Siberian Mathematical Journal, Vol. 42, No. 2, pp. 301–316, 2001
Original Russian Text Copyright © 2001 *Korshunov D. A.*

LIMIT THEOREMS FOR GENERAL MARKOV CHAINS

D. A. Korshunov UDC 519.214

§ 1. Introduction

Let S be a measurable space with a σ-algebra $\mathscr{B}(S)$ of measurable sets. Let $P_n(x, B)$, $x \in S$, $B \in \mathscr{B}(S)$, be some transition probability on S; in this article, the parameter n, time, ranges in the set $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ of nonnegative integers. The transition probability is not assumed homogeneous in time n. Consider the Markov chain $X = \{X_n,\ n \in \mathbb{Z}^+\}$ with values in S and transition probability $P_n(\cdot, \cdot)$; i.e.,

$$\mathbf{P}\{X_{n+1} \in B \mid X_n = x\} = P_n(x, B).$$

In § 2 we find out conditions under which $f(X_n)/n$ converges almost surely to some limit, where $f : S \to \mathscr{Y}$ is a function with values in a separable Banach space $\mathscr{Y}$. We consider the so-called p-smooth Banach space. In § 3 we study the behavior in time of the characteristic functional of a Markov chain with values in an arbitrary separable Banach space. In § 4, for a Markov chain with values in a finite-dimensional Euclidean space, we state conditions under which X_n satisfies the central limit theorem. In § 5 we derive an upper estimate for the probability to belong to a compact set for an $\mathbb{R}^d$-valued Markov chain asymptotically homogeneous in time and space (in some direction). In § 6 we prove a local central limit theorem for an asymptotically homogeneous Markov chain with values in the integral lattice $\mathbb{Z}^d$, and in § 7 we prove an analog of this theorem for a nonlattice Markov chain with values in $\mathbb{R}^d$.

Although in the conditions of the theorems we do not indicate explicitly whether the Markov chain $\{X_n\}$ is positive recurrent, null recurrent, or transient, our results are most meaningful for nonrecurrent chains. Moreover, in some theorems we suppose explicitly that the value of the chain X_n "goes to infinity in some direction."

From the viewpoint of the strong law of large numbers and the central limit theorem, most publications are devoted to positive Harris recurrent (ergodic) time homogeneous Markov chains (see, for instance, [1, § 17]). Some results relating to the central limit theorem for time nonhomogeneous ergodic Markov chains can be found in [2, 3]. Observe that, unlike transient chains (that are the main subject of study in the present article), the most natural problem for ergodic chains is that of the asymptotic behavior of the distribution of sums of the values of a function $f : S \to \mathbb{R}$ of a Markov chain, i.e., the distribution of $f(X_1) + \dots + f(X_n)$. In this case, employment of the cyclic structure of an ergodic chain (cycles of return to an atom) eventually reduces the problem to the familiar limit theorems for sums of independent random variables.

§ 2. Assertions of the Type of the Strong Law of Large Numbers for a Function of a Markov Chain

2.1. The strong law of large numbers for martingales in Banach spaces. Given $p \in [1, 2]$, say that a Banach space $\mathscr{Y}$ with a norm $\|\cdot\|$ is *p-smooth* if there is a constant $D < \infty$ such that the following inequality holds for arbitrary vectors $x, y \in \mathscr{Y}$ with $\|x\| = 1$ and $\|y\| \le 1$:

$$\|x + y\| + \|x - y\| \le 2 + D\|y\|^p. \tag{1}$$

The research was financially supported by the Russian Foundation for Basic Research (Grants 99–01–00504 and 99–01–00561)) and INTAS (Grant 1625).

Novosibirsk. Translated from *Sibirskiĭ Matematicheskiĭ Zhurnal*, Vol. 42, No. 2, pp. 354–371, March–April, 2001. Original article submitted August 9, 2000.

0037-4466/01/4202–0301 $25.00 © 2001 Plenum Publishing Corporation

By this definition, a p-smooth Banach space $\mathscr{Y}$ is p_1-smooth for every $p_1 \in [1, p]$.

Every Banach space is 1-smooth for $D = 2$ (the triangle inequality). A Banach space $\mathscr{Y}$ is 2-smooth if and only if there is a constant $D < \infty$ such that

$$\|x+y\|^2 + \|x-y\|^2 \le 2\|x\|^2 + D\|y\|^2 \tag{2}$$

(see [4]). If $\mathscr{Y}$ is a Hilbert space then $\mathscr{Y}$ is automatically 2-smooth, since (2) transforms into equality with $D = 2$ (the parallelogram identity).

The following strong law of large numbers [5, Theorem 2.2; 6] is known for martingales in separable Banach spaces (separability is needed for the addition of random elements with values in this space to be a measurable operation).

Theorem 1. *Suppose that a sequence Z_n, $n = 1, 2, \dots$, of random elements with values in a separable Banach space $\mathscr{Y}$ constitutes a martingale with respect to some filtration on the main probability space. If, for some $p \in [1, 2]$, the Banach space $\mathscr{Y}$ is p-smooth and the series*

$$\sum_{n=1}^{\infty} \frac{\mathbf{E}\|X_{n+1} - X_n\|^p}{n^p}$$

converges then $X_n/n \to 0$ almost surely as $n \to \infty$.

2.2. A function of a Markov chain. Let $\mathscr{Y}$ be a separable Banach space whose norm is $\|\cdot\|$ and let $\mathscr{B}(\mathscr{Y})$ be the Borel σ-algebra. Denote by $\mathscr{L}im_{n\to\infty} y_n$ the set of limit points of a sequence $\{y_n\}$, i.e., the set of all $y \in \mathscr{Y}$ such that $y_{n_k} \to y$ for some subsequence with subscripts $n_k \to \infty$. The set of limit points is closed by necessity.

Let f be a measurable function from S into $\mathscr{Y}$. Denote by $\eta_n(x)$, $x \in S$, the random vector corresponding to the jump of the process $Y_n = f(X_n)$, i.e., the vector such that

$$\mathbf{P}\{\eta_n(x) \in B\} = \mathbf{P}\{f(X_{n+1}) - f(X_n) \in B \mid X_n = x\}$$

for every $B \in \mathscr{B}(\mathscr{Y})$. Put $m_n^X(x) = \mathbf{E}\eta_n(x)$. Here and in the sequel, by expectation we mean the Bochner integral.

Introduce the partial order relation $\le_{\mathrm{st}}$ on the set of random variables: given two random variables η_1 and η_1, write $\eta_1 \le_{\mathrm{st}} \eta_2$ if $\mathbf{P}\{\eta_1 \ge x\} \le \mathbf{P}\{\eta_2 \ge x\}$ for every $x \in \mathbb{R}$.

Theorem 2. *Suppose that a Banach space $\mathscr{Y}$ is p-smooth for some $p \in (1, 2]$ and let a set $\widetilde{B} \in \mathscr{B}(\mathscr{Y})$ be such that*

$$\mathbf{P}\{f(X_n) \in \widetilde{B} \text{ for all } n \ge N\} \to 1 \tag{3}$$

as $N \to \infty$. Suppose also that the following inclusion is valid for some time $\widetilde{N}$ and a closed convex set M in $\mathscr{B}(\mathscr{Y})$:

$$\{m_n^X(x) : \ n \ge \widetilde{N}, \ f(x) \in \widetilde{B}\} \subseteq M. \tag{4}$$

Moreover, suppose that the family $\{\|\eta_n(x)\|, n \ge \widetilde{N}, f(x) \in \widetilde{B}\}$ of random variables possesses an integrable majorant, i.e., a random variable η with a finite mean such that

$$\|\eta_n(x)\| \le_{\mathrm{st}} \eta \quad \text{for all } n \ge \widetilde{N} \quad \text{and} \quad f(x) \in \widetilde{B}. \tag{5}$$

Then almost surely

$$\mathscr{L}im_{n\to\infty} \frac{f(X_n)}{n} \subseteq M.$$

Proof resembles in some details the arguments in the proof of Lemma 1 of [7]. Without loss of generality we may assume that $\widetilde{N} = 0$. To begin, we additionally suppose that the random variable

η is a majorant not only for the family $\{\|\eta_n(x)\|,\ n \ge 0,\ f(x) \in \widetilde{B}\}$ but also for the whole family $\{\|\eta_n(x)\|,\ n \ge 0,\ x \in S\}$; i.e.,

$$\|\eta_n(x)\| \le_{\text{st}} \eta \quad \text{for all } n \ge 0 \quad \text{and} \quad x \in S. \tag{6}$$

