded linear functional $\varphi_t \in \mathfrak{X}^*$ and a vector $h_t \in \mathfrak{X}$, all depending C^r-smoothly on t, such that

$$\mathcal{L}_t h_t = \lambda h_t, \quad \mathcal{L}_t^* \varphi_t = \lambda_t \varphi_t, \quad \varphi_t(h_t) = 1. \tag{C.1}$$

Then we will show that $\mathcal{L}_t$ has spectral gap with simple leading eigenvalue λ_t, and eigenprojection $P_t u := \varphi_t(u)h_t$.

Construction of h_t and λ_t: Let $\mathfrak{X}_0 := \ker(P_0) = \ker(\varphi_0)$. This is a Banach space.

Choose $0 < \varepsilon_0, \kappa_1 < a$ so small, that for all $|t| < \kappa_1$, $|\varphi_0(\mathcal{L}_t h_0)| > \varepsilon_0$. This is possible, because $t \mapsto \mathcal{L}_t$ is continuous, and $|\varphi_0(\mathcal{L}_0 h_0)| = |\lambda_0| > 0$.

It follows that for some $\varepsilon_1 > 0$,

$$\varphi_0(\mathcal{L}_t(h_0 + w)) \neq 0, \text{ whenever } |t| < \kappa_1, \|w\| < \varepsilon_1.$$

Let $B_1 := \{w \in \mathfrak{X}_0 : \|w\| < \varepsilon_1\}$, and define

$$F(t, w) = \frac{\mathcal{L}_t(h_0 + w)}{\varphi_0[\mathcal{L}_t(h_0 + w)]} - (h_0 + w).$$

F is C^r on $(-\kappa_1, \kappa_1) \times B_1$ by Lemma C.2, and $F(0, 0) = \frac{\lambda_0 h_0}{\lambda_0} - h_0 = 0$.

We claim that $(\partial_w F)(0, 0) : \mathfrak{X}_0 \to \mathfrak{X}_0$ is bounded, invertible, and has bounded inverse. Here is the proof. For every $w \in \mathfrak{X}_0$,

$$\varphi_0(\mathcal{L}_0(h_0 + w)) = (\mathcal{L}_0^* \varphi_0)(h_0 + w) = \lambda_0 \varphi_0(h_0 + w) = \lambda_0,$$

because $\mathfrak{X}_0 = \ker(\varphi_0)$. Therefore

$$F(0, w) - F(0, 0) = \frac{\mathcal{L}_0(h_0 + w)}{\varphi_0(\mathcal{L}_0(h_0 + w)))} - (h_0 + w) = \frac{\lambda_0 h_0 + \mathcal{L}_0 w}{\lambda_0} - (h_0 + w)$$
$$= \lambda_0^{-1} \mathcal{L}_0 w - w.$$

Therefore $(\partial_w F)(0,0) = (\lambda_0^{-1}\mathcal{L}_0 - I)|_{\mathfrak{X}_0}$. Since P_0 vanishes on $\mathfrak{X}_0$,

$$\rho(\mathcal{L}_0|_{\mathfrak{X}_0}) = \rho(\mathcal{L}_0(I - P_0)|_{\mathfrak{X}_0}) = \rho(N_0|_{\mathfrak{X}_0}) \le \rho(N_0) < |\lambda_0|.$$

By Lemma C.3, $(\lambda_0^{-1}\mathcal{L}_0 - I)|_{\mathfrak{X}_0}$ has a bounded inverse.

By the inverse function theorem, there is $0 < \kappa_2 < \kappa_1$ and a C^r map $t \mapsto w_t \in \mathfrak{X}_0$ on $(-\kappa_2, \kappa_2)$ such that $F(t, w_t) = 0$.

Equivalently, $\mathcal{L}_t(h_0 + w_t) = \varphi_0(\mathcal{L}_t(h_0 + w_t))(h_0 + w_t)$. Therefore

$$\mathcal{L}_t h_t = \lambda_t h_t,$$

where $h_t := h_0 + w_t$, $\lambda_t := \varphi_0(\mathcal{L}_t(h_0 + w_t))$. Since $w_t \in \mathfrak{X}_0$, $\varphi_0(h_t) = \varphi(h_0) = 1$.

By Lemma C.2, $t \mapsto \lambda_t, h_t$ are C^r-smooth on $(-\kappa_2, \kappa_2)$. By further decreasing κ_2, if necessary, we can guarantee that λ_t is close enough to λ_0 to be non-zero.

