

OCCUPATION TIMES FOR MARKOV AND SEMI-MARKOV CHAINS

BY

HARRY KESTEN⁽¹⁾

1. Introduction. We start with a recurrent, irreducible Markov chain Y_0, Y_1, \dots with a denumerable state space, say the integers, and n step transition probabilities $P_{i,j}^{(n)}$, and define a sequence T_0, T_1, \dots of non-negative random variables which represent the time spent in the i th state. More precisely we assume that there exists a set of distribution functions $H_j(t), j=0, \pm 1, \dots$ such that⁽²⁾

$$P\{T_0 \leq t_0, \dots, T_k \leq t_k \mid Y_0 = j_0, Y_1 = j_1, Y_2 = j_2, \dots\} = \prod_{i=0}^k H_{j_i}(t_i).$$

We then put

$$Y(t) = Y_k \quad \text{if} \quad \sum_{i=0}^{k-1} T_i < t \leq \sum_{i=0}^k T_i.$$

The process $Y(t)$ jumps at the time $\sum_{i=0}^k T_i, k=0, 1, \dots$, and is constant between jumps. $Y(t)$ is a semi-Markov chain. Our definition is somewhat narrower than the one of Smith in [15] but the case of [15] can be reduced to ours (cf. remarks after (3.4) and after Theorem 3). If the functions $H_j(t)$ are exponential distributions then $Y(t)$ becomes an ordinary Markov chain with continuous time. If $H_j(t)$ gives probability one to the value one then $1 = T_0 = T_1 = \dots$ and the process becomes equivalent to a Markov chain with discrete time.

Let $V(j)$ be some function on the integers. We shall be interested in the occupation time problem i.e. the study of the limiting behavior of $\sum_{i=0}^{\lambda} V(Y_i)$ and $\int_0^{\lambda} V(Y(t))dt$. In particular if V is the characteristic function of a set J of integers

$$\beta(\lambda) = \int_0^{\lambda} V(Y(t))dt$$

is the occupation time of J up till time λ .

For semi-Markov chains with 2 possible states $\beta(\lambda)$ was studied by Smith [15], Lamperti [13] and Takács [16]. Pyke [14] like us, studied the general

Presented to the Society, February 7, 1961; received by the editors May 23, 1961.

⁽¹⁾ "The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research of the Air Research and Development Command United States Air Force through its European office."

⁽²⁾ $P\{A\}$ = probability of the event A ; $P\{A|B\}$ = conditional probability of the event A given B .

case, and obtained results related to ours. For bona fide Markov processes the general problem of $\int_0^\lambda V(Y(t))dt$ was studied by Darling and Kac [2] for $V(j) \geq 0$. In particular, in the discrete time case, if

$$(1.1) \quad \lim_{z \uparrow 1} \frac{(1-z)^\alpha}{L\left(\frac{1}{1-z}\right)} \sum_{n=0}^{\infty} \sum_j P_{i,j}^{(n)} z^n V(j) = C > 0$$

for some $0 \leq \alpha \leq 1$, slowly varying L , and all i with $V(i) > 0$, they obtained for the limiting distribution of

$$(1.2) \quad \lambda^{-\alpha} L^{-1}(\lambda) \sum_{i=0}^{\lambda} V(Y_i)$$

a Mittag-Leffler distribution, and an analogous result for the continuous time case. For positive recurrent Markov chains (discrete time) the occupation time problem has been practically solved by Doeblin [4], Chung [1] and Kendall [11]. The only interesting case left, as far as Markov chains are concerned, seems therefore if V can be negative and $C=0$ in (1.1) (and if the chain is null recurrent for discrete time). Dobrusin [6] studied this problem for the simple random walk on the integers i.e.

$$P\{Y_{i+1} - Y_i = +1\} = P\{Y_{i+1} - Y_i = -1\} = \frac{1}{2}.$$

We shall follow the lines of [16]. First we extend the results of [16] to semi-Markov chains with more than 2 states. By specialization to Markov chains we then obtain extensions of [2] and [6] which apply also if C in (1.1) equals zero. Unfortunately, as Dobrusin, we have to restrict ourselves to functions V which vanish outside a *finite* set J . For finite J we obtain a result of the form

$$P\left\{\frac{\beta(\lambda)}{g(\lambda)} \leq x\right\} \rightarrow \int_0^\infty F(xy^{-\alpha}(1-xD)^{-\alpha\beta}) dG_1(y)$$

