Z. Wahrscheinlichkeitstheorie verw. Gebiete
54, 59–73 (1980)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

The Berry-Esseen Theorem for Functionals of Discrete Markov Chains

E. Bolthausen

Technische Universität Berlin, Fachbereich Mathematik, Straße des 17. Juni 135, 1000 Berlin 12

Summary. The error bound $O(1/\sqrt{n})$ is derived in the central limit theorem for partial sums $\sum_{j=1}^{n} f(\xi_j)$ where ξ_j is a recurrent discrete Markov chain and f is a real valued function on the state space. In particular it is shown that for bounded f and starting distribution dominated by some multiple of the stationary one, it is sufficient for the chain to have recurrence times with third moments on order to get this bound.

§ 1. Introduction

Let I be an at most countable set of states, $(p_{ij})_{i,j\in I}$ a stochastic matrix (i.e. $p_{ij}\geqq 0$, $\sum_{j\in I} p_{ij}=1$ for all $i\in I$) and $X=(\Omega, \mathfrak{A}, \xi_n, P_i)$ a Markov chain with transition probabilities p_{ij}; i.e. for $n\in\mathbb{N}_0=\mathbb{N}\cup\{0\}$ ξ_n: $\Omega\to I$ is $\mathfrak{A}$-measurable and for $i\in I$ P_i is a probability measure on $(\Omega,\mathfrak{A})$ with $P_i(\xi_0=i)=1$ and $P_i(\xi_n=i_n|\xi_0=i_0,\dots,\xi_{n-1}=i_{n-1})=p_{i_{n-1}i_n}$ if the left side is defined. We assume that I is one recurrent class, i.e. for each $i\in I$ ξ_n visits each state infinitely often with P_i-probability 1. For a probability μ on I P_μ is the probability $\sum_{i\in I}\mu(i)P_i$ on $(\Omega,\mathfrak{A})$.

We fix once for all a distinguished point $O\in I$. Let T_k: $\Omega\to\mathbb{N}_0\cup\{\infty\}$ be defined as follows:

$$T_0=\inf\{n\geqq 0\colon \xi_n=0\}$$
$$T_k=\inf\{n>T_{k-1}\colon \xi_n=0\},\qquad k\geqq 1.$$

It is well known that for any starting probability μ and all $k\in\mathbb{N}_0$ $T_k<\infty$ P_μ-a.s., so the

$$\tau_k=T_k-T_{k-1},\qquad k\geqq 1$$

are well defined if we restrict everything on a subspace of Ω which has full

0044-3719/80/0054/0059/\$03.00

measure for all P_μ. For the sake of notational convenience we write τ for τ_1. The τ_k are well known to be independent and identically distributed.

If $f\colon I \to \mathbb{R}$, we call the sequence $f(\xi_0), f(\xi_1), \ldots$ a functional of X.

If the chain is positive, i.e. $E_0(\tau) < \infty$, then there exists a unique stationary probability distribution $\Pi = (\pi(i))_{i \in I}$, i.e. we have $\Sigma_i \pi(i) p_{ij} = \pi(j)$ for all $j \in I$. $\xi_0, \xi_1, \ldots$ with the law P_π is then a stationary process. We call it the stationary chain.

In the sequel the chain X is assumed to be positive recurrent. The following central limit theorem is due to Doeblin (see [3], I. 16, Theorem 1).

Theorem A. *If the chain is positive and if* $E_0\left(\sum_{i=1}^{\tau} |f|(\xi_i)\right)^2 < \infty$ *then* $\Pi(f) = \sum_{i \in I} \pi(i) f(i)$ *is well defined, and if*

$$\sigma^2(f) = E_0\left(\sum_{n=1}^{\tau} (f(\xi_n) - \Pi(f))\right)^2 > 0$$

then

$$\lim_{n \to \infty} P_0\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{n} (f(\xi_j) - \Pi(f)) \leq t\right) = \Phi(t)$$

where Φ is the standard normal distribution function and $\alpha = E_0(\tau)$.

Our main result is the following Berry-Esseen type bound:

Theorem 1. *Let μ be a starting probability on I. If*

$$E_0(\tau^3) < \infty \tag{1.1}$$

$$E_0\left(\sum_{j=1}^{\tau} |f|(\xi_j)\right)^3 < \infty \tag{1.2}$$

$$E_\mu(T_0) < \infty \tag{1.3}$$

$$E_\mu\left(\sum_{j=1}^{T_0} |f|(\xi_j)\right) < \infty \tag{1.4}$$

then

$$\sup_t \left| P_\mu\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{n} (f(\xi_j) - \Pi(f)) \leq t\right) - \Phi(t) \right| = O(n^{-1/2}) \tag{1.5}$$

The proof will be given in § 3.

Taking in particular $\mu = \Pi$, (1.3) and (1.4) are entailed by (1.1) and (1.2). To see this, the following result of Pitman [9] is useful:

Pitman's Occupation Measure Identity

Let $g\colon I^{\mathbb{N}_0} \to [0, \infty)$ be measurable, let S be a stopping time for $\xi_0, \xi_1, \ldots$ and let v be the occupation measure on I defined by $v(i) = E_0\left(\sum_{n=0}^{S-1} 1_i(\xi_n)\right)$. Then

$$E_0\left(\sum_{n=0}^{S-1} g(\xi_n, \xi_{n+1}, \ldots)\right) = \sum_{i\in I} v(i)\, E_i(g(\xi_0, \xi_1, \ldots)).$$

With this result one easily proves the following

Lemma 1. *If* $\mu = \Pi$ *then* (1.3) *and* (1.4) *follow from* (1.1) *and* (1.2).

Proof. (1.3) is well known to follow from (1.1) (see e.g. [9]).