First of all, observe that, in view of (3) and (4), the following inclusion holds almost surely:

$$\mathop{\mathscr{L}im}_{n\to\infty} \mathbf{E}\{f(X_{n+1}) - f(X_n) \mid X_n\} \equiv \mathop{\mathscr{L}im}_{n\to\infty}\ m_n^X(X_n) \subseteq M. \tag{7}$$

Given a number $c > 0$ and a point $x \in S$, put

$$B_x^{[c]} = \{u \in S : \|f(u) - f(x)\| < c\}.$$

Define the transition probability $P_n^{[c]}(x, B)$ and the random vector $\eta_n^{[c]}(x)$ by

$$P_n^{[c]}(x,B) \equiv \begin{cases} P_n\big(x, B \cap B_x^{[c]}\big) & \text{if } x \notin B, \\ P_n\big(x, B \cap B_x^{[c]}\big) + P_n\big(x, \{S \setminus B_x^{[c]}\}\big) & \text{if } x \in B, \end{cases}$$

and

$$\eta_n^{[c]}(x) \equiv \begin{cases} \eta_n(x) & \text{if } \|\eta_n(x))\| < c, \\ 0 & \text{if } \|\eta_n(x))\| \ge c. \end{cases}$$

By construction, the distribution of the random vector $\eta_n^{[c]}(x)$ coincides with that of the difference $f(Z) - f(x)$, where Z has distribution $P_n^{[c]}(x, \cdot)$.

Assume that $A > 0$. Consider the Markov chain Y_n, $Y_0 = X_0$, with transition probabilities $P_n^{[An]}(\cdot,\cdot)$. We can define the chains Y_n and X_n on a common probability space so that the probability of discrepancy between the trajectories of Y_n and X_n be at most

$$\mathbf{P}\{Y_n \ne X_n \text{ for some } n\} \le \sum_{n=0}^{\infty} \mathbf{P}\{\|f(X_{n+1}) - f(X_n)\| \ge An\} \le \sum_{n=1}^{\infty} \mathbf{P}\{\eta \ge An\} \le \frac{\mathbf{E}\eta}{A}. \tag{8}$$

Consequently,

$$\mathbf{P}\left\{ \mathop{\mathscr{L}im}_{n\to\infty} \frac{f(X_n)}{n} \ne \mathop{\mathscr{L}im}_{n\to\infty} \frac{f(Y_n)}{n} \right\} \le \frac{\mathbf{E}\eta}{A}. \tag{9}$$

Put $m_n^Y(Y_n) = \mathbf{E}\{f(Y_{n+1}) - f(Y_n) \mid Y_n\} \equiv \mathbf{E}\eta_n^{[An]}(x)\big|_{x=Y_n}$ and $\Delta_n = f(Y_{n+1}) - f(Y_n) - m_n^Y(Y_n)$, so that

$$f(Y_n) - f(Y_0) = \sum_{k=0}^{n-1} m_k^Y(Y_k) + \sum_{k=0}^{n-1} \Delta_k \equiv Z_n^0 + Z_n^1.$$

By condition (6) on the jumps $\eta_n(x)$, we have the estimate

$$\big\|m_k^Y(Y_k) - m_k^X(Y_k)\big\| \le \mathbf{E}\{\eta; \eta \ge Ak\}.$$

Therefore,

$$\left\| \frac{Z_n^0}{n} - \frac{1}{n}\sum_{k=0}^{n-1} m_k^X(Y_k) \right\| \le \frac{1}{n}\sum_{k=0}^{n-1} \big\|m_k^Y(Y_k) - m_k^X(Y_k)\big\| \le \frac{1}{n}\sum_{k=0}^{n-1} \mathbf{E}\{\eta; \eta \ge Ak\}.$$

Since $\mathbf{E}\eta$ is finite,

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{E}\{\eta; \eta \ge Ak\} \to 0$$

as $n \to \infty$, and consequently

$$\mathop{\mathscr{L}im}_{n\to\infty} \frac{Z_n^0}{n} = \mathop{\mathscr{L}im}_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} m_k^X(Y_k) \tag{10}$$

on all elementary events. From (8) we obtain the estimate

$$\mathbf{P}\left\{\mathop{\mathscr{L}im}_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} m_k^X(Y_k) \neq \mathop{\mathscr{L}im}_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} m_k^X(X_k)\right\} \leq \frac{\mathbf{E}\eta}{A}. \tag{11}$$

Since M is closed and convex, (7) implies that the following inclusion holds almost surely:

$$\mathop{\mathscr{L}im}_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} m_k^X(X_k) \subseteq M. \tag{12}$$

Relations (10)–(12) imply the estimate

$$\mathbf{P}\left\{\mathop{\mathscr{L}im}_{n\to\infty} \frac{Z_n^0}{n} \subseteq M\right\} \geq 1 - \frac{\mathbf{E}\eta}{A}. \tag{13}$$

Since the sequence Y_n is a Markov chain, by the definition of Δ_n we have

$$\mathbf{E}\{\Delta_n \mid Y_0, \dots, Y_n\} = \mathbf{E}\{\Delta_n \mid Y_n\} = 0.$$

Therefore, the process Z_n^1 constitutes a martingale with respect to the filtration $\sigma(Y_0, \dots, Y_{n-1})$. Prove that the increments of this martingale satisfy

$$\sum_{n=1}^{\infty} \frac{\mathbf{E}\|\Delta_n\|^p}{n^p} < \infty. \tag{14}$$

By the construction of Δ_n, the inequality $\|a+b\|^p \leq 2^p\|a\|^p + 2^p\|b\|^p$, and (6), for every x we have

$$\begin{aligned}\mathbf{E}\{\|\Delta_n\|^p \mid Y_n = x\} &= \mathbf{E}\big\|\eta_n^{[An]}(x) - \mathbf{E}\eta_n^{[An]}(x)\big\|^p \\ &\leq 2^p\mathbf{E}\big\|\eta_n^{[An]}(u)\big\|^p + 2^p\big\|\mathbf{E}\eta_n^{[An]}(u)\big\|^p \leq 2^{p+1}\mathbf{E}\{\eta^p; \eta < An\}.\end{aligned}$$

Therefore,

$$\sum_{n=1}^{\infty} \frac{\mathbf{E}\|\Delta_n\|^p}{n^p} \leq 8\sum_{n=1}^{\infty} \frac{\mathbf{E}\{\eta^p; \eta < An\}}{n^p}.$$

The last series converges for every value of A, since

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mathbf{E}\{\eta^p; \eta < An\}}{n^p} &= \sum_{n=1}^{\infty} \frac{A^p}{n^p}\mathbf{E}\{(\eta/A)^p; \eta/A < n\} \\ \leq \sum_{n=1}^{\infty} \frac{A^p}{n^p}\sum_{k=1}^{n} k^p\mathbf{P}\{k-1 \leq \eta/A < k\} &= A^p\sum_{k=1}^{\infty} k^p\mathbf{P}\{k-1 \leq \eta/A < k\}\sum_{n=k}^{\infty}\frac{1}{n^p} < \infty\end{aligned}$$

by the equivalence $\sum_{n=k}^{\infty} 1/n^p \sim k/k^p$ (for $p > 1$) and the existence of $\mathbf{E}\eta$.

Thus, the martingale Z_n^1 satisfies (14) and we can use Theorem 1 by which $Z_n^1/n \to 0$ almost surely as $n \to \infty$. Consequently, the following equality holds almost surely:

$$\mathop{\mathscr{L}im}_{n\to\infty} \frac{f(Y_n) - f(Y_0)}{n} \equiv \mathop{\mathscr{L}im}_{n\to\infty} \frac{Z_n^0 + Z_n^1}{n} = \mathop{\mathscr{L}im}_{n\to\infty} \frac{Z_n^0}{n}.$$

Now, under the additional condition (6), the assertion of the theorem ensues from (13) by the arbitrariness of A.