(Theorem 1) where $g(\lambda)$, α , β , D are appropriately chosen, F and G_1 are distribution functions defined in terms of the distributions $H_j(t)$ of the times T_0, T_1, \dots . If $M(\lambda)$ denotes the number of visits to J up till time λ and $G_2(y)$ its limiting distribution function after proper normalization, then under appropriate conditions (Theorem 2)

$$(1.3) \quad \lim_{\lambda \rightarrow \infty} P\left\{\frac{\int_0^\lambda V(Y(t))dt - M(\lambda)B}{g(\lambda)} \leq x\right\} = \int_0^\infty \phi(xy^{-1/2}) dG_2(y)$$

$g(\lambda)$ is a normalization function, ϕ the standard normal distribution and B a certain constant. In the case of a discrete time Markov chain for instance (Theorem 3) $\int_0^\lambda V(Y(t))dt$ becomes $\sum_{i=0}^\lambda V(Y_i)$, G_2 will be a Mittag-Leffler distribution and if C in (1.1) equals zero so does B in (1.3). In case $B=0$ (1.3) takes an especially simple form and this is undoubtedly the most interesting case for (1.3).

The method of proof is in all cases the same. The random variable under consideration is shown to be the sum of a random number of random variables. In §2 we derive the limiting distributions of such a sum if the number of terms is "almost independent" of the summands themselves. The main lemmas (§3) show that in our situation "almost independence" really holds while in §4 we derive the theorems fairly quickly from the lemmas. We end with an example.

We have tried to obtain relatively simple forms of the theorems by not insisting on full generality. In principle our methods work for sums of the form $\sum V(Y_i, Y_{i+1}, T_i)$ as in [14]. To a certain extent one could, as in [16], also allow *non*-identical distributions for the consecutive times spent in a fixed state. This would require to strengthen the conditions (4.6) and (4.7). On the other hand it would be desirable to prove our results for functions $V(j)$ different from zero for *infinitely* many states. The author has been unable to prove anything more in this direction than what is mentioned in the remark to Theorem 3. The results of this paper are uninteresting if the recurrence times for the states of the process Y_i have finite expectations (like e.g. in a positive recurrent Markov chain). This case has been treated extensively in [14] however. Neither do our methods seem to work for a process with a nondiscrete state space. In this case one probably has to revert to the method of moments, similar to [2].

Someone interested in the results only could best skip §2, read the introductory part of §3 (up till Lemma 2), and then the theorems in the last section.

The author wishes to thank Professor S. Karlin and Dr. C. J. Stone for several helpful discussions. Reference [16] was brought to the author's attention by Professor Karlin, whereas Dr. Stone suggested the usefulness of the imbedded Markov chain.

2. Sums of a random number of random variables. The result of this section is quite intuitive. If the random variables X_0, X_1, \dots and N would be strictly independent Lemma 1 is almost obvious (cf. [5]). The lemma below states that the same formulae hold if N is "almost independent" of X_0, X_1, \dots .

Let Z_0, Z_1, \dots be an irreducible finite Markov chain (cf. [8, Chapter 15] for terminology of Markov chains) with state space $\{0, 1, \dots, r\}$ and let there be given $(r+1)$ distribution functions $F_0(x), F_1(x), \dots, F_r(x)$. We shall consider a sequence of random variables X_0, X_1, \dots satisfying⁽²⁾

$$(2.1) \quad \begin{aligned} &P\{X_0 \leq x_0, X_1 \leq x_1, \dots, X_k \leq x_k \mid Z_0 = j_0, Z_1 = j_1, Z_2 = j_2, \dots\} \\ &= \prod_{s=0}^k F_{j_s}(x_s). \end{aligned}$$