Let $S = \inf\{n > 0 \colon \xi_n = 0\}$, $h(i) = \max(|f(i)|, 1)$. Then

$$\begin{aligned} E_\Pi\left(\sum_{j=1}^{T_0} |f|(\xi_j)\right) &\leqq E_\Pi\left(\sum_{j=1}^{S} h(\xi_j)\right) \\ &\leqq E_\Pi\left(\sum_{j=0}^{S-1} h(\xi_j)\right) + h(0) \\ &\leqq E_\Pi\left(h(\xi_0) \sum_{j=0}^{S-1} h(\xi_j)\right) + h(0) \\ &= \pi(0)\, E_0\left(\sum_{i=0}^{S-1} h(\xi_i) \sum_{j=1}^{S-1} h(\xi_i)\right) + h(0) \\ &\leqq \pi(0)\, E_0\left(\sum_{i=0}^{S-1} h(\xi_i)\right)^2 + h(0), \\ &\leqq 2\pi(0)\, E_0\left(\sum_{i=0}^{S-1} |f|(\xi_i)\right)^2 + 2\pi(0)\, E_0(\tau^2) + h(0), \end{aligned}$$

where the equality is by Pitman's identity, using the fact that the occupation measure for S is $\pi(0)\,\Pi$. So it is seen that (1.4) follows from (1.1) and (1.2).

From Lemma 1 and Theorem 1 one derives the following

Corollary 1. *If the starting probability* μ *is dominated by some multiple of* Π *and if* (1.1) *and* (1.2) *hold then* (1.5) *is true.*

It is desirable to have conditions based on more familiar entities. The following so-called strong mixing coefficients have been introduced by Rosenblatt (see [10]):

Let $\mathfrak{F}_k = \sigma(\xi_0, \ldots, \xi_k)$ $\mathfrak{F}^k = \sigma(\xi_j, j \geqq k)$. $\alpha(k)$, $k \geqq 0$ is defined to be

$$\sup_{n\in\mathbb{N}_0} \sup_{A\in\mathfrak{F}_n} \sup_{B\in\mathfrak{F}^{n+k}} |P_\pi(A\cap B) - P_\pi(A)\, P_\pi(B)|$$

The following theorem will be proved in §4.

Theorem 2. *Let* $\lambda \geqq 0, \in \mathbb{R}$ *then* $\sum_{n=0}^{\infty} n^\lambda \alpha(n) < \infty$ *if and only if the chain is aperiodic and* $E_0(\tau^{\lambda+2}) < \infty$.

With this result and Corollary 1 one has

Corollary 2. *If some multiple of* Π *dominates* μ, *if* f *is bounded and* $\sum_n n\alpha(n) < \infty$ *then* (1.5) *holds true.*

For unbounded functions one obtains for $p>3$

$$E_0\left(\sum_{j=1}^{\tau}|f|(\xi_j)\right)^3 \leqq E_0\left(\tau^{p-1}\sum_{j=1}^{\tau}|f|^p(\xi_j)\right)^{3/p}$$
$$\leqq (E_0(\tau^{3(p-1)/(p-3)}))^{(p-3)/p}\left(E_0\left(\sum_{j=1}^{\tau}|f|^p\right)\right)^{3/p}.$$

So one has

Corollary 3. *If some multiple of Π dominates μ and for a real number $p>3$ $\Pi(|f|^p)<\infty$ and $\sum_n n^{(p+3)/(p-3)}\alpha(n)<\infty$ then* (1.5) *holds true.*

Bounds of order $O(n^{-1/2})$ for bounded functions f have been obtained by Lifshits [7] under conditions based on the maximum correlation coefficients, i.e. the cosinus of the angle between the spaces $L_2(\mathfrak{F}_n)$ and $L_2(\mathfrak{F}^{n+k})$. Such conditions seem to be quite strong for Markov chains. If any of these angles is larger than zero $\alpha(n)$ converges to zero exponentially fast ([7], Theorem 5). It follows from our theorem 2 that for any chain with recurrence times with moments only of a finite order all maximal correlation coefficients equal 1.

The method of proof used here is the renewal approach which goes back to Doeblin:

Let $\rho_n=\max\{k: T_k\leqq n\}$ and $l_n=T_{\rho_n}$; let further $X_n=\sum_{j=T_{n-1}+1}^{T_n}(f(\xi_j)-\Pi(f))$. The X_j are independent and identically distributed.

Obviously

$$\sum_{j=1}^{n}(f(\xi_j)-\Pi(f))=\sum_{j=1}^{T_0}(f(\xi_j)-\Pi(f))+\sum_{j=1}^{\rho_n}X_j + \sum_{j=l_n+1}^{n}(f(\xi_j)-\Pi(f)). \tag{1.6}$$

Theorem A then follows from the independence of the X_j, a central limit theorem with random summation and the asymptotic negligibility of first and third summand in (1.6) (after appropriate norming). However, error bounds of order $n^{-1/2}$ for central limit theorems with random summation are known only if X_j and ρ_n are independent, which certainly is not true in our case. Landers and Rogge in [5] derived bounds under rather general conditions, but applied to the situation in theorem 1 they only yield $O(n^{-1/4}(\log n)^{1/4})$ (see [6]). Bounds of order $O(n^{-1/3+\delta})$ under stronger conditions had previously been obtained by me with a modification of Landers' and Rogge's method [2]. Theorem 1 follows upon a close look at the dependence between ρ_n and the X_j.

A straightforward simplification of our proof also gives the following theorem which refutes the seemingly general belief that bounds of order $n^{-1/2}$ in central limit theorems with random summation are obtainable only in the independent case.