Construction of φ_t: Let $\mathfrak{X}_0^* := \{\varphi \in \mathfrak{X}^* : \varphi(h_0) = 0\}$. This is a Banach space. Let

$$\pi : \mathfrak{X}^* \to \mathfrak{X}_0^*,\ \pi(\varphi) := \varphi - \varphi(h_0)\varphi_0.$$

Note that $\pi|_{\mathfrak{X}_0^*} = I^*$ (the identity on $\mathfrak{X}_0^*$), and $\pi(\mathfrak{X}^*) = \mathfrak{X}_0^*$, because

$$\pi(\varphi)(h_0) = \varphi(h_0) - \varphi(h_0)\varphi_0(h_0) = 0.$$

Define $G : (-\kappa_1, \kappa_1) \times \mathfrak{X}_0^* \to \mathfrak{X}_0^*$ by

$$G(t, \varphi) := \pi[\mathcal{L}_t^*(\varphi_0 + \varphi) - \lambda_t(\varphi_0 + \varphi)] \overset{!}{=} \pi[\mathcal{L}_t^*(\varphi_0 + \varphi)] - \lambda_t \varphi, \quad \because \pi(\varphi_0) = 0.$$

G is C^r on its domain, and $G(0,0) = 0$. Since G is affine in its second coordinate,

$$(\partial_\varphi G)(0,0) = \pi \circ (\mathcal{L}_0^* - \lambda_0 I^*)|_{\mathfrak{X}_0^*} = (\pi\mathcal{L}_0^* - \lambda_0 I^*)|_{\mathfrak{X}_0^*} \overset{!}{=} (\mathcal{L}_0^* - \lambda_0 I^*)|_{\mathfrak{X}_0^*}.$$

To see the marked identity, note that if $\varphi \in \mathfrak{X}_0^*$, then $\mathcal{L}_0^*\varphi \in \mathfrak{X}_0^*$, because $\mathcal{L}_0 h_0 = \lambda_0 h_0$, and therefore $\pi[\mathcal{L}_0^*\varphi] = \mathcal{L}_0^*\varphi$ on $\mathfrak{X}_0^*$.

Clearly, $(\mathcal{L}_0^* - \lambda_0 I^*)|_{\mathfrak{X}_0^*}$ is bounded. It has a bounded inverse on $\mathfrak{X}_0^*$, because $P_0^*|_{\mathfrak{X}_0^*} = 0$, and therefore

$$\rho(\mathcal{L}_0^*|_{\mathfrak{X}_0^*}) = \rho(\mathcal{L}_0^* - \lambda_0 P_0^*|_{\mathfrak{X}_0^*}) = \rho(N_0^*|_{\mathfrak{X}_0^*}) \le \rho(N_0^*) = \rho(N_0) < |\lambda_0|.$$

We can now apply the inverse function theorem, and construct $0 < \kappa_3 < \kappa_2$ and a C^r-smooth function $t \mapsto \psi_t$ on $(-\kappa_3, \kappa_3)$ so that $G(t, \psi_t) = 0$. Equivalently,

$$\mathcal{L}_t^*(\varphi_0 + \psi_t) - \lambda_t(\varphi_0 + \psi_t) = [\mathcal{L}_t^*(\varphi_0 + \psi_t) - \lambda_t(\varphi_0 + \psi_t)](h_0) \cdot \varphi_0. \tag{C.2}$$

Evaluating the two sides of (C.2) at h_t, we obtain

$$0 = [\mathcal{L}_t^*(\varphi_0 + \psi_t) - \lambda_t(\varphi_0 + \psi_t)](h_t) = [\mathcal{L}_t^*(\varphi_0 + \psi_t) - \lambda_t(\varphi_0 + \psi_t)](h_0) \tag{C.3}$$

(The left-hand side of (C.2) vanishes because $\mathcal{L}_t h_t = \lambda_t h_t$, the right-hand side of (C.3) reduces to the right-hand of (C.2) because $\varphi_0(h_t) = 1$). Thus the right-hand-side of (C.2) vanishes. It follows that the left-hand-side (C.2) vanishes, and

$$\mathcal{L}_t^*(\varphi_0 + \psi_t) = \lambda_t(\varphi_0 + \psi_t).$$

Since $t \mapsto (\varphi_0 + \psi_t)(h_t)$ is smooth, and $(\varphi_0 + \psi_0)(h_0) = \varphi_0(h_0) = 1$, there is $0 < \kappa_4 < \kappa_3$ such that $(\varphi_0 + \psi_t)(h_t) \neq 0$ for $|t| < \kappa_4$. Now take

$$\varphi_t := \frac{\varphi_0 + \psi_t}{(\varphi_0 + \psi_t)(h_t)}.$$

Completion of the Proof Define the operators $P_t u := \varphi_t(u) h_t$ and $N_t := \mathcal{L}_t - \lambda_t P_t$. Then it is straightforward to verify, using (C.1), that

$$P_t^2 = P_t\,,\ P_t\mathcal{L}_t = \mathcal{L}_t P_t = \lambda_t P_t\,,\ P_t N_t = N_t P_t = 0.$$

Clearly, $\dim\{P_t u : u \in \mathfrak{X}\} = \mathrm{Span}\,\{h_t\} = 1$, and by Lemma C.2, $t \mapsto P_t$ is C^r-smooth on $(-\kappa_4, \kappa_4)$.