I.e., given Z_0, Z_1, \dots the X 's are independent and

$$P\{X_i \leq x_i \mid Z_i = j_i\} = F_{j_i}(x_i).$$

Repeated use will be made of (positive) normalizing functions $f(k)$ and $h(\lambda)$ which in all applications are of the form

$$(2.2) \quad f(k) = k^\alpha L_1(k), \quad \alpha > 0$$

and

$$(2.3) \quad h(\lambda) = \lambda^\beta L_2(\lambda), \quad \beta \geq 0, \quad h(\lambda) \rightarrow \infty \text{ as } \lambda \rightarrow \infty$$

where $L_1(k)$ and $L_2(\lambda)$ are slowly varying functions. $L(t)$ is called slowly varying if it is continuous and if $L(ct)/L(t) \rightarrow 1$ as $t \rightarrow \infty$ for all $c > 0$. Notice that it follows from the properties of slowly varying functions that [7, p. 400 and the references cited there]

$$(2.4) \quad \frac{f(k_2)}{f(k_1)} \rightarrow 0 \quad \text{as} \quad \frac{k_2}{k_1} \rightarrow 0, \quad k_2 \geq 1.$$

Even though several results hold for more general functions we shall restrict ourselves to this class.

$F(\cdot)$, $G(\cdot)$ and $H(\cdot)$ with or without indices always stand for distribution functions and convergence of distribution functions is always meant in the weak sense. Thus expressions like

$$\lim_{k \rightarrow \infty} P\left\{\sum_{i=0}^k X_i \leq x\right\} = F(x)$$

are only meant to hold for the continuity points of $F(x)$.

LEMMA 1. *Let Z_0, Z_1, \dots be an irreducible finite Markov chain and X_0, X_1, \dots a sequence of random variables satisfying (2.1) and let $N(\lambda)$ be a non-negative, integer valued, random variable such that⁽³⁾*

$$(2.5) \quad \begin{aligned} &P\{N(\lambda) \leq k \mid X_0, X_1, X_2, \dots, Z_0, Z_1, Z_2, \dots\} \\ &= P\{N(\lambda) \leq k \mid X_0, \dots, X_k, Z_0, \dots, Z_k\}. \end{aligned}$$

If there exist functions $f(k)$ and $h(\lambda)$ of the form (2.2), (2.3) such that⁽⁴⁾

⁽³⁾ In the usual terminology, one is given a stopping rule, which determines when to stop observing the process. The decision to stop or continue after k steps depends on the past k steps only and not on the future.

⁽⁴⁾ In $\lim_{k \rightarrow \infty; \lambda \rightarrow \infty}$, both k and λ go to infinity in any (dependent or independent) manner.

$$(2.6) \quad \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{f(k)} \leq x \text{ \& } \frac{N(\lambda)}{h(\lambda)} \leq y \right\} = F(x) \cdot G(y)$$

then⁽⁵⁾

$$(2.7) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{N(\lambda)} X_i}{f(h(\lambda))} \leq x \right\} = H(x) = \int_{0-}^{\infty} F(xy^{-\alpha}) dG(y).$$

If (2.5) holds and

$$(2.8) \quad \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - ak}{bk^{\alpha}} \leq x \text{ \& } \frac{N(\lambda) - c\lambda}{d\lambda^{\beta}} \leq y \right\} = F(x) \cdot G(y)$$

with $ac \neq 0$, $0 < \alpha, \beta < 1$, then⁽⁶⁾

$$(2.9) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{N(\lambda)} X_i - ac\lambda}{\lambda^{\max(\alpha, \beta)}} \leq x \right\} = \begin{cases} F(xb^{-1}c^{-\alpha}) & \text{if } \alpha > \beta, \\ F(xb^{-1}c^{-\alpha}) * G(xa^{-1}d^{-1}) & \text{if } \alpha = \beta, \\ G(xa^{-1}d^{-1}) & \text{if } \alpha < \beta. \end{cases}$$

Proof. Let us prove (2.7). From (2.6)

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{f(k)} \leq x \right\} = F(x).$$

This is of course independent of the value of Z_0 and hence, for any $0 \leq j \leq r$