Theorem 3. *Let $(\eta_i, r_i)_{i\in\mathbb{N}}$ be independent identically distributed two-dimensional random variables with*

$$E(\eta_i)=0,\ E(\eta_i^2)=1,\ E(|\eta_i|^3)<\infty, \qquad r_i\in\mathbb{N},\ E(r_i^3)<\infty.$$

Let $\alpha=E(r_i)$ *and* $\rho_n=\max\left\{k\colon \sum_{j=1}^{k} r_j\leqq n\right\}$. *Then*

$$\sup_t\left|P\left(\sqrt{\alpha/n}\sum_{j=1}^{\rho_n}\eta_j\leqq t\right)-\Phi(t)\right|=O(n^{-1/2}).$$

§ 2. A Semi-Local Berry-Esseen Bound

We prepare for the proof of Theorem 1 with a special Berry-Esseen theorem for two-dimensional i.i.d. random variables $((\zeta_n,\gamma_n)$, $n\in\mathbb{N}$, which are lattice in one component. So we assume there is a $\rho\in\mathbb{R}$ such that $\gamma_n\in\rho+\mathbb{Z}$ a.s. It is further assumed that $E\zeta_n=E\gamma_n=0$, $E|\zeta_n|^3<\infty$, $E|\gamma_n|^3<\infty$ and that the covariance matrix $\Sigma=(\sigma_{ij})_{i,j=1,2}$ has full rank 2.

Let $\Lambda=\{n\in\mathbb{N}\colon\ \exists\, k\in\mathbb{Z}$ with $P(\gamma-\rho=k)>0,\ P(\gamma-\rho=k+n)>0\}$. Clearly $\Lambda\neq\emptyset$, and for the sake of convenience we assume the largest common divisor d of Λ to be 1. This is not essential. The modifications needed in the case when this is not true are straightforward and therefore omitted.

Let φ be the two-dimensional density function of the centred normal distribution with covariance Σ, and let $\psi(x,y)=\int_{-\infty}^{x}\varphi(s,y)\,ds$. Let $S_n=\sum_{i=1}^{n}\zeta_i$, $T_n=\sum_{i=1}^{n}\gamma_i$, $\lambda_n(t_1,t_2)$ be the characteristic function of $(S_n/\sqrt{n},\ T_n/\sqrt{n})$ and $g(t_1,t_2)$ be the characteristic function of $(\zeta_i,\gamma_i-\rho)$. Obviously

$$\lambda_n(t_1,t_2)=[g(t_1/\sqrt{n},t_2/\sqrt{n})\exp(it_2\rho/\sqrt{n})]^n. \tag{2.1}$$

Lemma 1. *Given* $\delta>0$, *there exist* $\delta'>0$, $0<r<1$ *and* $C>0$ *such that* $\lambda_n(t_1,t_2)$ *and all partial derivatives up to the third (or any fixed) order are dominated in absolute value by* Cr^{-n} *for* $|t_1|\leqq\delta'\sqrt{n}$, $\delta\sqrt{n}\leqq|t_2|\leqq\pi\sqrt{n}$.

Proof. From the assumption $d=1$ it follows that $|g(0,v)|$ is bounded away from 1 uniformly in $\delta\leqq|v|\leqq\pi$. From continuity of g it follows that there is a $\delta'>0$, $r<1$ with $|g(u,v)|\leqq r$ for $|u|\leqq\delta'$, $\delta\leqq|v|\leqq\pi$. The lemma now follows from (2.1) and the chain rule.

Proofs of the following two propositions may be found in [1] (Theorem 9.10 and Theorem 22.1).

Proposition A. *Let* $\lambda_0(t_1,t_2)=\exp\left(-\frac{1}{2}\sum_{j,k=1}^{2}t_jt_k\sigma_{jk}\right)$. *There exist constants* $\tilde{\varepsilon},\beta,c>0$ *(depending only on* Σ *and* $E|\zeta_i|^3$, $E|\gamma_i|^3$) *such that for* $u,v\in\mathbb{N}_0$, $u+v\leqq 3$

$$\left|\frac{\partial^{u+v}}{\partial t_1^u\,\partial t_2^v}(\lambda_n(t_1,t_2)-\lambda_0(t_1,t_2))\right|\leqq\frac{c}{\sqrt{n}}|t|^{3-u-v}e^{-\beta|t|^2}$$

for $|t_1|, |t_2| \leqq \tilde{\varepsilon}\sqrt{n}$, *where* $t=(t_1, t_2)$ *and* $|t|=(t_1^2+t_2^2)^{1/2}$.
For $\alpha \in \mathbb{Z}$ let $y_{\alpha,n}=(n\rho+\alpha)/\sqrt{n}$.

Proposition B.

$$\sup_{\alpha \in \mathbb{Z}} (1+|y_{\alpha,n}|^3)\,|P(T_n/\sqrt{n}=y_{\alpha,n})-\psi(\infty, y_{\alpha,n})/\sqrt{n}|=O(n^{-1}).$$

The main result of this section is

Theorem 4. *Under the above stated conditions*

$$\sup_{x\in\mathbb{R}} \sup_{\alpha\in\mathbb{Z}} (1+y_{\alpha,n}^2)\left|P\left(\frac{S_n}{\sqrt{n}} \leqq x, \frac{T_n}{\sqrt{n}}=y_{\alpha,n}\right)-\frac{1}{\sqrt{n}}\psi(x, y_{\alpha,n})\right|=O\left(\frac{1}{n}\right).$$

Remark. The proof given below easily gives the stronger statement where $y_{\alpha,n}^2$ is replaced by $|y_{\alpha,n}|^{3-\delta}$ ($\delta>0$). The statement with $|y_{\alpha,n}|^3$ may also be true but would probably require more refined techniques. The above theorem is sufficient for our purpose.

Proof. Let $F_n(x, y_{\alpha,n})=P(S_n/\sqrt{n}\leqq x, \; T_n/\sqrt{n}=y_{\alpha,n})$ and $\hat{F}_n(x, y_{\alpha,n}) = F_n(x, y_{\alpha,n})/F_n(\infty, y_{\alpha,n})$ if $F_n(\infty, y_{\alpha,n})>0$, $\hat{F}_n(x, y_{\alpha,n})=1_{[0,\infty)}(x)$ if $F_n(\infty, y_{\alpha,n})=0$. For fixed $\alpha\in\mathbb{Z}$ $\hat{F}_n(\cdot, y_{\alpha,n})$ is a distribution function. Let $\hat{\psi}(x, y_{\alpha,n}) = \psi(x, y_{\alpha,n})/\psi(\infty, y_{\alpha,n})$. For $T>0$ let $v_T(x)=(1-\cos(Tx))/(\pi T x^2)$. v_T is the density of a probability distribution with characteristic function $\omega_T(\lambda)=\max(0, 1-|\lambda|/T)$.