We claim that there is $0 < \kappa_5 < \kappa_4$ such that

$$\sup_{|t|<\kappa_5} \rho(N_t) < \inf_{|t|<\kappa_4} |\lambda_t|.$$

By assumption $\rho(N_0) < |\lambda_0|$, therefore there exists some n_0 such that $\|N_0^{n_0}\| < |\lambda_0|^{n_0}$. Choose some $\gamma > 0$ such that

$$\|N_0^{n_0}\| \le (1-2\gamma)^{n_0}|\lambda_0|^{n_0}.$$

Since $t \mapsto P_t$ and $t \mapsto \mathcal{L}_t$ are continuous at $t = 0$, $t \mapsto \|N_t^{n_0}\|$ is continuous at $t = 0$. Therefore there exists $0 < \kappa_5 < \kappa_4$ such that

$$\|N_t^{n_0}\| \le (1-\gamma)^{n_0}|\lambda_t|^{n_0} \text{ for all } |t| < \kappa_5.$$

Necessarily, $\rho(N_t) \le (1-\gamma)|\lambda_t|$ for $|t| < \kappa_5$. □

C.4 Notes and References

For a comprehensive account of the theory of perturbations of linear operators on Banach spaces, see [105]. The proof we gave here is taken from [98, §XIV.2].

Suppose $\mathcal{L}_t$ is a smooth perturbation of an operator $\mathcal{L}_0$ with spectral gap. The perturbation theorem provides smooth solutions to

$$\mathcal{L}_t h_t = \lambda_t h_t.$$

In §7.3.2 we considered a similar eigenvector problem, but in the inhomogeneous setup. There the unperturbed operators are

$$\mathcal{L}_n : L^\infty(\mathfrak{S}_{n+1}) \to L^\infty(\mathfrak{S}_n)\,,\ (\mathcal{L}_n\varphi)(x) = \mathbb{E}(\varphi(X_{n+1})|X_n = x),$$

the perturbations are $(\mathcal{L}_n^t\varphi)(x) = \mathbb{E}(\mathrm{e}^{t f_{n+1}(X_n, X_{n+1})}\varphi(X_{n+1})|X_n = x)$, and the eigenvector problem is to find smooth families of h_n^t and λ_n^t such that $\mathcal{L}_{n+1}^t h_{n+1}^t = \lambda_n h_n^t$.

In §7.3.2 we solved the problem, using the contraction properties of linear maps mapping cones into other cones. It is also possible to solve the problem using “inhomogeneous versions" of the perturbation theorem.

Suppose first that f_n are uniformly bounded. Hafouta and Kifer ([93]) proved the *sequential complex Ruelle-Perron Frobenius Theorem*[1], which provides holomorphic solutions of $\mathcal{L}_t h_t = \lambda_t h_t$, for t in a small *complex* neighborhood of $t = 0$. By contrast, the methods of §7.3.2 rely on the positivity of the function $\mathrm{e}^{t f_n}$ in a crucial way, and are therefore limited to *real* t. To deal with complex t, Hafouta and Kifer use a highly non-trivial extension of Hilbert projective metric to complex cones, due to Rugh [173] and Dubois [67].

Yeor Hafouta informed us of a different proof of the complex Ruelle-Perron Frobenius Theorem, based on the approach of [98], which also covers unbounded f_n with finite L^2 norm.

The analyticity of complex perturbations opens the way to employing perturbative calculations similar to those done in Lemma 7.17 and Proposition 8.21, but for complex t and inhomogeneous (X, f). For applications to Edgeworth expansions and Berry-Esseen estimates for inhomogeneous Markov chains, see [59] and the notes to Chapters 6 and 8. (The homogenous setup can be analyzed using the spectral theorem, see [26, 73, 74, 85, 98, 99].)

---

[1] For the non-sequential case, see [161].