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{f(k)} \leq x \mid Z_0 = j \right\} = F(x).$$

Thus by (2.4) there exists for every $\epsilon > 0$ a $\delta > 0$ such that (for any $0 \leq j \leq r$)

$$(2.10) \quad P \left\{ \left| \frac{\sum_{i=1}^{k_2} X_i}{f(k_1)} \right| \geq \epsilon \mid Z_0 = j \right\} = P \left\{ \left| \frac{\sum_{i=1}^{k_2} X_i}{f(k_2)} \right| \geq \epsilon \frac{f(k_1)}{f(k_2)} \mid Z_0 = j \right\} \leq \epsilon$$

whenever $k_2/k_1 \leq \delta$, $k_2 \geq 1$. Choose δ so as to satisfy (2.10) and also such that the points $i\delta$, $i = 1, 2, \dots$ are continuity points of $G(y)$. Define a new ran-

⁽⁵⁾ $F(x0^{-1})$ is taken to be one if $x \geq 0$ and zero if $x < 0$.

⁽⁶⁾ $F(xb^{-1}c^{-\alpha}) * G(xa^{-1}d^{-1})$ is the convolution $\int_{-\infty}^{\infty} F((x-y)b^{-1}c^{-\alpha}) d_y G(ya^{-1}d^{-1})$. The same notation for convolution is used in (3.38) and (4.50).

dom variable $M(\lambda)$ by

$$(2.11) \quad \begin{cases} M(\lambda) = [\delta h(\lambda)] & \text{if } N(\lambda) \leq \delta h(\lambda), \\ M(\lambda) = [k\delta h(\lambda)] & \text{if } (k-1)\delta h(\lambda) < N(\lambda) \leq k\delta h(\lambda), \quad k = 2, 3, \dots \end{cases}$$

Thus

$$(2.12) \quad 0 \leq M(\lambda) - N(\lambda) \leq \delta h(\lambda)$$

and by (2.5) the conditional distribution of $\sum_{i=N(\lambda)+1}^{M(\lambda)} X_i$, given $Z_{N(\lambda)} = j$ is the same as of $\sum_{i=1}^{M(\lambda)-N(\lambda)} X_i$ given $Z_0 = j$. Since $Z_{N(\lambda)}$ can take only the values $0, 1, \dots, r$ it follows from (2.10) and (2.12) that

$$(2.13) \quad P \left\{ \left| \frac{\sum_{i=0}^{M(\lambda)} X_i - \sum_{i=0}^{N(\lambda)} X_i}{f(h(\lambda))} \right| \geq \epsilon \right\} \leq \epsilon.$$

We may therefore restrict ourselves to the study of $\{f(h(\lambda))\}^{-1} \sum_{i=0}^{M(\lambda)} X_i$. From now on x will be chosen a continuity point of $H(x)$ ($H(x)$ is obviously a distribution function). Put $k_i = [i\delta h(\lambda)]$. By (2.6) we can choose an A such that for all λ

$$\sum_{j=A+1}^{\infty} P\{M(\lambda) = k_j\} \leq \epsilon.$$

Then

$$(2.14) \quad \begin{aligned} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} X_i}{f(h(\lambda))} \leq x \right\} &= \sum_{j=1}^A P \left\{ \frac{\sum_{i=0}^{k_j} X_i}{f(k_j)} \leq \frac{xf(h(\lambda))}{f(k_j)} \& M(\lambda) = k_j \right\} + \theta\epsilon \\ &= \sum_{j=1}^A P \left\{ \frac{\sum_{i=0}^{k_j} X_i}{f(k_j)} \leq \frac{xf(h(\lambda))}{f(k_j)} \& (j-1)\delta < \frac{N(\lambda)}{h(\lambda)} \leq j\delta \right\} + \theta\epsilon \end{aligned}$$

where $|\theta| \leq 1$. Letting λ tend to infinity we get by (2.2) and (2.6)