Let $F_n^T(x,y)=\int_{-\infty}^{\infty} F_n(x-u, y)\, v_T(u)\, du$ and $\hat{F}_n^T, \psi^T, \hat{\psi}^T$ be defined by similar convolutions. (We drop indices α, n in $y_{\alpha,n}$ for the sake of notational simplicity.)

From Lemma 3.1, Ch. XVI of [4]

$$\begin{aligned}&\sup_x |\hat{F}_n(x,y)-\hat{\psi}(x,y)|\\ &\quad\leqq 2\sup_x |\hat{F}_n^T(x,y)-\hat{\psi}^T(x,y)|+\frac{12}{\pi T}\sup_x\left|\frac{\partial}{\partial x}\hat{\psi}(x,y)\right|.\end{aligned}$$

After some elementary calculations it follows that

$$\begin{aligned}\sup_x |F_n(x,y)-\psi(x,y)/\sqrt{n}| &\leqq 2\sup_x |F_n^T(x,y)-\psi^T(x,y)/\sqrt{n}|\\ &\quad+3\,|F_n(\infty,y)-\psi(\infty,y)/\sqrt{n}|+12\sup_x \varphi(x,y)/(\pi T\sqrt{n}).\end{aligned}$$

Combining this with Proposition B and taking $T\sim\sqrt{n}$ one has

$$\begin{aligned}&\sup_{x\in\mathbb{R},\,\alpha\in\mathbb{Z}} (1+y_{\alpha,n})^2\,|F_n(x,y_{\alpha,n})-\psi(x,y_{\alpha,n})/\sqrt{n}|\\ &\quad\leqq 2\sup_{x,\alpha}(1+y_{\alpha,n}^2)\,|F_n^T(x,y_{\alpha,n})-\psi^T(x,y_{\alpha,n})/\sqrt{n}|+O\left(\frac{1}{n}\right).\end{aligned} \tag{2.2}$$

From now on we take $T=\varepsilon\sqrt{n}$, $\varepsilon=\min(\tilde{\varepsilon}, \delta')$ where $\tilde{\varepsilon}$ comes from Proposition A, δ' from lemma 1, and in this lemma $\delta=\tilde{\varepsilon}$. Now

$$F_n^T(x,y)-\psi^T(x,y)/\sqrt{n}=\frac{1}{(2\pi)^2\sqrt{n}}\int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\frac{1}{-it_1}e^{-it_1x-it_2y}\omega_T(t_1)\cdot(\lambda_n(t_1,t_2)-\lambda_0(t_1,t_2))\,dt_2\,dt_1 \tag{2.3}$$

and therefore if $z<x$.

$$\begin{aligned}
&y^2(F_n^T(x,y)-\psi^T(x,y)/\sqrt{n}-(F_n^T(z,y)-\psi^T(z,y)/\sqrt{n}))\\
&=\frac{1}{(2\pi)^2\sqrt{n}}\int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\frac{1}{it_1}(e^{-it_1x}-e^{-it_1z})e^{-it_2y}\\
&\quad\cdot\omega_T(t_1)\frac{\partial^2}{\partial t_2^2}(\lambda_n(t_1,t_2)-\lambda_0(t_1,t_2))\,dt_2\,dt_1\\
&=\frac{1}{(2\pi)^2\sqrt{n}}\int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}}\left\{\int_{-\tilde{\varepsilon}\sqrt{n}}^{\tilde{\varepsilon}\sqrt{n}}\cdots\frac{\partial^2}{\partial t_2^2}(\lambda_n(t_1,t_2)-\lambda_0(t_1,t_2))\,dt_2\right. \qquad (2.4)\\
&\quad+\int_{|t_2|\in\sqrt{n}[\tilde{\varepsilon},\pi]}\cdots\frac{\partial^2}{\partial t_2^2}\lambda_n(t_1,t_2)\,dt_2\\
&\quad\left.-\int_{|t_2|\in\sqrt{n}[\tilde{\varepsilon},\pi]}\cdots\frac{\partial^2}{\partial t_2^2}\lambda_0(t_1,t_2)\,dt_2\right\}dt_1\\
&=I_1+I_2+I_3\quad\text{say.}
\end{aligned}$$

We write

$$\begin{aligned}
&h(t_1,t_2)=\frac{\partial^2}{\partial t_2^2}(\lambda_n(t_1,t_2)-\lambda_0(t_1,t_2))\\
&h(t_1,t_2)=(1-e^{-\beta t_1^2})\,h(t_1,t_2)+e^{-\beta t_1^2}(h(t_1,t_2)-h(0,t_2))\\
&\qquad+e^{-\beta t_1^2}h(0,t_2),
\end{aligned}\tag{2.5}$$

where β is from Proposition A. From this proposition one has for $|t_1|\leqq\varepsilon$, $|t_2|\leqq\tilde{\varepsilon}$:

$$|h(t_1,t_2)|\leqq\frac{c}{\sqrt{n}}|t|\,e^{-\beta|t|^2},\qquad |h(t_1,t_2)-h(0,t_2)|\leqq\frac{c}{\sqrt{n}}|t_1|\,e^{-\beta t_1^2},$$

$$|h(0,t_2)|\leqq\frac{c}{\sqrt{n}}e^{-\beta t_2^2}.$$

Further $|1-e^{-\beta t_1^2}|\leqq\beta t_1^2$. Implementing these estimates in (2.4) and (2.5) one has