References

1. J. Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.
2. J. Aaronson and M. Denker. Distributional limits for hyperbolic infinite volume geodesic flows. *Tr. Mat. Inst. Steklova*, 216(1):181–192, 1997.
3. J. Aaronson and M. Denker. The Poincaré series of $\mathbf{C} \setminus \mathbf{Z}$. *Ergodic Theory Dynam. Systems*, 19(1):1–20, 1999.
4. J. Aaronson and M. Denker. A local limit theorem for stationary processes in the domain of attraction of a normal distribution. In *Asymptotic methods in probability and statistics with applications (St. Petersburg, 1998)*, Stat. Ind. Technol., pages 215–223. Birkhäuser Boston, Boston, MA, 2001.
5. J. Aaronson and M. Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.*, 1(2):193–237, 2001.
6. J. Aaronson and M. Denker. Group extensions of Gibbs-Markov maps. *Probab. Theory Related Fields*, 123(1):28–40, 2002.
7. J. Aaronson, H. Nakada, O. Sarig, and R. Solomyak. Invariant measures and asymptotics for some skew products. *Israel J. Math.*, 128:93–134, 2002.
8. J. Aaronson and B. Weiss. Remarks on the tightness of cocycles. *Colloq. Math.*, 84/85(part 2):363–376, 2000.
9. R. Aimino, M. Nicol, and S. Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields*, 162(1-2):233–274, 2015.
10. Martine Babillot and François Ledrappier. Lalley's theorem on periodic orbits of hyperbolic flows. *Ergodic Theory Dynam. Systems*, 18(1):17–39, 1998.
11. R. R. Bahadur and R. Ranga Rao. On deviations of the sample mean. *Ann. Math. Statist.*, 31:1015–1027, 1960.
12. V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings. I. *Izv. Ross. Akad. Nauk Ser. Mat.*, 58(2):40–72, 1994.
13. G. Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.
14. D. Blackwell and J. L. Hodges, Jr. The probability in the extreme tail of a convolution. *Ann. Math. Statist.*, 30:1113–1120, 1959.
15. T. Bogenschütz and V. M. Gundlach. Ruelle's transfer operator for random subshifts of finite type. *Ergodic Theory Dynam. Systems*, 15(3):413–447, 1995.
16. R. C. Bradley. *Introduction to strong mixing conditions. Vol. 1,2,3*. Kendrick Press, Heber City, UT, 2007.
17. L. Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
18. J. Brémont. Gibbs measures at temperature zero. *Nonlinearity*, 16(2):419–426, 2003.
19. M. I. Brin. Topological transitivity of a certain class of dynamical systems, and flows of frames on manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 9(1):9–19, 1975.
20. A. Broise. Transformations dilatantes de l'intervalle et théorèmes limites. Etudes spectrales d'opérateurs de transfert et applications. *Astérisque*, (238):1–109, 1996.
21. H. Bruin, M. Holland, and M. Nicol. Livšic regularity for Markov systems. *Ergodic Theory Dynam. Systems*, 25(6):1739–1765, 2005.
22. P. J. Bushell. Hilbert's metric and positive contraction mappings in a Banach space. *Arch. Rational Mech. Anal.*, 52:330–338, 1973.
23. N. R. Chaganty and J. Sethuraman. Large deviation local limit theorems for arbitrary sequences of random variables. *Ann. Probab.*, 13(1):97–114, 1985.
24. N. R. Chaganty and J. Sethuraman. Strong large deviation and local limit theorems. *Ann. Probab.*, 21(3):1671–1690, 1993.
25. N. I. Chernov. Markov approximations and decay of correlations for Anosov flows. *Ann. of Math. (2)*, 147(2):269–324, 1998.
26. Zaqueu Coelho and William Parry. Central limit asymptotics for shifts of finite type. *Israel J. Math.*, 69(2):235–249, 1990.
27. R. Cogburn. Markov chains in random environments: the case of Markovian environments. *Ann. Probab.*, 8(5):908–916, 1980.
28. R. Cogburn. On the central limit theorem for Markov chains in random environments. *Ann. Probab.*, 19(2):587–604, 1991.

29. G. Contreras and R. Iturriaga. *Global minimizers of autonomous Lagrangians*. 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.
30. J.-P. Conze, S. Le Borgne, and M. Roger. Central limit theorem for stationary products of toral automorphisms. *Discrete Contin. Dyn. Syst.*, 32(5):1597–1626, 2012.
31. J.-P. Conze and A. Raugi. Limit theorems for sequential expanding dynamical systems on [0, 1]. In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 89–121. Amer. Math. Soc., Providence, RI, 2007.
32. J.-P. Conze and A. Raugi. On the ergodic decomposition for a cocycle. *Colloq. Math.*, 117(1):121–156, 2009.
33. I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinaĭ. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.
34. H. Cramér. Sur une nouveau théorème-limite de la théorie des probabilités, 1938. *Reprinted in H. Cramér, Collected works, A. Martin-Löf (Ed.), vol II, Springer, Berlin, 1994*, pages 895–913.
35. Christophe Cuny and Florence Merlevède. Strong invariance principles with rate for "reverse" martingale differences and applications. *J. Theoret. Probab.*, 28(1):137–183, 2015.
36. K. È. Dambis. On decomposition of continuous submartingales. *Teor. Verojatnost. i Primenen.*, 10:438–448, 1965.
37. R. de la Llave, J. M. Marco, and R. Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math. (2)*, 123(3):537–611, 1986.
38. A. de Moivre. *The doctrine of chances or, a method of calculating the probabilities of events in play*. New impression of the second edition, with additional material. Cass Library of Science Classics, No. 1. Frank Cass & Co., Ltd., London, 1967.
39. A. Dembo and O. Zeitouni. Large deviations via parameter dependent change of measure, and an application to the lower tail of Gaussian processes. In *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progr. Probab.*, pages 111–121. Birkhäuser, Basel, 1995.
40. A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
41. M. F. Demers, F. Pène, and H.-K. Zhang. Local limit theorem for randomly deforming billiards. *Comm. Math. Phys.*, 375(3):2281–2334, 2020.
42. M. Denker and M. Gordin. The central limit theorem for random perturbations of rotations. *Probab. Theory Related Fields*, 111(1):1–16, 1998.
43. M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems*, 4(4):541–552, 1984.
44. M. Denker and X. Zheng. On the local times of stationary processes with conditional local limit theorems. *Stochastic Process. Appl.*, 128(7):2448–2462, 2018.
45. B. Derrida and T. Sadhu. Large deviations conditioned on large deviations I: Markov chain and Langevin equation. *J. Stat. Phys.*, 176(4):773–805, 2019.
46. Y. Derriennic and M. Lin. The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Related Fields*, 119(4):508–528, 2001.
47. Z. Dietz and S. Sethuraman. Large deviations for a class of nonhomogeneous Markov chains. *Ann. Appl. Probab.*, 15(1A):421–486, 2005.
48. Z. Dietz and S. Sethuraman. Occupation laws for some time-nonhomogeneous Markov chains. *Electron. J. Probab.*, 12:no. 23, 661–683, 2007.
49. J. Dieudonné. *Foundations of modern analysis*. Academic Press, New York-London, 1969. Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.
50. R. Dobrushin. Central limit theorem for non-stationary Markov chains. i, ii. *Theory of Probab. & Appl.*, 1:65–80, 329–383, 1956.
51. W. Doeblin. Le cas discontinu des probabilités en chaîne. *Publ. Fac. Sci. Univ. Masaryk (Brno)*, (236), 1937.
52. W. Doeblin. Sur les propriétes asymptotiques de mouvement régis par certains types de chaînes simples. *Bull. Math. Soc. Roum. Sci.*, 39(1):57–115, 1937.
53. W. Doeblin. Sur les sommes d'un grand nombres de variables aleatoires independantes. *Bull. Soc. math. France*, 53:23–64, 1939.
54. D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
55. D. Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergodic Theory Dynam. Systems*, 18(5):1097–1114, 1998.
56. D. Dolgopyat. A local limit theorem for sums of independent random vectors. *Electron. J. Probab.*, 21:Paper No. 39, 15, 2016.
57. D. Dolgopyat, C. Dong, A. Kanigowski, and P. Nándori. Flexibility of statistical properties for smooth systems satisfying the central limit theorem. *Invent. Math.*, 230(1):31–120, 2022.
58. D. Dolgopyat, C. Dong, A. Kanigowski, and P. Nándori. Mixing properties of generalized T, T^{-1} transformations. *Israel J. Math.*, 247(1):21–73, 2022.
59. D. Dolgopyat and Y. Hafouta. A Berry-Esseen theorem and Edgeworth expansions for uniformly elliptic inhomogeneous Markov chains. *To appear in PTRF*.
60. D. Dolgopyat and P. Nándori. Infinite measure mixing for some mechanical systems. *Adv. Math.*, 410(B):Paper No. 108757, 56, 2022.
61. D. Dolgopyat, D. Szász, and T. Varjú. Recurrence properties of planar Lorentz process. *Duke Math. J.*, 142(2):241–281, 2008.