Now, we remove condition (6). Fix $\varepsilon > 0$. By (3), there is a time $N_1 \geq 1$ such that

$$\mathbf{P}\{f(X_n) \in \widetilde{B} \text{ for all } n \geq N_1\} \geq 1 - \varepsilon. \tag{15}$$

Consider an auxiliary time nonhomogeneous Markov chain $\widetilde{X}_n$ whose transition probabilities $\widetilde{P}_n(x, \cdot)$ coincide with $P_n(x, \cdot)$ for $f(x) \in \widetilde{B}$ and are equal to $\mathbf{I}\{x \in \cdot\}$ for $f(x) \notin \widetilde{B}$. In particular, $\widetilde{\eta}_n(x) = \eta_n(x)$ for $f(x) \in \widetilde{B}$ and $\widetilde{\eta}_n(x) = 0$ for $f(x) \notin \widetilde{B}$. Therefore, by (5) the increments $\widetilde{\eta}_n(x)$ satisfy (6). Consequently, by Theorem 2 the following inclusion is valid almost surely:

$$\mathop{\mathscr{L}im}_{n\to\infty} f(\widetilde{X}_n)/n \subseteq M.$$

Moreover, if we consider not all values of the Markov chain $\widetilde{X}_n$ but only $\{\widetilde{X}_n,\ n \geq N_1\}$ and suppose that the distribution of $\widetilde{X}_{N_1}$ coincides with that of X_{N_1} then from (15) we conclude that the trajectories of the chains X_n and $\widetilde{X}_n$ coincide for $n \geq N_1$ with a probability at least $1 - \varepsilon$. Therefore,

$$\mathbf{P}\{\mathop{\mathscr{L}im}_{n\to\infty} f(X_n)/n \subseteq M\} \geq 1 - \varepsilon.$$

Since $\varepsilon > 0$ was chosen arbitrarily, we arrive at the conclusion of the theorem.

2.3. The strong law of large numbers for a one-dimensional Markov chain with asymptotically homogeneous drift. In this section, we consider a Markov chain $\{X_n\}$ with values on the real axis $\mathbb{R}$. Denote by $\xi_n(x)$ the random variable whose distribution corresponds to the distribution of the jump of the chain $\{X_n\}$ from a state x at time n; i.e., $\mathbf{P}\{x + \xi_n(x) \in B\} = P_n(x, B)$, $B \in \mathscr{B}(\mathbb{R})$.

We say that the chain X_n with values in $\mathbb{R}$ is a *chain with asymptotically homogeneous drift* (*in time and space*) if $\mathbf{E}\xi_n(x)$ converges to some number $\mu \in \mathbb{R}$ as $n, x \to \infty$ (here we do not presume existence of $\mathbf{E}\xi_n(x)$ for all values of n and x).

Theorem 3. *Suppose that the Markov chain X_n has asymptotically homogeneous mean drift (in time and space) $\mu \geq 0$ and*

$$X_n \to \infty \quad \text{as } n \to \infty \tag{16}$$

almost surely. Suppose that, for some space level U and some time N, the family $\{|\xi_n(u)|,\ n \geq N,\ u \geq U\}$ of random variables possesses an integrable majorant; i.e., there is a random variable ξ with finite mean such that $|\xi_n(u)| \leq_{\rm st} \xi$ for arbitrary $n \geq N$ and $u \geq U$. Then $X_n/n \to \mu$ almost surely as $n \to \infty$.

Remark 1. The simplest examples of Markov chains satisfying the conditions of the theorem are: (a) the usual partial sum process $X_n = \xi_1 + \cdots + \xi_n$ of independent identically distributed random variables $\xi_1, \xi_2, \ldots$ with a positive mean and (b) a random walk $X_{n+1} = \max(0, X_n + \xi_n)$ with delay at zero.

Remark 2. For an irreducible countable Markov chain with values in $\mathbb{Z}^+$, condition (16) is equivalent to transience of the chain.

Proof of Theorem 3. Fix $\varepsilon > 0$. Since $\mathbf{E}\xi_n(x) \to \mu$, there exist $\widetilde{N} > N$ and $\widetilde{U} > U$ such that

$$\mathbf{E}\xi_n(x) \in [\mu - \varepsilon, \mu + \varepsilon] \quad \text{for } n \geq \widetilde{N} \text{ and } x \geq \widetilde{U}.$$

In view of (16),

$$\mathop{\mathscr{L}im}_{N_1\to\infty} \mathbf{P}\{X_n \geq \widetilde{U} \text{ for every } n \geq N_1\} \to 1.$$

Thus, the chain X_n satisfies the conditions of Theorem 2 for $S = \mathscr{Y} = \mathbb{R}$, $f(x) = x$, $\widetilde{B} = [\widetilde{U}, \infty)$, and $M = [\mu - \varepsilon, \mu + \varepsilon]$. Consequently, the following inclusion is valid almost surely:

$$\mathop{\mathscr{L}im}_{n\to\infty} X_n/n \subseteq [\mu - \varepsilon, \mu + \varepsilon].$$

Since $\varepsilon > 0$ is arbitrary, the theorem is proven.

§ 3. Time Behavior of the Characteristic Functional of a Markov Chain

The characteristic functional of the sum of independent variables equals the product of the characteristic functionals of the summands. In this section we clarify the extent to which this assertion remains valid for a Markov chain with values in a separable Banach space $\mathscr{Y}$.

3.1. A proximity estimate for the value of the characteristic functional of a chain and the product of the characteristic functionals of jumps.

Lemma 1. *Suppose that $\lambda : \mathscr{Y} \to \mathbb{R}$ is a linear functional. The following inequality is valid for all $n \geq 1$ and $k \leq n$ and a complex number $\varphi \in \mathbb{C}$, $|\varphi| \leq 1$ (here i is the imaginary unity):*

$$|\mathbf{E}e^{i\lambda(X_n)} - \varphi^{n-k}\mathbf{E}e^{i\lambda(X_k)}| \leq \sum_{j=k}^{n-1} \delta_j |\varphi|^{n-j-1},$$

where

$$\delta_j = \sup_{x\in\mathscr{Y}} |\mathbf{E}e^{i\lambda(\xi_j(x))} - \varphi|. \tag{17}$$

Proof. Take $j \in [k+1, n]$. Since λ is linear and $\{X_n\}$ is a Markov chain, we have

$$\begin{aligned}\mathbf{E}e^{i\lambda(X_j)} &= \mathbf{E}\{\mathbf{E}\{e^{i\lambda(X_j - X_{j-1})}e^{i\lambda(X_{j-1})}|X_{j-1}\}\} \\ &= \int_{\mathscr{Y}} (\mathbf{E}e^{i\lambda(\xi_{j-1}(x))})e^{i\lambda(x)}\mathbf{P}\{X_{j-1}{\in}dx\}.\end{aligned}$$

Consequently,

$$|\mathbf{E}e^{i\lambda(X_j)} - \varphi\mathbf{E}e^{i\lambda(X_{j-1})}| = \left|\int_{\mathscr{Y}} (\mathbf{E}e^{i\lambda(\xi_{j-1}(x))} - \varphi)e^{i\lambda(x)}\mathbf{P}\{X_{j-1}{\in}dx\}\right| \leq \delta_{j-1}$$

in view of (17). Hence, we derive the inequality

$$\begin{aligned}&|\mathbf{E}e^{i\lambda(X_n)} - \varphi^{n-k}\mathbf{E}e^{i\lambda(X_k)}| \\ &\leq \sum_{j=k+1}^{n} |\varphi^{n-j}\mathbf{E}e^{i\lambda(X_j)} - \varphi^{n-(j-1)}\mathbf{E}e^{i\lambda(X_{j-1})}| \leq \sum_{j=k+1}^{n} \delta_{j-1}|\varphi|^{n-j},\end{aligned}$$

completing the proof of the lemma.

3.2. A proximity estimate in terms of high probability sets. We defined δ_j in (17) as the maximal deviation of the value of the characteristic functional of a jump of a chain from some complex number $\varphi \in \mathbb{C}$ over all phase space. In the lemma below, we define δ_j as the maximal deviation of the value of the characteristic functional of a jump of the chain from φ on some set rather than the whole phase space. In applications of this lemma in the next sections the corresponding sets have probability close to 1.

Suppose that $B_0, B_1, \dots$ are some sets in $\mathscr{Y}$. Given $k \leq n$, consider the event

$$B_{k,n} = \{X_j \in B_j \text{ for every } j \in [k, n]\}.$$

The events $B_{k,n}$ constitute a nonincreasing sequence in n.