$$\begin{aligned}
|I_1|\leqq\frac{c'}{\sqrt{n}}&\left\{\int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}}\left\{\int_{-\tilde{\varepsilon}\sqrt{n}}^{\varepsilon\sqrt{n}}|t_1|\,\omega_T(t_1)\,e^{-\beta|t|^2}\,dt_2\right.\right.\\
&\left.+\int_{-\tilde{\varepsilon}\sqrt{n}}^{\tilde{\varepsilon}\sqrt{n}}|\omega_T(t_1)|\,e^{-\beta|t|^2}\,dt_2\right\}dt_1\\
&\left.+\left|\int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}}\frac{1}{it_1}(e^{-it_1x}-e^{-it_1z})\,\omega_T(t_1)\,e^{-\beta t_1^2}\,dt_1\right|\int_{-\tilde{\varepsilon}\sqrt{n}}^{\tilde{\varepsilon}\sqrt{n}}e^{-\beta t_2^2}\,dt_2\right\}\\
=O(n^{-1})&
\end{aligned}$$

uniformly in x, z, α,

I_2 can be handled similarly by using Lemma 1 instead of Proposition B and splitting as follows:

$$\frac{\partial^2}{\partial t_2^2}\lambda_n(t_1,t_2)=\frac{\partial^2}{\partial t_2^2}(\lambda_n(t_1,t_2)-\lambda_n(0,t_2))+\frac{\partial^2}{\partial t_2^2}\lambda_n(0,t_2).$$

In this way one obtains $|I_2|=O(\delta^{-n})$ and in the same way, using exponential decrease of λ_0, one has exponential decrease of $|I_3|$. So $|I_1|+|I_2|+|I_3|=O(n^{-1})$, and letting $z\to-\infty$ one has

$$y^2(F_n^T(x,y)-\psi^T(x,y)/\sqrt{n})=O(n^{-1}).$$

In the same way

$$(F_n^T(x,y)-\psi^T(x,y)/\sqrt{n})=O(n^{-1})$$

and from these estimates and (2.2) the theorem follows.

§ 3. Proof of Theorem 1

In this section $c, c', \varepsilon, \varepsilon'$ are always constants >0, c, c' "sufficiently large" and $\varepsilon, \varepsilon'$ "sufficiently small" which do not depend on n, m, t, s etc. They may vary from formula to formula but not in the same.

We resume the notation of § 1. We have from (1.6)

$$\begin{aligned}&\left\{\frac{\sqrt{\alpha}}{\sigma\sqrt{n}}\sum_{j=1}^{n}(f(\xi_j)-\Pi(f))\leqq t\right\}\\&=\bigcup_{m=0}^{n}\bigcup_{s=0}^{n}\bigcup_{r=0}^{n}\left\{\frac{\sqrt{\alpha}}{\sigma\sqrt{n}}\left(\sum_{j=1}^{r}(f(\xi_j)-\Pi(f))+\sum_{j=1}^{m}X_j\right.\right.\\&\left.\left.\quad+\sum_{j=n-s+1}^{m}(f(\xi_j)-\Pi(f))\right)\leqq t,\ T_0=r,\ \sum_{j=1}^{m}\tau_j=n-s-r,\ \tau_{m+1}>s\right\}.\end{aligned}$$

By the Markov property

$$\begin{aligned}&P_\mu\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}}\sum_{j=1}^{n}(f(\xi_j)-\Pi(f))\leqq t\right)\\&=\sum_{m=0}^{n}\sum_{s=0}^{n}\sum_{r=0}^{n}\iint P_0\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}}\sum_{j=1}^{m}X_j\leqq t-u-v,\ \sum_{j=1}^{m}\tau_j=n-s-r\right)\\&\quad\cdot P_\mu(R_r\in dv,\ T_0=r)\,P_0(R_s\in du,\ \tau>s)\end{aligned}\tag{3.1}$$

where

$$R_k=\frac{\sqrt{\alpha}}{\sigma\sqrt{n}}\sum_{j=1}^{k}(f(\xi_j)-\Pi(f)).$$

The sum on the right side of (3.1) may be splitted as

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=0}^{n} + \sum_{s=\sqrt{n}+1}^{n} \sum_{r=0}^{n} \sum_{m=0}^{n} + \sum_{r=\sqrt{n}+1}^{n} \sum_{s=0}^{n} \sum_{m=0}^{n}$$

where it is understood that summation begins or ends at the integer part of a number. The second summand is bounded by

$$\sum_{s=\sqrt{n}+1}^{\infty} P_0(\tau > s) = O(n^{-1})$$

and the third by

$$\sum_{r=\sqrt{n}+1}^{\infty} P_\mu(T_0 = r) = O(n^{-1/2})$$

so in order to prove the theorem it suffices to consider summation over s, r up to $\sqrt{n}$ in (3.1). Clearly $m=0$ may then be excluded.

Let $\zeta_k = X_k/\sigma$, $\gamma_k = \tau_k - \alpha$. For the moment we assume that (ζ_k, τ_k) has covariance matrix of rank 2, so we can apply theorem 4 of § 2. We assumed there $d = 1$ which means that the chain is aperiodic. However, this is only for notational convenience and is easily seen to be of no importance. We have

$$\begin{aligned} &P_\mu\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{m} X_j \leqq t-u-v, \sum_{j=1}^{m} \tau_j = n-s-r\right) \\ &= \frac{1}{\sqrt{m}} \psi\left(\sqrt{\frac{n}{\alpha m}}(t-u-v), \lambda_{r,s,m}\right) + O\left(\frac{1}{m(1+\lambda_{r,s,m}^2)}\right) \end{aligned} \tag{3.3}$$

for $m \geqq 1$ where $\lambda_{r,s,m} = (n-s-r-\alpha m)/\sqrt{m}$.