62. J. L. Doob. *Stochastic processes*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1990. Reprint of the 1953 original, A Wiley-Interscience Publication.
63. D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti. Almost sure invariance principle for random piecewise expanding maps. *Nonlinearity*, 31(5):2252–2280, 2018.
64. D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti. A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys.*, 360(3):1121–1187, 2018.
65. D. Dragičević and Y. Hafouta. Almost sure invariance principle for random dynamical systems via Gouëzel's approach. *Nonlinearity*, 34(10):6773–6798, 2021.
66. L. E. Dubins and G. Schwarz. On continuous martingales. *Proc. Nat. Acad. Sci. U.S.A.*, 53:913–916, 1965.
67. L. Dubois. Real cone contractions and analyticity properties of the characteristic exponents. *Nonlinearity*, 21(11):2519–2536, 2008.
68. R. Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. 5th edition.
69. R. S. Ellis. Large deviations for a general class of random vectors. *Ann. Probab.*, 12(1):1–12, 1984.
70. R. S. Ellis. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original.
71. C. G. Esseen. On the Kolmogorov-Rogozin inequality for the concentration function. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 5:210–216, 1966.
72. C. G. Esseen. On the concentration function of a sum of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 9:290–308, 1968.
73. K. Fernando and C. Liverani. Edgeworth expansions for weakly dependent random variables. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(1):469–505, 2021.
74. K. Fernando and F. Pène. Expansions in the local and the central limit theorems for dynamical systems. *Comm. Math. Phys.*, 389(1):273–347, 2022.
75. P. Ferrero and B. Schmitt. Produits aléatoires d'opérateurs matrices de transfert. *Probab. Theory Related Fields*, 79(2):227–248, 1988.
76. N. G. Gamkrelidze. On a local limit theorem for lattice random variables. *Teor. Verojatnost. i Primenen*, 9:733–736, 1964.
77. J. Gärtner. On large deviations from an invariant measure. *Teor. Verojatnost. i Primenen.*, 22(1):27–42, 1977.
78. B. V. Gnedenko. On a local limit theorem of the theory of probability. *Uspehi Matem. Nauk (N. S.)*, 3(3(25)):187–194, 1948.
79. B. V. Gnedenko. On a local theorem for the region of normal attraction of stable laws. *Doklady Akad. Nauk SSSR (N.S.)*, 66:325–326, 1949.
80. B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954. Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob.
81. M. I. Gordin. The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR*, 188:739–741, 1969.
82. M. I. Gordin and B. A. Lifšic. Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, 239(4):766–767, 1978.
83. W. H. Gottschalk and G. A. Hedlund. *Topological dynamics*. American Mathematical Society Colloquium Publications, Vol. 36. American Mathematical Society, Providence, R. I., 1955.
84. S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Related Fields*, 128(1):82–122, 2004.
85. S. Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(6):997–1024, 2005.
86. S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.*, 38(4):1639–1671, 2010.
87. S. Gouëzel. Limit theorems in dynamical systems using the spectral method. In *Hyperbolic dynamics, fluctuations and large deviations*, volume 89 of *Proc. Sympos. Pure Math.*, pages 161–193. Amer. Math. Soc., Providence, RI, 2015.
88. Y. Guivarc'h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. *Ann. Inst. H. Poincaré Prob. Stat.*, 24(1):73–98, 1988.
89. Y. Hafouta. Limit theorems for some skew-products with mixing base maps. *Preprint (2018), arXiv:1808.00735v5*.
90. Y. Hafouta. A sequential RPF theorem and its applications to limit theorems for time dependent dynamical systems and inhomogeneous markov chains. *Preprint (2019), arXiv:1903.04018v4*.
91. Y. Hafouta. Limit theorems for some time-dependent expanding dynamical systems. *Nonlinearity*, 33(12):6421–6460, 2020.
92. Y. Hafouta and Yu. Kifer. A nonconventional local limit theorem. *J. Theoret. Probab.*, 29(4):1524–1553, 2016.
93. Y. Hafouta and Yu. Kifer. *Nonconventional limit theorems and random dynamics*. World Scientific, Hackensack, NJ, 2018.
94. J. Hajnal. Weak ergodicity in non-homogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 54:233–246, 1958.
95. P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press., New York-London, 1980. Probability and Mathematical Statistics.
96. P. Hartman and A. Wintner. On the law of the iterated logarithm. *Amer. J. Math.*, 63:169–176, 1941.
97. N. Haydn, M. Nicol, A. Török, and S. Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.*, 369(8):5293–5316, 2017.
98. H. Hennion and L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