$$\begin{aligned} \sum_{j=1}^{\infty} F \left(\frac{x - \epsilon}{(j\delta)^\alpha} \right) \{G(j\delta) - G((j-1)\delta)\} + F \left(\frac{x - \epsilon}{\delta^\alpha} \right) G(0) - 2\epsilon \\ \leq \liminf_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} X_i}{f(h(\lambda))} \leq x \right\} \leq \limsup_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} X_i}{f(h(\lambda))} \leq x \right\} \\ \leq \sum_{j=1}^{\infty} F \left(\frac{x + \epsilon}{(j\delta)^\alpha} \right) \{G(j\delta) - G((j-1)\delta)\} + F \left(\frac{x + \epsilon}{\delta^\alpha} \right) G(0) + 2\epsilon. \end{aligned}$$

Letting first $\delta \downarrow 0$, then $\epsilon \downarrow 0$ and using (2.13) (2.7) follows. The proof of (2.9) is very similar if we take into account that $\sum_{i=0}^{N(\lambda)} X_i - ac\lambda = \sum_{i=0}^{N(\lambda)} (X_i - a) + a(N(\lambda) - c\lambda)$. The details can easily be supplied.

This lemma is quite analogous to the result of [5]. As in [5] one might also want to consider the cases $a=0, c \neq 0$ and $a \neq 0, c=0$ in (2.8). These, however, do not lead to anything more than the first part of our lemma. E.g. if $a=0, c \neq 0$, Dobrusin [5] also only uses the consequence " $N(\lambda)/c\lambda \rightarrow 1$ in probability" of (2.8).

COROLLARY. *If Z_0, Z_1, \dots is a finite irreducible Markov chain and X_0, X_1, \dots a sequence of random variables satisfying (2.1) and, as $k \rightarrow \infty$ $\sum_{i=0}^k X_i/k \rightarrow C \neq 0$ in probability and if (2.5) holds and*

$$\lim_{\lambda \rightarrow \infty} P \left\{ \frac{N(\lambda)}{h(\lambda)} \leq y \right\} = G(y),$$

then

$$(2.15) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{N(\lambda)} X_i}{h(\lambda)} \leq x \right\} = \begin{cases} G(xC^{-1}) & \text{if } C > 0, \\ 1 - G(xC^{-1}) & \text{if } C < 0. \end{cases}$$

Proof. Putting

$$(2.16) \quad \bar{F}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

one has

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{k} \leq x \right\} = \bar{F}(x - C).$$

Since \bar{F} can take only the values zero and one, one easily verifies

$$\begin{aligned} & \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{k} \leq x \text{ \& } \frac{N(\lambda)}{h(\lambda)} \leq y \right\} \\ &= \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i}{k} \leq x \right\} \cdot P \left\{ \frac{N(\lambda)}{h(\lambda)} \leq y \mid \frac{\sum_{i=0}^k X_i}{k} \leq x \right\} \\ &= \bar{F}(x - C) \cdot G(y). \end{aligned}$$

(2.15) follows then from (2.7).

3. The main lemmas. In this section and the next one Y_0, Y_1, \dots will

be a denumerable irreducible recurrent Markov chain. (If one has a chain which is not irreducible one has to go through the usual method of separating the state space into several irreducible and one transient part.)

For convenience we identify the state space with the set of integers. $H_j(t)$ ($j=0, \pm 1, \dots$) will be a set of distribution functions and T_0, T_1, \dots a sequence of non-negative random variables satisfying

$$(3.1) \quad P\{T_0 \leq t_0, \dots, T_k \leq t_k \mid Y_0 = j_0, Y_1 = j_1, Y_2 = j_2, \dots\} = \prod_{s=0}^k H_{j_s}(t_s)$$

and

$$(3.2) \quad H_j(0-) = 0.$$

We construct a semimarkovian process $Y(t)$ by defining

$$(3.3) \quad \begin{cases} Y(0) = Y_0, \\ Y(t) = Y_k & \text{if } \sum_{i=0}^{k-1} T_i < t \leq \sum_{i=0}^k T_i. \end{cases}$$