The O-term on the right side of (3.3) does not depend on u, v, so from (3.1)–(3.3) follows

$$\begin{aligned} &P_\mu\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{n} (f(\xi_j) - \Pi(f)) \leqq t\right) \\ &= \sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} \left\{ \iint \psi\left(\sqrt{\frac{n}{\alpha m}}(t-u-v), \lambda_{r,s,m}\right) \right. \\ &\quad \cdot P_\mu(R_r \in dv, T_0 = r)\, P_0(R_s \in du, \tau_1 > s) \\ &\quad \left. + O\left(\frac{1}{m}(1+\lambda_{r,s,m}^2)^{-1}\right) P_\mu(T_0 = r)\, P_0(\tau_1 > s) \right\}. \end{aligned} \tag{3.4}$$

The theorem then follows from the following three relations

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} O\left(\frac{1}{m}(1+\lambda_{r,s,m}^2)^{-1}\right) = O(n^{-1/2}) \tag{3.5}$$

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} \Bigg| \iint \frac{1}{\sqrt{m}} \psi\left(\sqrt{\frac{n}{\alpha m}}(t-u-v), \lambda_{r,s,m}\right)$$
$$\cdot P_0(R_s \in du, \tau > s)\, P_\mu(R_r \in dv, T_0 = r) \tag{3.6}$$
$$- \frac{1}{\sqrt{m}} \psi(t, \lambda_{r,s,m})\, P_0(\tau > s)\, P_\mu(T_0 = r)\Bigg| = O(n^{-1/2})$$

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \left| \sum_{m=1}^{n} \frac{1}{\sqrt{m}} \psi(t, \lambda_{r,s,m}) - \frac{1}{\alpha} \int_{-\infty}^{\infty} \psi(t,x)\,dx \right| P_0(\tau > s)\, P_\mu(T_0 = r) = O(n^{-1/2}) \tag{3.7}$$

everything uniformly in t.

Indeed (3.3)-(3.7) imply

$$P\left(\frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{n} (f(\xi_j) - \Pi(f)) \leqq t\right)$$
$$= \left(\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{n} \frac{1}{\alpha} P_0(\tau > s)\, P_\mu(T_0 = r)\right) \int_{-\infty}^{\infty} \psi(t,x)\,dx + O(n^{-1/2})$$
$$= \phi(t) + O(n^{-1/2}).$$

So the theorem is proved in the case where the covariance matrix of (ζ, γ) is nondegenerated. It remains to prove (3.5)-(3.7) for this case.

Proof of (3.5). For $r, s \leqq \sqrt{n}$

$$\frac{1}{m}(1+\lambda_{r,s,m}^2)^{-1} \leqq \begin{cases} 1/m \leqq c/n & \text{for } |n-\alpha m| \leqq 2\sqrt{n} \\ (n-\alpha m)^{-2} & \text{for } \alpha m > n + 2\sqrt{n} \\ (n-2\sqrt{n}-\alpha m)^{-2} & \text{for } \alpha m < n - 2\sqrt{n} \end{cases}$$

From this (3.5) follows by some elementary calculations

Proof of (3.6). Let $I_m(t,u,v)$ be the interval between t and $\sqrt{\frac{n}{\alpha m}}(t-u-v)$. We have

$$\left|\psi\left(\sqrt{\frac{n}{\alpha m}}(t-u-v), \lambda_{r,s,m}\right) - \psi(t, \lambda_{r,s,m})\right|$$
$$\leqq \left(\sqrt{\frac{n}{\alpha m}}(|u|+|v|) \sup_{x\in\mathbb{R}} \varphi(x, \lambda_{r,s,m})\right.$$
$$\left. + |t| \left|\sqrt{\frac{n}{\alpha m}} - 1\right| \sup_{x\in I_m} \varphi(x, \lambda_{r,s,m}).\right.$$

So

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} \Bigg| \iint \frac{1}{\sqrt{m}} \psi\left(\sqrt{\frac{n}{\alpha m}}(t-u-v), \lambda_{r,s,m}\right)$$
$$\cdot P_0(R_s \in du, \tau > s)\, P_\mu(R_r \in dv, T_0 = r)$$
$$- \frac{1}{\sqrt{m}} \psi(t, \lambda_{r,s,m})\, P_0(\tau > s)\, P_\mu(T_0 = r)\Bigg|$$

$$\leqq \sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} \left\{ \int \frac{1}{\sqrt{m}} \sqrt{\frac{n}{\alpha m}} |u| \sup_{x\in\mathbb{R}} \varphi(x, \lambda_{r,s,m}) P_0(R_s \in du, \tau > s) P_\mu(T_0 = r) \right. \tag{3.8}$$

$$+ \int \frac{1}{\sqrt{m}} \sqrt{\frac{n}{\alpha m}} |v| \sup_{x\in\mathbb{R}} \varphi(x, \lambda_{r,s,m}) P_0(\tau > s) P_\mu(R_r \in dv, T = r)$$

$$+ \iint \frac{1}{\sqrt{m}} |t| \left| \sqrt{\frac{n}{\alpha m}} - 1 \right| \sup_{x \in I_m(t,u,v)} |\varphi(x, \lambda_{r,s,m})| P_0(R_s \in du, \tau > s)$$

$$\cdot P_\mu(R_r \in dv, T_0 = r)$$

$$= A_1 + A_2 + A_3 \quad \text{say.}$$

Obviously $\sup_{x\in\mathbb{R}} \varphi(x,\lambda) \leqq c \exp(-\varepsilon\lambda^2)$ and on $\{\tau > s\}$

$$|R_s| \leqq \frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{\tau} |f(\xi_j) - \Pi(f)| = Z \qquad \text{say.}$$

So

$$\int |u| P_0(R_s \in du, \tau > s) = E_0(|R_s| 1_{\{\tau > s\}}) \leqq E(Z 1_{\{\tau > s\}})$$

$$\leqq \frac{1}{s^2} E(Z\tau^2) \leqq c/(\sqrt{n}\, s^2)$$

by Hölder's inequality. So

$$A_1 \leqq c \sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \left(\frac{1}{s^2} \wedge 1\right) P_\mu(T_0 = r) \sum_{m=1}^{n} \frac{1}{m} \exp(-\varepsilon \lambda_{r,s,m}^2). \tag{3.9}$$

Let now

$$\Lambda_0 = \{m \colon n - 3\sqrt{n} \leqq \alpha m \leqq n + \sqrt{n}\}$$

$$\Lambda_k = \{m \colon n + k\sqrt{n} < \alpha m \leqq n + (k+1)\sqrt{n}\}; \qquad k \geqq 1$$

$$\Lambda'_k = \{m \colon n - (k+1)\sqrt{n} \leqq \alpha m < n - k\sqrt{n}\}; \qquad k \geqq 3.$$

Remarking now that for $s, r \leqq \sqrt{n}$, $m \leqq n$, $m \in \Lambda_k$ one has $|\lambda_{r,s,m}| \geqq \frac{k}{\sqrt{2}}$, and for $m \in \Lambda'_k$ $|\lambda_{r,s,m}| \geqq (k-2)$ one obtains by splitting the sum $\sum_{m=1}^{n}$ into the sums over the Λ's after some elementary calculations

$$\sum_{m=1}^{n} \frac{1}{m} \exp(-\varepsilon \lambda_{r,s,m}^2) = O(n^{-1/2}) \qquad \text{for } s, r \leqq \sqrt{n},$$

so $A_1 = O(n^{-1/2})$ follows. A_2 can be handled similarly. We consider now A_3.