Lemma 2. *Suppose that $\lambda : \mathscr{Y} \to \mathbb{R}$ is a linear functional. The following inequality is valid for all $n \geq 1$ and $k \leq n$ and a complex number $\varphi \in \mathbb{C}$, $|\varphi| \leq 1$:*

$$|\mathbf{E}e^{i\lambda(X_n)} - \varphi^{n-k}\mathbf{E}e^{i\lambda(X_k)}| \leq \sum_{j=k}^{n-1} \delta_j |\varphi|^{n-j-1} + 2(1 - \mathbf{P}\{B_{k,n-1}\}),$$

where

$$\delta_j = \sup_{x \in B_j} |\mathbf{E}e^{i\lambda(\xi_j(x))} - \varphi|. \tag{18}$$

PROOF. Without loss of generality we may assume that $k = 0$. Given n, take an arbitrary point x_n in B_n. Define the auxiliary chain $\widetilde{X}_n$ with jumps $\tilde{\xi}_n(x)$, by putting $\tilde{\xi}_n(x) = \xi_n(x)$ for $x \in B_n$ and $\tilde{\xi}_n(x) = \xi_n(x_n)$ for $x \notin B_n$. By construction and in view of (18), for the chain $\widetilde{X}_n$ we have

$$\delta_j = \sup_{x \in \mathscr{Y}} |\mathbf{E}e^{i\lambda(\tilde{\xi}_j(x))} - \varphi|.$$

Therefore, from Lemma 1 we obtain the estimate

$$|\mathbf{E}e^{i\lambda(\widetilde{X}_n)} - \varphi^n \mathbf{E}e^{i\lambda(\widetilde{X}_0)}| \leq \sum_{j=0}^{n-1} \delta_j |\varphi|^{n-j-1}.$$

Put $\widetilde{X}_0 = X_0$. Then, due to the Markov property and coincidence of the jumps of two chains on the event $B_{0,n-1}$, the values of X_n and $\widetilde{X}_n$ coincide with probability at least $\mathbf{P}\{B_{0,n-1}\}$. Hence,

$$|\mathbf{E}e^{i\lambda(X_n)} - \mathbf{E}e^{i\lambda(\widetilde{X}_n)}| \leq 2(1 - \mathbf{P}\{B_{0,n-1}\}),$$

which completes the proof of the estimate of the lemma.

§ 4. A Central Limit Theorem for a Markov Chain with Values in Euclidean Space

In this section we study a Markov chain in the Euclidean space $\mathbb{R}^d$. The inner product of two row vectors ξ and $\eta \in \mathbb{R}^d$ is denoted by $\langle \xi, \eta \rangle$. By a column vector ξ^T we mean a transposed row vector ξ.

4.1. A central limit theorem. In the following theorem we find out sufficient conditions under which a Markov chain with values in $\mathbb{R}^d$ satisfies the central limit theorem.

Theorem 4. *Suppose that a nonincreasing sequence $B_0 \supseteq B_1 \supseteq \dots$ of sets in $\mathbb{R}^d$ is such that*

$$\mathbf{P}\{X_n \in B_n \text{ for every } n \geq N\} \to 1 \tag{19}$$

as $N \to \infty$. Suppose that, for some time $\widetilde{N}$, the family $\{\|\xi_n(x)\|^2,\ n \geq \widetilde{N},\ x \in B_0\}$ is integrable uniformly in n and x. If the relations

$$\sup_{x \in B_n} \|\mathbf{E}\xi_n(x) - \mu\| = o(1/\sqrt{n}), \tag{20}$$

$$\sup_{x \in B_n} \|\mathbf{Cov}(\xi_n(x), \xi_n(x)) - \sigma^2\| \to 0 \tag{21}$$

as $n \to \infty$ hold for some vector $\mu \in \mathbb{R}^d$ and a nonnegative definite symmetric $(d \times d)$-matrix σ^2 then the distribution of the random vector $n^{-1/2}(X_n - n\mu)$ converges weakly as $n \to \infty$ to the d-dimensional normal law with mean zero and covariance matrix σ^2.

PROOF. We use the method of characteristic functions. Henceforth $\lambda \in \mathbb{R}^d$. In view of the condition of uniform integrability of the family of the squares of the jumps, we have the expansion

$$\mathbf{E}e^{i\langle \lambda, \xi_j(x) - \mu \rangle} = 1 + i\langle \lambda, \mathbf{E}\xi_j(x) - \mu \rangle - \frac{1}{2}\lambda \mathbf{E}(\xi_j(x) - \mu)^T(\xi_j(x) - \mu)\lambda^T + o(\|\lambda\|^2)$$

as $\lambda \to 0$ uniformly in $j \ge \widetilde{N}$ and $x \in B_0$. Recalling (20) and (21), we obtain the relation

$$\mathbf{E}e^{i\langle \lambda, \xi_j(x)-\mu\rangle} = 1 - \lambda\sigma^2\lambda^T/2 + \varepsilon_j(\lambda, x)(\|\lambda\|/\sqrt{j} + \|\lambda\|^2),$$

where $\varepsilon_j(\lambda, x) \to 0$ as $\lambda \to 0$ and $j \to \infty$ uniformly in $x \in B_j$. Fix arbitrary $\lambda \in \mathbb{R}^d$ and $\varepsilon > 0$. By the last relation and (19), there is $k \ge \widetilde{N}$ such that for every $j \ge k$

$$|\mathbf{E}e^{i\langle \lambda/\sqrt{n}, \xi_j(x)-\mu\rangle} - (1 - \lambda\sigma^2\lambda^T/2n)| \le \varepsilon/\sqrt{nj}$$

uniformly in $x \in B_n$, and

$$\mathbf{P}\{X_j \in B_j \text{ for every } j \ge k\} \ge 1 - \varepsilon.$$

Applying Lemma 2 with $\varphi = 1 - \lambda\sigma^2\lambda^T/2n$ to the chain $\frac{X_n - n\mu}{\sqrt{n}}$, we now obtain the estimate

$$\left|\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle} - \left(1 - \frac{\lambda\sigma^2\lambda^T}{2n}\right)^{n-k} \mathbf{E}e^{i\langle \lambda, \frac{X_k - k\mu}{\sqrt{n}}\rangle}\right| \le \frac{\varepsilon}{\sqrt{n}} \sum_{j=k}^{n-1} \frac{1}{\sqrt{j}} + 2\varepsilon \le 3\varepsilon.$$

Since for every fixed k

$$\mathbf{E}e^{i\langle \lambda, \frac{X_k - k\mu}{\sqrt{n}}\rangle} \to 1$$

and

$$\left(1 - \frac{\lambda\sigma^2\lambda^T}{2n}\right)^{n-k} \to e^{-\lambda\sigma^2\lambda^T/2}$$

as $n \to \infty$, we derive the estimate

$$\limsup_{n\to\infty} |\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle} - e^{-\lambda\sigma^2\lambda^T/2}| \le 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the theorem is proven.

4.2. A central limit theorem for a one-dimensional Markov chain with asymptotically homogeneous drift. In this section we specify the result of Theorem 4 for a Markov chain $\{X_n\}$ with values on the real axis $\mathbb{R}$.

Theorem 5. *Suppose that a chain X has asymptotically homogeneous mean drift (in time and space) $\mu > 0$ and that (16) is satisfied. Suppose that the family $\{\xi_n^2(x),\ n \ge \widetilde{N},\ x \ge \widetilde{U}\}$ of the squares of the jumps is integrable uniformly in n and x for some time $\widetilde{N}$ and some space level $\widetilde{U}$. If*

$$\mathbf{E}\xi_n(x) = \mu + o(1/\sqrt{n} + 1/\sqrt{x}), \tag{22}$$

$$\mathbf{Var}\,\xi_n(x) \to \sigma^2 > 0 \tag{23}$$

as $n,\ x \to \infty$ then the distribution of the random variable $(X_n - n\mu)/\sqrt{n\sigma^2}$ converges weakly to the standard normal law as $n \to \infty$.