99. L. Hervé and F. Pène. The Nagaev-Guivarc'h method via the Keller-Liverani theorem. *Bull. Soc. Math. France*, 138(3):415–489, 2010.
100. D. Hilbert. Über die gerade linie als kürzeste verbindung zweier punkte. *Math. Ann.*, 46:91–96, 1895.
101. L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
102. L. Hörmander. *The analysis of linear partial differential operators. III*, volume 274 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985. Pseudodifferential operators.
103. I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.
104. K. Itô and H. P. McKean, Jr. *Diffusion processes and their sample paths*. Die Grundlehren der mathematischen Wissenschaften, Band 125. Springer-Verlag, Berlin-New York, 1974. Second printing, corrected.
105. T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
106. A. Katok and A. Kononenko. Cocycles' stability for partially hyperbolic systems. *Math. Res. Lett.*, 3(2):191–210, 1996.
107. A. Katsuda and T. Sunada. Closed orbits in homology classes. *Inst. Hautes Études Sci. Publ. Math.*, (71):5–32, 1990.
108. A. Ya. Khinchin. Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. *Math. Ann.*, 92(1-2):115–125, 1924.
109. A. Ya. Khinchin. *Mathematical foundations of quantum statistics*. Dover Publications, Mineola, NY, 1998. Translated from the Russian by E. J. Kelly, Jr., M. D. Friedman, W. H. Furry and A. H. Halperin.
110. Yu. Kifer. *Ergodic theory of random transformations*, volume 10 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1986.
111. Yu. Kifer. Perron-Frobenius theorem, large deviations, and random perturbations in random environments. *Math. Z.*, 222(4):677–698, 1996.
112. Yu. Kifer. Limit theorems for random transformations and processes in random environments. *Trans. Amer. Math. Soc.*, 350(4):1481–1518, 1998.
113. Yu. Kifer. Thermodynamic formalism for random transformations revisited. *Stoch. Dyn.*, 8(1):77–102, 2008.
114. C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
115. A. N. Kolmogorov. Über das Gesetz des iterierten Logarithmus. *Math. Ann.*, 101(1):126–135, 1929.
116. A. N. Kolmogorov. A local limit theorem for classical Markov chains. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 13:281–300, 1949.
117. A. N. Kolmogorov and A. Ya. Khinchin. Über konvergenz von reihen, deren glieder durch den zufall bestimmt weden. *Mat. Sb.*, 32:668–667, 1925.
118. J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32:111–131, 1975.
119. T. Komorowski, C. Landim, and S. Olla. *Fluctuations in Markov processes. Time symmetry and martingale approximation*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012.
120. A. Korepanov, Z. Kosloff, and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(4):859–885, 2018.
121. Z. Kosloff and D. Volny. Local limit theorem in deterministic systems. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(1):548–566, 2022.
122. S. P. Lalley. Closed geodesics in homology classes on surfaces of variable negative curvature. *Duke Math. J.*, 58(3):795–821, 1989.
123. S. P. Lalley. Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits. *Acta Math.*, 163(1-2):1–55, 1989.
124. P.-S. Laplace. *Théorie analytique des probabilités. Vol. I*. Éditions Jacques Gabay, Paris, 1995. Introduction: Essai philosophique sur les probabilités. [Introduction: Philosophical essay on probabilities], Livre I: Du calcul des fonctions génératrices. [Book I: On the calculus of generating functions], Reprint of the 1820 3d edition.
125. P.-S. Laplace. *Théorie analytique des probabilités. Vol. II*. Éditions Jacques Gabay, Paris, 1995. Livre II: Théorie générale des probabilités. [Book II: General probability theory], Suppléments. [Supplements], Reprint of the 1820 3d edition.
126. F. Ledrappier and O. Sarig. Unique ergodicity for non-uniquely ergodic horocycle flows. *Discrete Contin. Dyn. Syst.*, 16(2):411–433, 2006.
127. F. Ledrappier and O. Sarig. Fluctuations of ergodic sums for horocycle flows on $\mathbb{Z}^d$-covers of finite volume surfaces. *Discrete Contin. Dyn. Syst.*, 22(1-2):247–325, 2008.
128. M. Lemańczyk. Analytic nonregular cocycles over irrational rotations. *Comment. Math. Univ. Carolin.*, 36(4):727–735, 1995.
129. V. P. Leonov. On the dispersion of time means of a stationary stochastic process. *Teor. Verojatnost. i Primenen.*, 6:93–101, 1961.
130. P. Lévy. *Théorie de l'addition des variables aléatoires*. Gautier-Villars, 1937.
131. B. A. Lifšic. The convergence of moments in the central limit theorem for inhomogeneous Markov chains. *Teor. Verojatnost. i Primenen.*, 20(4):755–772, 1975.