Thus the process starts with $Y(0) = Y_0 = j_0$ for some j_0 . It remains a time T_0 in this state. Then it jumps to some state, say j_1 with probability $P\{Y_1 = j_1 \mid Y_0 = j_0\}$. Given j_0, T_0 and j_1 are independent by (3.1). It stays in j_1 a time T_1 whose distribution depends on j_1 and jumps then to some state j_2 , etc. In order to make (3.3) meaningful for every t we have to assume

$$(3.4) \quad \sum_{i=0}^{\infty} T_i \text{ diverges with probability 1.}$$

The definition of the $Y(t)$ process is almost the same as the definition of a regular semimarkovian process of Smith in [15]. Our definition may seem less general since given Y_i, T_i and Y_{i+1} are independent. The general case where the distribution of T_i may depend on Y_i and Y_{i+1} can be reduced to our case by considering the Markov chain of pairs, $Y'_i = (Y_i, Y_{i+1})$. With this definition T_i and Y'_{i+1} are independent if Y'_i is given (cf., however, the remarks after Theorem 3).

In addition to the T_0, T_1, \dots we consider a sequence of random variables V_0, V_1, \dots . We assume that there exist two dimensional distribution functions $G_j(x, t)$ ($j=0, \pm 1, \dots$) such that

$$(3.5) \quad \lim_{x \rightarrow \infty} G_j(x, t) = H_j(t)$$

and

$$(3.6) \quad P\{V_0 \leq x_0, \dots, V_k \leq x_k, T_0 \leq t_0, \dots, T_k \leq t_k \mid Y_0 = j_0, Y_1 = j_1, Y_2 = j_2, \dots\} = \prod_{s=0}^k G_{j_s}(x_s, t_s).$$

Thus, if the V 's are given the V 's are independent, the T 's are independent and V_i is independent of T_j $i \neq j$. V_i may, however, be dependent on T_i or even be a function of T_i .

We assume that there exists a *finite* set

$$J = \{j_0, j_1, \dots, j_r\}$$

such that V_j is zero with probability one for indices j outside J , i.e. (cf. (2.16))

$$(3.7) \quad \lim_{t \rightarrow \infty} G_j(x, t) = \bar{F}(x) \quad \text{for all } j \notin J.$$

Define the successive times

$$r_0 < r_1 < \dots$$

of visits to J by

$$(3.8) \quad Y_{r_j} \in J \quad \text{but} \quad Y_k \notin J \quad \text{if } k \text{ is not equal to some } r_j.$$

Put

$$(3.9) \quad Z_i = Y_{r_i},$$

$$(3.10) \quad \dagger X_i = V_{r_i}.$$

Z_0, Z_1, \dots is the imbedded Markov chain with state space J . Its transition probabilities are given by

$$(3.11) \quad \begin{aligned} Q_{u,v} &= P\{Z_{k+1} = j_v \mid Z_k = j_u\} \\ &= \sum_{n=1}^{\infty} P\{Y_n = j_v, Y_k \notin J \text{ for } 1 \leq k \leq n-1 \mid Y_0 = j_u\}. \end{aligned}$$

Since we assumed the original chain irreducible, so is the finite Markov chain Z_0, Z_1, \dots . If the stationary probability of j_u in the last chain is $q(u)$, then (cf. [8])

$$(3.12) \quad q(u) > 0, \quad \sum_{u=0}^r q(u) = 1, \quad \sum_{u=0}^r q(u) Q_{u,v} = q(v).$$

We introduce in addition

$$(3.13) \quad W_0 = \sum_{j=0}^{r_0-1} T_j \quad (=0 \text{ if } r_0 = 0),$$

$$(3.14) \quad W_i = \sum_{j=r_{i-1}+1}^{r_i-1} T_j \quad (=0 \text{ if } r_i = r_{i-1} + 1).$$

W_i is the time spent outside J between the i th and $(i+1)$ st visit to J . Finally, $K(\lambda)$ is defined by

$$(3.15) \quad \sum_{i=0}^{K(\lambda)} W_i < \lambda \leq \sum_{i=0}^{K(\lambda)+1} W_i.$$