Proof. Since the family $\{\xi_n^2(x), n \ge \widetilde{N}, x \ge \widetilde{U}\}$ of the squares of the jumps is uniformly integrable, the family $\{|\xi_n(x)|,\ n \ge \widetilde{N},\ x \ge \widetilde{U}\}$ of random variables possesses an integrable majorant and the chain X_n satisfies the conditions of Theorem 3. By Theorem 3, $X_n/n \to \mu$ as $n \to \infty$. Consequently, the sets $B_n = [n\mu/2, \infty)$ satisfy the condition

$$\mathbf{P}\{X_n \in B_n \text{ for every } n \ge N\} \to 1$$

as $N \to \infty$. In view of (22) and (23), relations (20) and (23) are valid for the sets B_n. Application of Theorem 4 completes the proof.

§ 5. A Local Estimate for the Distribution of a Markov Chain with Values in Euclidean Space

5.1. An upper estimate for the probability for a Markov chain to belong to a compact set. Suppose that random variables ξ_n, $n = 1, 2, \dots$, with values in $\mathbb{R}$ are independent and identically distributed. The following estimate is well known for the concentration function of the distribution of the sum $S_n = \xi_1 + \dots + \xi_n$ (see, for instance, Theorem 9 of [8, Chapter III]): there is a constant c depending only on the distribution of ξ_1 such that, for arbitrary $x \in \mathbb{R}$ and $n \geq 1$,

$$\mathbf{P}\{S_n \in [x, x+1)\} \leq c/\sqrt{n}.$$

In the theorem below, we generalize this assertion to Markov chains with values in the Euclidean space $\mathbb{R}^d$. As above, $\xi_n(x)$ stands for the jumps of the chain at time n from a state x. Denote the characteristic function of the jump $\xi_n(x)$ by $\varphi_n(\lambda, x)$, $\lambda \in \mathbb{R}^d$.

Let $\varphi(\lambda)$ be the characteristic function of some random vector $\xi \in \mathbb{R}^d$ with nondegenerate distribution. Nondegeneracy means that the distribution of ξ is concentrated in none hyperplane; in other words, the (nonnegative definite symmetric) covariance matrix of the distribution F_r, $F_r(B) = \{\xi \in B \mid \|\xi\| < r\}$, is nondegenerate for at least one value $r > 0$ (and hence for all sufficiently large r).

Lemma 3. *For every random vector ξ with nondegenerate distribution, there is a positive number $\delta > 0$ such that $|\mathbf{E}e^{i\langle\lambda,\xi\rangle}| \leq e^{-\delta\|\lambda\|^2}$ for $\|\lambda\| \leq \delta$.*

Proof. Suppose that r is such that the covariance matrix σ^2 of the distribution F_r is nondegenerate. Then there is $\delta_1 > 0$ such that $\langle\lambda\sigma^2, \lambda\rangle \geq \delta_1\|\lambda\|^2$ for every $\lambda \in \mathbb{R}^d$. Denote by μ the mean of the distribution F_r. Since

$$|\mathbf{E}\{e^{i\langle\lambda,\xi\rangle} \mid \|\xi\| < r\}| = |\mathbf{E}\{e^{i\langle\lambda,\xi-\mu\rangle} \mid \|\xi\| < r\}| = 1 - \langle\lambda\sigma^2, \lambda\rangle/2 + o(\|\lambda\|^2)$$

as $\lambda \to 0$, there exists $\delta > 0$ such that for $\|\lambda\| < \delta$

$$|\mathbf{E}\{e^{i\langle\lambda,\xi\rangle} \mid \|\xi\| < r\}| \leq 1 - \delta\|\lambda\|^2.$$

Consequently, the following inequality is valid for $\|\lambda\| < \delta$:

$$\begin{aligned}|\mathbf{E}e^{i\langle\lambda,\xi\rangle}| &= |\mathbf{E}\{e^{i\langle\lambda,\xi\rangle} | \|\xi\| < r\}\mathbf{P}\{\|\xi\| < r\} + \mathbf{E}\{e^{i\langle\lambda,\xi\rangle} | \|\xi\| \geq r\}\mathbf{P}\{\|\xi\| \geq r\}| \\ &\leq |\mathbf{E}\{e^{i\langle\lambda,\xi\rangle} | \|\xi\| < r\}|\mathbf{P}\{\|\xi\| < r\} + \mathbf{P}\{\|\xi\| \geq r\} \leq 1 - \delta\|\lambda\|^2\mathbf{P}\{\|\xi\| < r\}.\end{aligned}$$

Now, the assertion of the lemma ensues from the inequality $1 + h \leq e^h$ which is valid for every real h.

Denote the unit cube with the "left lower" vertex $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ by $\square(x)$:

$$\square(x) = \{y = (y_1, \dots, y_d) : y_j \in [x_j, x_j + 1] \ \text{ for every } j = 1, \dots, d\}.$$

Theorem 6. *Suppose that the following relations hold for some nonincreasing sequence $B_0 \supseteq B_1 \supseteq \dots$ and a number $\varepsilon > 0$:*

$$\mathbf{P}\{X_k \notin B_k \text{ for some } k \geq n\} = O(n^{-d/2}) \tag{24}$$

as $n \to \infty$ and

$$\sup_{x \in B_n, \|\lambda\| \leq \varepsilon} |\varphi_n(\lambda, x) - \varphi(\lambda)| = O(\delta_n), \tag{25}$$

where

$$\delta_n = \begin{cases} n^{-1} & \text{for } d = 1, \\ (n \log n)^{-1} & \text{for } d = 2, \\ n^{-d/2} & \text{for } d \geq 3. \end{cases} \tag{26}$$

Then there is a constant c_1 such that the inequality

$$\mathbf{P}\{X_n \in \square(x)\} \le c_1 n^{-d/2}$$

holds uniformly in n and $x \in \mathbb{R}^d$.

REMARK 3. Conditions sufficient for validity of (24) with $d = 1$ for sets $B_n = [na, \infty)$, $a \in \mathbb{R}$, are given in the end of the current section.

PROOF. Without loss of generality we may assume that the number $\delta > 0$ of Lemma 3 equals ε. By this lemma, the following estimate holds for every $j \ge 1$:

$$\int\limits_{[-\varepsilon,\varepsilon]^d} |\varphi(\lambda)|^j \, d\lambda \le \int\limits_{\mathbb{R}^d} e^{-j\varepsilon\|\lambda\|^2} d\lambda = \left(\frac{2\pi}{j\varepsilon}\right)^{d/2}. \tag{27}$$

Estimate the value of the characteristic function $\mathbf{E}e^{i\langle\lambda,X_n\rangle}$ in a neighborhood of zero. Apply Lemma 2 to the chain X_n for $\varphi = \varphi(\lambda)$ and $k = n/2$. Using (24) and (25), we conclude that there exists $c_2 < \infty$ such that the following estimate holds uniformly in $\|\lambda\| \le \varepsilon$:

$$|\mathbf{E}e^{i\langle\lambda,X_n\rangle}| \le |\varphi(\lambda)|^{n/2} + c_2\delta_n \sum_{j=1}^{n/2} |\varphi(\lambda)|^j + O(n^{-d/2}). \tag{28}$$

Consider the random vector $\eta = (\eta_1, \dots, \eta_d)$ with independent coordinates each of which has a common distribution with the density (see [9, Chapter XVI, § 3])

$$p(z) = (1 - \cos z)/\pi z^2, \quad z \in \mathbb{R}, \tag{29}$$

and the characteristic function, $\lambda_1 \in \mathbb{R}$,

$$\psi(\lambda_1) = \mathbf{E}e^{i\lambda_1\eta_1} = \begin{cases} 1 - |\lambda_1| & \text{if } |\lambda_1| \le 1, \\ 0 & \text{if } |\lambda_1| > 1. \end{cases} \tag{30}$$

We suppose that η is independent of X_n. The characteristic function of the sum $X_n + \eta/\varepsilon$ equals $\mathbf{E}e^{i\langle\lambda,X_n\rangle}\psi(\lambda/\varepsilon)$, where $\psi(\lambda) = \psi(\lambda_1)\dots\psi(\lambda_d)$, $\lambda = (\lambda_1, \dots, \lambda_d)$. Since the characteristic function $|\psi(\lambda)|$ is integrable, the sum $X_n + \eta/\varepsilon$ has the bounded continuous density $p_{X_n+\eta/\varepsilon}(z)$ which is reconstructed by means of the following inversion formula (see, for instance, [9; Chapter XV § 3, § 7]):

$$p_{X_n+\eta/\varepsilon}(z) = \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d} e^{-i\langle\lambda,z\rangle} \mathbf{E}e^{i\langle\lambda,X_n\rangle} \psi(\lambda/\varepsilon) \, d\lambda, \quad z \in \mathbb{R}^d.$$