132. J. W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, 15(1):211–225, 1922.
133. Ju. V. Linnik. Limit theorems for the sums of independent variables taking into account the large deviations. I. *Teor. Verojatnost. i Primenen.*, 6:145–163, 1961.
134. A. N. Livšic. Certain properties of the homology of Y-systems. *Mat. Zametki*, 10:555–564, 1971.
135. A. N. Livšic. Cohomology of dynamical systems. *Izv. Akad. Nauk SSSR Ser. Mat.*, 36:1296–1320, 1972.
136. A.M. Lyapunov. Sur une proposition de la théorie des probabilités. *Bull. de l'Academie Imperiale des Sci. de St. Petersbourg*, 13(4):359–386, 1900.
137. A.A. Markov. Extension of the law of large numbers to dependent events (russian). *Bull. Soc. Phys. Math. Kazan*, 15(2):135–156, 1906.
138. P. Matula. A note on the almost sure convergence of sums of negatively dependent random variables. *Statist. Probab. Lett.*, 15(3):209–213, 1992.
139. M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, 28(2):713–724, 2000.
140. D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability*, 2:620–628, 1974.
141. I. Melbourne and D. Terhesiu. Analytic proof of multivariate stable local large deviations and application to deterministic dynamical systems. *Electron. J. Probab.*, 27:Paper No. 21, 17, 2022.
142. F. Merlevède, M. Peligrad, and C. Peligrad. On the local limit theorems for psi-mixing Markov chains. *ALEA Lat. Am. J. Probab. Math. Stat.*, 18(2):1221–1239, 2021.
143. F. Merlevède, M. Peligrad, and S. Utev. *Functional Gaussian approximation for dependent structures*, volume 6 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2019.
144. J. Mineka and S. Silverman. A local limit theorem and recurrence conditions for sums of independent non-lattice random variables. *Ann. Math. Statist.*, 41:592–600, 1970.
145. C. C. Moore and K. Schmidt. Coboundaries and homomorphisms for nonsingular actions and a problem of H. Helson. *Proc. London Math. Soc. (3)*, 40(3):443–475, 1980.
146. D. A. Moskvin. A local limit theorem for large deviations in the case of differently distributed lattice summands. *Teor. Verojatnost. i Primenen.*, 17:716–722, 1972.
147. D. A. Moskvin, G. A. Freĭman, and A. A. Judin. Structural theory of set summation, and local limit theorems for independent lattice random variables. *Teor. Verojatnost. i Primenen.*, 19:52–62, 1974.
148. A. B. Mukhin. Local limit theorems for lattice random variables. *Teor. Veroyatnost. i Primenen.*, 36(4):660–674, 1991.
149. S. V. Nagaev. Some limit theorems for stationary Markov chains. *Teor. Veroyatnost. i Primenen.*, 2:389–416, 1957.
150. S. V. Nagaev. Large deviations of sums of independent random variables. *Ann. Probab.*, 7(5):745–789, 1979.
151. M. Nicol, A. Török, and S. Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dynam. Systems*, 38(3):1127–1153, 2018.
152. W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.
153. W. Parry and M. Pollicott. Skew products and Livsic theory. In *Representation theory, dynamical systems, and asymptotic combinatorics*, volume 217 of *Amer. Math. Soc. Transl. Ser. 2*, pages 139–165. Amer. Math. Soc., Providence, RI, 2006.
154. Y. Peres. Domains of analytic continuation for the top Lyapunov exponent. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(1):131–148, 1992.
155. V. V. Petrov. Generalization of Cramér's limit theorem. *Uspehi Matem. Nauk (N.S.)*, 9(4(62)):195–202, 1954.
156. V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.
157. W. Philipp and W. Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.*, 2(161,, 161):iv+140, 1975.
158. D. Plachky and J. Steinebach. A theorem about probabilities of large deviations with an application to queuing theory. *Period. Math. Hungar.*, 6(4):343–345, 1975.
159. M. Pollicott and R. Sharp. Asymptotic expansions for closed orbits in homology classes. *Geom. Dedicata*, 87(1-3):123–160, 2001.
160. M. Pollicott and R. Sharp. Chebotarev-type theorems in homology classes. *Proc. Amer. Math. Soc.*, 135(12):3887–3894, 2007.
161. Mark Pollicott. A complex Ruelle-Perron-Frobenius theorem and two counterexamples. *Ergodic Theory Dynam. Systems*, 4(1):135–146, 1984.
162. G. Pólya. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, 84(1-2):149–160, 1921.
163. Yu. V. Prokhorov. On a local limit theorem for lattice distributions. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:535–538, 1954.
164. F. Rassoul-Agha and T. Seppäläinen. *A course on large deviations with an introduction to Gibbs measures*, volume 162 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.
165. A. Raugi. Mesures invariantes ergodiques pour des produits gauches. *Bull. Soc. Math. France*, 135(2):247–258, 2007.
166. A. Rényi. Contributions to the theory of independent random variables. *Acta Math. Acad. Sci. Hungar.*, 1:99–108, 1950.
167. W. Richter. Lokale Grenzwertsätze für grosse Abweichungen. *Teor. Veroyatnost i Primenen.*, 2:214–229, 1957.
168. J. Rousseau-Egele. Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux. *Ann. Probab.*, 11(3):772–788, 1983.