Let Z be as above and

$$Z' = \frac{\sqrt{\alpha}}{\sigma\sqrt{n}} \sum_{j=1}^{T_0} |f(\xi_j) - \Pi(f)|.$$

For $|u| + |v| \leqq \frac{2|t|}{3}$

$$\sup_{x \in I_m(t,u,v)} \varphi(x, \lambda_{r,s,m}) \leqq c \exp\left(-\varepsilon \frac{n}{\alpha m} t^2\right) \exp(-\varepsilon \lambda_{r,s,m}^2).$$

Therefrom

$$\begin{aligned}
&\sup_{t\in\mathbb{R}} \iint |t| \sup_{x\in I_m(t,u,v)} \varphi(x,\lambda_{r,s,m})\, P_0\left(R_s\in du, Z\leqq\frac{|t|}{3}, \tau>s\right)\\
&\quad\cdot P_\mu\left(R_r\in dv, Z'\leqq\frac{|t|}{3}, T_0=r\right) \qquad (3.10)\\
&\leqq c\sqrt{\frac{\alpha m}{n}}\,\exp(-\varepsilon\lambda_{r,s,m}^2)\, P_0(\tau>s)\, P_\mu(T_0=r)
\end{aligned}$$

$$\sup_{t\in\mathbb{R}} |t|\, P_\mu\left(Z'>\frac{|t|}{3}, T_0=r\right)\leqq c\,E_\mu(Z'\,1_{T_0=r}) \qquad (3.11)$$

$$\sup_{t\in\mathbb{R}} |t|\, P_0\left(Z>\frac{|t|}{3}, \tau>s\right)\leqq c\left(\frac{1}{s^2}\wedge 1\right). \qquad (3.12)$$

Combining (3.10)–(3.12) gives

$$\begin{aligned}
A_3\leqq c\sum_{s=0}^{\sqrt{n}}\sum_{r=0}^{\sqrt{n}}\sum_{m=1}^{n}\frac{1}{\sqrt{m}}\left|\sqrt{\frac{n}{\alpha m}}-1\right|\exp(-\varepsilon\lambda_{r,s,m}^2)\\
\cdot\left(\left(\frac{1}{s^2}\wedge 1\right)(E_\mu(Z'\,1_{T_0=r})+P_\mu(T_0=r))\right).
\end{aligned}$$

Splitting the sum over m into subsummations over the Λ's one obtains after some elementary calculations for $s,r\leqq\sqrt{n}$

$$\sum_{m=1}^{n}\frac{1}{\sqrt{m}}\left|\sqrt{\frac{n}{\alpha m}}-1\right|\exp(-\varepsilon\lambda_{r,s,m}^2)=O(n^{-1/2}).$$

So $A_3=O(n^{-1/2})$ follows.

Proof of (3.7). For fixed s,r $\lambda_{r,s,m}$ decreases as m increases and

$$\lambda_{r,s,m}-\lambda_{r,s,m+1}=\alpha\frac{1}{\sqrt{m}}+\lambda_{r,s,m+1}\underbrace{\left(\sqrt{1+\frac{1}{m}}-1\right)}_{=O\left(\frac{1}{m}\right)}.$$

From this one easily derives

$$\begin{aligned}
\sum_{s=0}^{\sqrt{n}}\sum_{r=0}^{\sqrt{n}}\left|\frac{1}{\sqrt{m}}\psi(t,\lambda_{r,s,m})-\frac{1}{\alpha}(\lambda_{r,s,m}-\lambda_{r,s,m+1})\,\psi(t,\lambda_{r,s,m})\right|\\
\cdot P_0(\tau>s)\,P_\mu(T_0=r)=O(n^{-1/2}).
\end{aligned} \qquad (3.13)$$

Further

$$\sup_{x\in[\lambda_{m+1},\lambda_m]}|\psi(t,x)-\psi(t,\lambda_m)|\leqq c(\lambda_m-\lambda_{m+1})\exp(-\varepsilon\lambda_m^2)$$

and therefore

$$\sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \left| \sum_{m=1}^{n} \frac{1}{\alpha}(\lambda_{r,s,m}-\lambda_{r,s,m+1})\,\psi(t,\lambda_{r,s,m}) \cdot P_0(\tau>s)\,P_\mu(T_0=r) - \int_{\lambda_{r,s,n}}^{\lambda_{r,s,1}} \psi(t,x)\,dx\, P_0(\tau>s)\,P_\mu(T_0=r) \right|$$

$$\leq c \sum_{s=0}^{\sqrt{n}} \sum_{r=0}^{\sqrt{n}} \sum_{m=1}^{n} (\lambda_{r,s,m}-\lambda_{r,s,m+1})^2 \exp(-\varepsilon \lambda_{r,s,m}^2) \cdot P_0(\tau>s)\,P_\mu(T_0=r) = O(n^{-1/2}).$$

Obviously

$$\int_{\lambda_n}^{\lambda_1} \psi(t,x)\,dx = \int_{-\infty}^{\infty} \psi(t,x) + O(n^{-1/2}),$$

so (3.13)-(3.14) entail (3.7).