Consequently,

$$p_{X_n+\eta/\varepsilon}(z) \le \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d} |\mathbf{E}e^{i\langle\lambda,X_n\rangle}| \psi(\lambda/\varepsilon) \, d\lambda \le \frac{1}{(2\pi)^d} \int\limits_{[-\varepsilon,\varepsilon]^d} |\mathbf{E}e^{i\langle\lambda,X_n\rangle}| \, d\lambda$$

in view of the definition of $\psi(\lambda)$. Using (28) and (27), we arrive at the inequalities

$$p_{X_n+\eta/\varepsilon}(z) \le \left(\frac{2}{n\delta}\right)^{d/2} + c_2\delta_n \sum_{j=1}^{n/2} \left(\frac{1}{j\delta}\right)^{d/2} + O(n^{-d/2}).$$

The sum in the second summand on the right-hand side of the inequality is $O(\sqrt{n})$ for $d = 1$, $O(\log n)$ for $d = 2$, and $O(1)$ for $d \ge 3$. Therefore, there is c_3 such that

$$p_{X_n+\eta/\varepsilon}(z) \le c_3 n^{-d/2},$$

and

$$\mathbf{P}\{X_n + \eta \in \Box_u(x)\} \le c_3(1+2u)^d n^{-d/2} \tag{31}$$

for every u-neighborhood $\Box_u(x)$ of the cube $\Box(x)$. Take $u > 0$ so that $\mathbf{P}\{\|\eta/\varepsilon\| \le u\} \ge 1/2$. Then the following inequalities are valid in view of independence of X_n and η:

$$\begin{aligned}\mathbf{P}\{X_n + \eta/\varepsilon \in \Box_u(x)\} &\ge \mathbf{P}\{X_n \in \Box(x),\ \|\eta/\varepsilon\| \le u\} \\ &= \mathbf{P}\{X_n \in \Box(x)\}\mathbf{P}\{\|\eta/\varepsilon\| \le u\} \ge \mathbf{P}\{X_n \in \Box(x)\}/2.\end{aligned}$$

The last inequality together with (31) yields the estimate of the lemma.

5.2. Sufficient conditions for validity of (24) in the one-dimensional case. Suppose that $d = 1$, the left tail of the initial distribution of the chain satisfies the condition

$$\mathbf{P}\{X_0 \le -x\} = O(1/\sqrt{x}) \quad \text{as } x \to \infty,$$

and there is a random variable η with mean $m = a + \delta$, $\delta > 0$, and finite variance such that the inequality

$$\xi_n(x) \ge_{\mathrm{st}} \eta$$

holds for all values of n and x. Then

$$\mathbf{P}\{X_k < na \text{ for some } k \ge n\} = O(1/\sqrt{n})$$

as $n \to \infty$.

Indeed, let η_n, $n \in \mathbf{Z}^+$, be independent copies of η, and put $S_n = \eta_0 + \dots + \eta_{n-1}$. Define the chain X and η_n, $n \in \mathbf{Z}^+$, on the same probability space so that $\xi_n(x) \ge \eta_n$ almost surely for all n and x. Then we have $X_k \ge S_k - k\delta/2$ on the event $X_0 \ge -n\delta/2$ for every $k \ge n$. Therefore,

$$\begin{gathered}\mathbf{P}\{X_k < ka \text{ for some } k \ge n\} \\ \le \mathbf{P}\{X_0 < -n\delta/2\} + \mathbf{P}\{S_k - k\delta/2 < ka \text{ for some } k \ge n\} \\ = O(1/\sqrt{n}) + \mathbf{P}\{(S_k - km)/k < -\delta/2 \text{ for some } k \ge n\}.\end{gathered}$$

The sequence $(S_k - km)/k$, $k = 1, 2, \dots$, constitutes a reverse martingale and from Kolmogorov's inequality for martingales we obtain

$$\begin{gathered}\mathbf{P}\left\{\frac{S_k - km}{k} < -\frac{\delta}{2} \text{ for some } k \ge n\right\} \\ \le \mathbf{P}\left\{\sup_{k \ge n} \frac{|S_k - km|}{k} > \frac{\delta}{2}\right\} \le \frac{4\mathbf{E}(S_n - nm)^2}{\delta^2 n^2} = O(1/n) = O(1/\sqrt{n}),\end{gathered}$$

as required.

§ 6. A Local Central Limit Theorem in the Lattice Case

Suppose that $B_0 \supseteq B_1 \supseteq \dots$ is some nonincreasing sequence of sets in the Euclidean space $\mathbb{R}^d$. In this and next sections we consider a *Markov chain* with values in $\mathbb{R}^d$ *asymptotically homogeneous in time and space* (*in the direction of* B_n), i.e., a chain X_n such that the distribution of the jump $\xi_n(x)$ converges weakly to the distribution of some random variable ξ as $n \to \infty$ uniformly in $x \in B_n$.

We suppose that the "limit jump" ξ has finite mean $\mu \in \mathbb{R}^d$ and covariance matrix $\sigma^2 > 0$ of size d. We assume that the distribution of ξ is concentrated in none hyperplane; i.e., the nonnegative definite symmetric matrix σ^2 is nondegenerate. Denote by Q the inverse matrix of σ^2.

Suppose that the independent identically distributed random variables ξ_n, $n = 1, 2, \dots$, with values in $\mathbb{Z}$ have finite mean μ and variance $\sigma^2 = \mathbf{Var}\,\xi_1$. Suppose that the greatest common divisor of the numbers in the set $\{k \in \mathbb{Z} : \mathbf{P}\{\xi_1 = k\} > 0\}$ is equal to one; this means that $\mathbb{Z}$ is a minimal lattice for the distribution of ξ_1. The well-known local theorem for the distribution of the sum $S_n = \xi_1 + \cdots + \xi_n$ (see, for instance, Theorem 3 of [9, Chapter XV, § 5]) claims that

$$\mathbf{P}\{S_n = k\} = \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-(k-n\mu)^2/2n\sigma^2} + o\left(\frac{1}{\sqrt{n}}\right)$$

as $n \to \infty$ uniformly in all $k \in \mathbb{Z}$. In the theorem stated below, we generalize this assertion to a Markov chain X with values on the integer lattice $\mathbb{Z}^d$.

Denote $\varphi_n(\lambda, x) \equiv \mathbf{E}e^{i\langle\lambda,\xi_n(x)-\mu\rangle}$ and $\varphi(\lambda) \equiv \mathbf{E}e^{i\langle\lambda,\xi-\mu\rangle}$. In this section we suppose that ξ is a lattice distribution and $\mathbb{Z}^d$ is a minimal lattice in the sense that for every $\varepsilon > 0$

$$\sup_{\lambda\in[-\pi,\pi]^d\setminus[-\varepsilon,\varepsilon]^d} |\varphi(\lambda)| < 1. \tag{32}$$

In the one-dimensional case $d = 1$ the last relation is equivalent to the fact that the lattice $\mathbb{Z}$ is minimal for the distribution of ξ.

Theorem 7. *Suppose that*

$$(X_n - n\mu)n^{-1/2} \Rightarrow N(0, \sigma^2) \tag{33}$$

as $n \to \infty$. Moreover, suppose that the following two relations hold:

$$\mathbf{P}\{X_k \notin B_k \text{ for some } k \ge n\} = o(n^{-d/2}) \tag{34}$$

as $n \to \infty$ and

$$\sup_{x\in B_n,\ \lambda\in[-\pi,\pi]^d} |\varphi_n(\lambda, x) - \varphi(\lambda)| = o(\delta_n), \tag{35}$$

where δ_n is defined by (26). *Then the relation*

$$\mathbf{P}\{X_n = k\} = \frac{\sqrt{\det Q}}{(2\pi n)^{d/2}} e^{-\langle(k-n\mu)Q,k-n\mu\rangle/2n} + o(n^{-d/2})$$

holds uniformly in all $k \in \mathbb{Z}^d$.

Remark 4. Sufficient conditions for the weak convergence (3) are given in Theorem 5.