169. Yu. A. Rozanov. On a local limit theorem for lattice distributions. *Teor. Veroyatnost. i Primenen.*, 2:275–281, 1957.
170. W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, 3d edition, 1987.
171. D. Ruelle. *Thermodynamic formalism. The mathematical structures of classical equilibrium statistical mechanics*, volume 5 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1978.
172. D. Ruelle. Analycity properties of the characteristic exponents of random matrix products. *Adv. in Math.*, 32(1):68–80, 1979.
173. H. H. Rugh. Cones and gauges in complex spaces: spectral gaps and complex Perron-Frobenius theory. *Ann. of Math. (2)*, 171(3):1707–1752, 2010.
174. E. L. Rvačeva. On domains of attraction of multidimensional distributions. *L′vov. Gos. Univ. Uč. Zap. Ser. Meh.-Mat.*, 29(6):5–44, 1954.
175. O. Sarig. Invariant Radon measures for horocycle flows on abelian covers. *Invent. Math.*, 157(3):519–551, 2004.
176. L. Saulis. Asymptotic expansion for probabilities with large deviations. *Litovsk. Mat. Sb.*, 9:605–625, 1969.
177. L. Saulis and V. A. Statulevičius. *Limit theorems for large deviations*, volume 73 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers, Dordrecht, 1991. Translated and revised from the 1989 Russian original.
178. K. Schmidt. *Cocycles on ergodic transformation groups*. Macmillan Company of India, Ltd., Delhi, 1977. Macmillan Lectures in Mathematics, Vol. 1.
179. E. Seneta. On the historical development of the theory of finite inhomogeneous Markov chains. *Proc. Cambridge Philos. Soc.*, 74:507–513, 1973.
180. T. Seppäläinen. Large deviations for Markov chains with random transitions. *Ann. Probab.*, 22(2):713–748, 1994.
181. S. Sethuraman and S. R. S. Varadhan. A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains. *Electron. J. Probab.*, 10:no. 36, 1221–1235, 2005.
182. R. Sharp. Local limit theorems for free groups. *Math. Ann.*, 321(4):889–904, 2001.
183. L. A. Shepp. A local limit theorem. *Ann. Math. Statist.*, 35:419–423, 1964.
184. Ya. G. Sinai. *Theory of phase transitions: rigorous results*, volume 108 of *International Series in Natural Philosophy*. Pergamon Press, Oxford-Elmsford, N.Y., 1982. Translated from the Russian by J. Fritz, A. Krámli, P. Major and D. Szász.
185. S. H. Siraždinov. *Limit theorems for stationary Markov chains*. Izdat. Akad. Nauk Uzbekskoĭ SSR, Taškent, 1955.
186. A. V. Skorokhod. *Studies in the theory of random processes*. Kiev Univ. Publ., 1961.
187. S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
188. V. A. Statulevičius. On large deviations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 6:133–144, 1966.
189. V. Statuljavičjus. Local limit theorems and asymptotic expansions for non-stationary Markov chains. *Litovsk. Mat. Sb.*, 1(1-2):231–314, 1961.
190. V. A. Statuljavičus. Limit theorems for densities and asymptotic expansions for distributions of sums of independent random variables. *Teor. Verojat. i Primene*, 10:645–659, 1965.
191. C. Stone. A local limit theorem for nonlattice multi-dimensional distribution functions. *Ann. Math. Statist.*, 36:546–551, 1965.
192. W. F. Stout. A martingale analogue of Kolmogorov's law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 15:279–290, 1970.
193. V. Strassen. An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 3:211–226 (1964), 1964.
194. V. Strassen. Almost sure behavior of sums of independent random variables and martingales. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1*, pages 315–343. Univ. California Press, Berkeley, Calif., 1967.
195. D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, 1979.
196. P. Tchebycheff. Sur deux théorèmes relatifs aux probabilités. *Acta Math.*, 14(1):305–315, 1890.
197. S. R. S. Varadhan. *Large deviations*, volume 27 of *Courant Lecture Notes in Mathematics*. American Mathematical Society, Providence, RI, 2016.
198. A. D. Wentzell. *Limit theorems on large deviations for Markov stochastic processes*, volume 38 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990. Translated from the Russian.
199. E. F. Whittlesey. Analytic functions in Banach spaces. *Proc. Amer. Math. Soc.*, 16:1077–1083, 1965.

Index