The case where Σ is degenerated is much more simple. First, if τ is norandom it can easily be reduced to the standard Berry-Esseen theorem. If τ is nondegenerated but Σ has rank 1, then there exists a constant $a \in \mathbb{R}$ such that $\zeta_i = a\gamma_i$ a.s. A typical example for this is if $f = 1_{\{0\}}$. In this special case the statement of theorem 1 (with fixed starting point) has been proved by Landers and Rogge in [6] (theorem 1). Their proof can easily be adapted to the general case where $\zeta_i = a\gamma_i$. We omit the details.

§ 4. Proof of Theorem 2

Rosenblatt ([10], VII.3, Lemma 1) obtained the result that a Markov chain is strongly mixing, i.e. $\lim_{n\to\infty} \alpha(n) = 0$, if and only if

$$\sup \{\sum_{i\in I} \pi(i)\,|E_i(f(\xi_n)) - \Pi(f)| : f: I \to \mathbb{R},\ \|f\|_\infty \leq 1\}$$

goes to 0 as $n \to \infty$. His proof easily gives the following stronger statement:

Lemma 2.
$$\alpha(n) \leq \tfrac{1}{2} \sup_{\|f\|_\infty \leq 1} \Pi(|E_\cdot(f(\xi_n)) - \Pi(f)|)$$
$$\leq 2 \sup_{A,B\subset I} |P_\Pi(\xi_0 \in A, \xi_n \in B) - \pi(A)\pi(B)|$$
$$\leq 2\alpha(n).$$

Proof of Theorem 2. (I) It is assumed that $\sum_{n=1}^{\infty} n^p \alpha(n) < \infty$ for some $p \geq 0$. Let $A_n = \{\xi_j \neq 0$ for $j = n+1, n+2, \ldots, 2n\}$. For any starting probability μ

$$P_\mu(A_n) = \sum_{m=1}^{n} P_\mu(\xi_m = 0)\,P_0(S > 2n-m) + P_\mu(S > 2n).$$

Taking $\mu=\delta_0$ and $\mu=\Pi$ one obtains

$$|P_0(S>2n)-P_\pi(S>2n)|$$
$$\leqq |P_0(A_n)-P_\Pi(A_n)|+P_0(S>n)\sum_{m=1}^{\infty}|P_0(\xi_m=0)-\pi(0)|.$$

Now $|P_0(A_n)-P_\Pi(A_n)|\leqq\alpha(n)/\pi(0)$ and $|P_0(\xi_m=0)-\pi(0)|\leqq\alpha(m)/\pi(0)$. So for q, $0\leqq q\leqq p$,

$$\pi(0)\sum_{n=1}^{\infty}n^q|P_0(S>2n)-P_\Pi(S>2n)|$$
$$\leqq\sum_{n=1}^{\infty}n^q\alpha(n)+\left(\sum_{n=1}^{\infty}\alpha(n)\right)\left(\sum_{n=1}^{\infty}n^q P_0(S>n)\right).$$

So it follows that if $E_0(S^{q+1})<\infty$ then $E_\Pi(S^{q+1})<\infty$. On the other hand it is well known that for any $r\geqq 0$ $E_0(S^{r+1})<\infty$ if and only if $E_\Pi(S^r)<\infty$. So it clearly follows that $E_0(S^{p+2})<\infty$.

(II) Let us prove the converse, so we assume $E_0(S^{p+2})<\infty$ for some $p\geqq 0$ or, what is the same, $E_\Pi(S^{p+1})<\infty$.

We use the Pitman coupling technique (see [8]), so let ξ_n,ξ'_n be two independent chains with transition probabilities p_{ij}. We write $\hat{P}.$ for the law of the pair (ξ_n,ξ'_n). Let $R=\inf\{n\geqq 0\colon \xi_n=0,\ \xi'_n=0\}$ and let $f\colon I\to\mathbb{R}$, $\|f\|_\infty\leqq 1$. As in Pitman [8]

$$|E_i(f(\xi_n))-E_\Pi(f(\xi_n))|\leqq 2\hat{P}_{\delta_i\times\Pi}(R\geqq n)$$

so

$$\Pi(|E.(f(\xi_n))-\Pi(f)|)\leqq 2\hat{P}_{\Pi\times\Pi}(R\geqq n).$$

If $E_\Pi(S^{p+1})<\infty$ Pitman proved in [8] that $E_{\Pi\times\Pi}(R^{p+1})<\infty$ so $\sum_{n=1}^{\infty}n^p\hat{P}_{\Pi\times\Pi}(R\geqq n)<\infty$ and $\sum_{n=1}^{\infty}n^p\alpha(n)<\infty$ follows from Lemma 2.

References

1. Battacharga, R.N., Rao, R.R.: Normal approximation and asymptotic expansions. New York: Wiley 1976
2. Bolthausen, E.: On rates of convergence in a random central limit theorem and in the central limit theorem for Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **38**, 279–286 (1977)
3. Chung, K.L.: Markov chains with stationary transition probabilities. Berlin, Heidelberg, New York: Springer 1967
4. Feller, W.: An introduction to probability theory and its applications, Vol. II. New York: Wiley 1966
5. Landers, D., Rogge, L.: The exact approximation order in the central limit theorem for random summation. Z. Wahrscheinlichkeitstheorie verw. Gebiete **36**, 169–183 (1976)

6. Landers, D., Rogge, L.: On the rate of convergence in the central limit theorem for Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **35**, 169–183 (1976)
7. Lifshits, B.A.: On the central limit theorem for Markov chains. Theory Probab. Appl. **23**, 279–295 (1978)
8. Pitman, J.W.: Uniform rates of convergence for Markov chain transition probabilities. Z. Wahrscheinlichkeitstheorie verw. Gebiete **29**, 193–227 (1974)
9. Pitman, J.W.: Occupation measures for Markov chains. Adv. in Appl. Probab. **9**, 69–86 (1977)
10. Rosenblatt, M.: Markov processes. Structure and asymptotic behaviour. Berlin, Heidelberg, New York: Springer 1971

Received January 15, 1980