Proof. The following inversion formula is valid:

$$\mathbf{P}\{X_n = k\} = \frac{1}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} e^{-i\langle\lambda,k\rangle}\mathbf{E}e^{i\langle\lambda,X_n\rangle}\, d\lambda.$$

Consequently,

$$n^{d/2}\mathbf{P}\{X_n = k\} = \frac{1}{(2\pi)^d} \int\limits_{[-\pi\sqrt{n},\pi\sqrt{n}]^d} e^{-i\langle\lambda,k-n\mu\rangle/\sqrt{n}}\mathbf{E}e^{i\langle\lambda,\frac{X_n-n\mu}{\sqrt{n}}\rangle}\, d\lambda.$$

Recalling also that

$$\frac{\sqrt{\det Q}}{(2\pi)^{d/2}} e^{-\langle zQ,z\rangle/2} = \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d} e^{-i\langle\lambda,z\rangle-\langle\lambda\sigma^2,\lambda\rangle/2}\, d\lambda \tag{36}$$

for every $z \in \mathbb{R}^d$, we obtain the following equality and inequality for $z = (k - n\mu)/\sqrt{n}$ and every $A > 0$:

$$\begin{aligned}
&\left| n^{d/2}\mathbf{P}\{X_n = k\} - \frac{\sqrt{\det Q}}{(2\pi)^{d/2}} e^{-\langle (k-n\mu)Q, k-n\mu\rangle/2n} \right| \\
&= \frac{1}{(2\pi)^d} \Bigg| \int\limits_{[-\pi\sqrt{n},\pi\sqrt{n}]^d} e^{-i\langle \lambda, k-n\mu\rangle/\sqrt{n}} (\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle} - e^{-\langle \lambda\sigma^2, \lambda\rangle/2})\, d\lambda \\
&\quad - \int\limits_{\mathbb{R}^d \setminus [-\pi\sqrt{n},\pi\sqrt{n}]^d} e^{-i\langle \lambda, k-n\mu\rangle/\sqrt{n} - \langle \lambda\sigma^2, \lambda\rangle/2}\, d\lambda \Bigg| \\
&\le \frac{1}{(2\pi)^d} \int\limits_{[-A,A]^d} |\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle} - e^{-\langle \lambda\sigma^2, \lambda\rangle/2}|\, d\lambda \\
&\quad + \frac{1}{(2\pi)^d} \int\limits_{[-\pi\sqrt{n},\pi\sqrt{n}]^d \setminus [-A,A]^d} |\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle}|\, d\lambda \\
&\quad + \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d \setminus [-A,A]^d} e^{-\langle \lambda\sigma^2, \lambda\rangle/2}\, d\lambda \equiv I_1 + I_2 + I_3.
\end{aligned}$$

In view of the weak convergence (33) the characteristic function $\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle}$ converges to the characteristic function $e^{-\langle \lambda\sigma^2, \lambda\rangle/2}$ of the normal law as $n \to \infty$ uniformly in λ on every compact set. Therefore, $I_1 \to 0$ for every fixed $A > 0$ as $n \to \infty$. Moreover, $I_3 \to 0$ as $A \to \infty$, since the matrix σ^2 is strictly positive definite. To prove the theorem, it therefore suffices to demonstrate that

$$I_2 \to 0 \quad \text{as } n, A \to \infty. \tag{37}$$

Apply Lemma 2 to the chain X_n for $k = n/2$ and $\varphi = \varphi(\lambda/\sqrt{n})$. By (34) and (35), we have the estimate

$$|\mathbf{E}e^{i\langle \lambda, \frac{X_n - n\mu}{\sqrt{n}}\rangle}| \le |\varphi(\lambda/\sqrt{n})|^{n/2} + o(\delta_n) \sum_{j=1}^{n/2} |\varphi(\lambda/\sqrt{n})|^j + o(n^{-d/2}) \tag{38}$$

as $n \to \infty$ uniformly in $\lambda \in [-\pi\sqrt{n}, \pi\sqrt{n}]^d$. By (32) and Lemma 3, there is a number $\delta > 0$ such that for every $\lambda \in [-\pi, \pi]^d$

$$|\varphi(\lambda)| \le e^{-\delta\|\lambda\|^2}.$$

Consequently, the following estimate is valid for every $j \ge 1$:

$$\begin{aligned}
\int\limits_{[-\pi\sqrt{n},\pi\sqrt{n}]^d \setminus [-A,A]^d} |\varphi(\lambda/\sqrt{n})|^j\, d\lambda &\le \int\limits_{\mathbb{R}^d \setminus [-A,A]^d} e^{-j\delta\|\lambda\|^2/n}\, d\lambda \\
&= \left(\frac{n}{j\delta}\right)^{d/2} \int\limits_{\mathbb{R}^d \setminus [-A\sqrt{k\delta/n}, A\sqrt{k\delta/n}]^d} e^{-\|\lambda\|^2}\, d\lambda.
\end{aligned}$$

Inserting this estimate in (38), we find that I_2 does not exceed

$$\frac{1}{(2\pi)^d} \left(\left(\frac{2}{\delta}\right)^{d/2} \int\limits_{\mathbb{R}^d \setminus [-A\sqrt{\delta/2}, A\sqrt{\delta/2}]^d} e^{-\|\lambda\|^2}\, d\lambda + o(\delta_n) \sum_{j=1}^{n/2} \left(\frac{n}{j\delta}\right)^{d/2} \right) + o(1).$$

The last value vanishes as $n, A \to \infty$, since the sum in the second summand is $O(n)$ for $d = 1$, $O(n \log n)$ for $d = 2$, and $O(n^{d/2})$ for $d \ge 3$. We have proven (37) and hence the theorem.

§ 7. A Local Central Limit Theorem in the Nonlattice Case

As in the preceding section, here we consider a Markov chain X asymptotically homogeneous in time and space (again in the direction of the sets B_n), but now the distribution of the limit random variable ξ is supposed to be nonlattice. Moreover, the distribution of the inner product $\langle\lambda,\xi\rangle$ is assumed to be nonlattice for any $\lambda\in\mathbb{R}^d$. Therefore, for every $0<\varepsilon<a$ we have the inequality

$$\sup_{\lambda\in[-a,a]^d\setminus[-\varepsilon,\varepsilon]^d}|\varphi(\lambda)|<1. \tag{39}$$

Denote the parallelepiped with the "left lower" vertex $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ and side lengths $r_1,\dots,r_d$ by $\Box(x,r_1,\dots,r_d)=\Box(x,r)$, $r=(r_1,\dots,r_d)$:

$$\Box(x,r)=\{y=(y_1,\dots,y_d):y_j\in[x_j,x_j+r_j]\text{ for every }j=1,\dots,d\}.$$

Theorem 8. *Suppose that the conditions of Theorem 7 are satisfied. Then the following relation holds uniformly in $x\in\mathbb{R}^d$ for every fixed collection of numbers $r_1>0,\dots,r_d>0$:*

$$\mathbf{P}\{X_n\in\Box(x,r)\}=\frac{\sqrt{\det Q}}{(2\pi n)^{d/2}}e^{-\langle(x-n\mu)Q,x-n\mu\rangle/2n}\prod_{j=1}^d r_j+o(n^{-d/2}).$$

Proof. The assertion of the theorem is equivalent to the following:

$$\mathbf{P}\{X_n\in\Box(x,r)\}=\frac{1}{(2\pi n)^{d/2}}\int\limits_{\Box(x,r)}e^{-\langle(t-n\mu)Q,t-n\mu\rangle/2n}\,dt+o(n^{-d/2})$$

as $n\to\infty$ uniformly in x.

Consider the random vector $\eta=(\eta_1,\dots,\eta_d)$ with independent coordinates each of which has the common distribution with density (29) and characteristic function (30). We suppose that η is independent of X_n.

Fix $\varepsilon>0$. The characteristic function of the sum $X_n+\varepsilon^2\eta$ equals $\mathbf{E}e^{i\langle\lambda,X_n\rangle}\psi(\varepsilon^2\lambda)$, where $\psi(\lambda)=\psi(\lambda_1)\dots\psi(\lambda_d)$, $\lambda=(\lambda_1,\dots,\lambda_d)$. Since the characteristic function $|\psi(\lambda)|$ is integrable in $\mathbb{R}^d$, the sum $X_n+\varepsilon^2\eta$ has the bounded continuous density $p_{X_n+\varepsilon^2\eta}(z)$ which can be reconstructed for every $z\in\mathbb{R}^d$ by means of the following inversion formula (see, for instance, [9; Chapter XV § 3, § 7]):

$$p_{X_n+\varepsilon^2\eta}(z)=\frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d}\psi(\varepsilon^2\lambda)e^{-i\langle\lambda,z\rangle}\mathbf{E}e^{i\langle\lambda,X_n\rangle}\,d\lambda.$$

Hence,

$$n^{d/2}p_{X_n+\varepsilon^2\eta}(z)=\frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d}\psi(\varepsilon^2\lambda/\sqrt{n})e^{-i\langle\lambda,z-n\mu\rangle/\sqrt{n}}\mathbf{E}e^{i\langle\lambda,\frac{X_n-n\mu}{\sqrt{n}}\rangle}\,d\lambda.$$

From this equality and (36) we derive

$$\begin{aligned}\Delta_n&\equiv\left|n^{d/2}p_{X_n+\varepsilon^2\eta}(z)-\frac{\sqrt{\det Q}}{(2\pi)^{d/2}}e^{-\langle(z-n\mu)Q,z-n\mu\rangle/2n}\right|\\&\le\frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d}|\psi(\varepsilon^2\lambda/\sqrt{n})\mathbf{E}e^{i\langle\lambda,\frac{X_n-n\mu}{\sqrt{n}}\rangle}-e^{-\langle\lambda\sigma^2,\lambda\rangle/2}|\,d\lambda.\end{aligned}$$

Using compactness of the support of the function $\psi(\lambda)$, for every $A>0$ we hence obtain the inequality

$$\Delta_n\le\frac{1}{(2\pi)^d}\int\limits_{[-A,A)^d}|\psi(\varepsilon^2\lambda/\sqrt{n})\mathbf{E}e^{i\langle\lambda,\frac{X_n-n\mu}{\sqrt{n}}\rangle}-e^{-\langle\lambda\sigma^2,\lambda\rangle/2}|\,d\lambda$$

$$+\frac{1}{(2\pi)^d}\int\limits_{[-\sqrt{n}/\varepsilon^2,\sqrt{n}/\varepsilon^2]^d\setminus[-A,A]^d}|\mathbf{E}e^{i\langle\lambda,\frac{X_n-n\mu}{\sqrt{n}}\rangle}|\,d\lambda+\frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d\setminus[-A,A]^d}e^{-\langle\lambda\sigma^2,\lambda\rangle/2}\,d\lambda.$$

Arguing as in the proof of Theorem 7, we show that each of three summands on the right-hand side of the last estimate vanishes as $n, A \to \infty$ uniformly in $z \in \mathbb{R}^d$; here we use (39) in place of the property (32) of separateness of the module of the characteristic function from 1. Thus,

$$p_{X_n+\varepsilon^2\eta}(z) = \frac{\sqrt{\det Q}}{(2\pi n)^{d/2}} e^{-\langle (z-n\mu)Q, z-n\mu\rangle/2n} + o(n^{-d/2}) \tag{40}$$

as $n \to \infty$ uniformly in $z \in \mathbb{R}^d$.

Put

$$\Box_\varepsilon^+(x,r) = \{y = (y_1, \dots, y_d) : y_j \in [x_j - \varepsilon, x_j + r_j + \varepsilon]\},$$
$$\Box_\varepsilon^-(x,r) = \{y = (y_1, \dots, y_d) : y_j \in [x_j + \varepsilon, x_j + r_j - \varepsilon]\}.$$

It follows from the definition of the density (29) that $\mathbf{P}\{|\eta_1| \ge t\} \le 1/t$. Hence,

$$\mathbf{P}\{\varepsilon^2\eta \notin [-\varepsilon,\varepsilon]^d\} \le \sum_{j=1}^d \mathbf{P}\{\varepsilon^2\eta_j \notin [-\varepsilon,\varepsilon]\} \le d\varepsilon.$$

Since

$$\mathbf{P}\{X_n + \varepsilon^2\eta \in \Box_\varepsilon^+(x,r)\} \ge \mathbf{P}\{X_n \in \Box(x,r),\ \varepsilon^2\eta \in [-\varepsilon,\varepsilon]^d\}$$
$$= \mathbf{P}\{X_n \in \Box(x,r)\}\mathbf{P}\{\varepsilon^2\eta \in [-\varepsilon,\varepsilon]^d\} \ge \mathbf{P}\{X_n \in \Box(x,r)\}(1 - d\varepsilon)$$

in view of independence of X_n and η, from (40) we obtain the estimate from above

$$\mathbf{P}\{X_n \in \Box(x,r)\} \le \frac{\mathbf{P}\{X_n + \varepsilon^2\eta \in \Box_\varepsilon^+(x,r)\}}{1 - d\varepsilon}$$
$$= \frac{\sqrt{\det Q}}{(2\pi n)^{d/2}} e^{-\langle (x-n\mu)Q, x-n\mu\rangle/2n} \frac{\prod_{j=1}^d (r_j + 2\varepsilon)}{1 - d\varepsilon} + o(n^{-d/2}) \tag{41}$$

as $n \to \infty$ uniformly in $x \in \mathbb{R}^d$. Therefore (we can also use Theorem 6), there is a constant c such that for every $z \in \mathbb{R}^d$

$$\mathbf{P}\{X_n \in \Box(z,r)\} \le cn^{-d/2}.$$

This estimate implies the inequality

$$\mathbf{P}\{X_n + \varepsilon^2\eta \in \Box_\varepsilon^-(x,r),\ \varepsilon^2\eta \notin [-\varepsilon,\varepsilon]^d\}$$
$$= \int\limits_{u \notin [-\varepsilon,\varepsilon]^d} \mathbf{P}\{X_n \in \Box_\varepsilon^-(x-u,r)\}\mathbf{P}\{\varepsilon^2\eta \in du\} \le cn^{-d/2}\mathbf{P}\{\varepsilon^2\eta \notin [-\varepsilon,\varepsilon]^d\} \le cn^{-d/2}d\varepsilon.$$

From here and the inequality

$$\mathbf{P}\{X_n + \varepsilon^2\eta \in \Box_\varepsilon^-(x,r),\ \varepsilon^2\eta \in [-\varepsilon,\varepsilon]^d\} \le \mathbf{P}\{X_n \in \Box(x,r)\}$$

we obtain the following estimate from below:

$$\mathbf{P}\{X_n \in \Box(x,r)\} \ge \mathbf{P}\{X_n + \varepsilon^2\eta \in \Box_\varepsilon^-(x,r)\} - cn^{-d/2}d\varepsilon$$
$$= \frac{\sqrt{\det Q}}{(2\pi n)^{d/2}} e^{-\langle (x-n\mu)Q, x-n\mu\rangle/2n} \prod_{j=1}^d (r_j - 2\varepsilon) + o(n^{-d/2}) - cn^{-d/2}d\varepsilon \tag{42}$$

by (40). Combining (41) and (42) and using the arbitrariness of $\varepsilon > 0$, we complete the proof of the theorem.

References

1. *Meyn S. P. and Tweedie R. L.*, Markov Chains and Stochastic Stability, Springer-Verlag, London, Berlin, etc. (1993).
2. *Dobrushin R. L.*, “Central limit theorem for non-stationary Markov chains. I,” Teor. Veroyatnost. i Primenen., **1**, 72–89 (1956).
3. *Dobrushin R. L.*, “Central limit theorem for non-stationary Markov chains. II,” Teor. Veroyatnost. i Primenen., **1**, 365–425 (1956).
4. *Pinelis I.*, “Optimum bounds for the distribution of martingales in Banach spaces,” Ann. Probab., **22**, 1679–1706 (1994).
5. *Hoffman-Jørgensen J. and Pisier G.*, “The law of large numbers and the central limit theorem in Banach spaces,” Ann. Probab., **4**, 587–599 (1976).
6. *Pisier G.*, “Martingales with values in uniformly convex spaces,” Israel J. Math., **20**, 326–350 (1975).
7. *Borovkov A. A. and Korshunov D. A.*, “Large deviation probabilities for one-dimensional Markov chains. I: Prestationary distributions in the subexponential case,” Teor. Veroyatnost. i Primenen. (to appear).
8. *Petrov V. V.*, Sums of Independent Random Variables, Akademie-Verlag, Berlin (1975).
9. *Feller W.*, An Introduction to Probability Theory and Its Applications. Vol. 2, J. Wiley and Sons, New York, London, etc. (1971).