We shall be interested in the limiting behavior of sums like

$$\sum_{i=0}^{K(\lambda)} X_i = \sum_{i=0}^{K(\lambda)} V_{r_i}.$$

LEMMA 2. If (3.5), (3.6) and (3.7) are satisfied and if there exist functions $f(k)$, $h(\lambda)$ of the form (2.2), (2.3) such that

$$(3.16) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \right\} = F(x)$$

and

$$(3.17) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \right\} = G(y) \quad (c \geq 0, \text{ since } K(\lambda) \geq 0)$$

and, in case $c \neq 0$ also

$$(3.18) \quad \frac{h(\lambda)}{\lambda} \rightarrow 0, \quad \frac{h(\lambda)}{\lambda^{1/2}} \rightarrow \infty,$$

then

$$(3.19) \quad \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \right\} = F(x) \cdot G(y).$$

Proof. We shall only prove the case where $c \neq 0$, the other case being quite similar. Consider the imbedded Markov chain Z_0, Z_1, \dots and let

$$(3.20) \quad n_k(u, v) = \text{number of indices } i, 1 \leq i \leq k, \text{ for which } Z_{i-1} = j_u, Z_i = j_v.$$

Define

$$A(k) = \left(\frac{h(k)}{k^{1/2}} \right)^{1/2}.$$

By (3.18) $A(k) \rightarrow \infty$ as $k \rightarrow \infty$ and consequently (e.g. by Theorem 4 in [1]) there exists for every $\epsilon > 0$ a k_0 such that

$$(3.21) \quad P \{ q(u) Q_{u,v} k (1 - A(k) k^{-1/2}) \leq n_k(u, v) \leq q(u) Q_{u,v} k (1 + A(k) k^{-1/2}) \} \geq 1 - \epsilon/r^2$$

whenever $k \geq k_0$. Introduce the abbreviation

$$(3.22) \quad s = s(y, \lambda) = [c\lambda + yh(\lambda)] + 1$$

and the event

$$(3.23) \quad E(x, y, k, \lambda) = \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \text{ \& } |n_s - sq(u)Q_{u,v}| \leq q(u)Q_{u,v}A(s)s^{1/2} \text{ for } 0 \leq u, v \leq r \right\}.$$

Then, if $s \geq k_0(|\theta_i| \leq 1)$

$$(3.24) \quad \begin{aligned} & P \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \right\} \\ &= P \left\{ E(x, y, k, \lambda) \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \right\} + \theta_1 \epsilon \\ &= P \{ E(x, y, k, \lambda) \} \cdot P \left\{ \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \mid E(x, y, k, \lambda) \right\} + \theta_1 \epsilon \\ &= P \{ E(x, y, k, \lambda) \} \cdot P \left\{ \sum_{i=0}^{s(y, \lambda)} W_i \geq \lambda \mid E(x, y, k, \lambda) \right\} + \theta_1 \epsilon \\ &= P \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \right\} \cdot P \left\{ \sum_{i=0}^{s(y, \lambda)} W_i \geq \lambda \mid E(x, y, k, \lambda) \right\} + 2\theta_2 \epsilon. \end{aligned}$$

Moreover, because of (3.6), (3.7) and (3.10) one has

$$(3.25) \quad \begin{aligned} & P \left\{ \sum_{i=1}^s W_i \geq \lambda \mid X_0, X_1, \dots, n_1(u, v), n_2(u, v), \dots, 0 \leq u, v \leq r \right\} \\ &= P \left\{ \sum_{i=1}^s W_i \geq \lambda \mid n_s(u, v) \text{ } 0 \leq u, v \leq r \right\}. \end{aligned}$$

In fact, write $H_{u,v}(x)$ for the conditional distribution of W_i given that $Y_{r_{i-1}} = j_u, Y_{r_i} = j_v$ i.e.

$$\begin{aligned} H_{u,v}(x) &= \frac{1}{Q_{u,v}} \sum_{n=1}^{\infty} P \left\{ \sum_{i=1}^{n-1} T_i \leq x, Y_n = j_v, Y_k \notin J \text{ for } 1 \leq k \leq n-1 \mid Y_0 = j_u \right\} \\ &= P \{ W_1 \leq x \mid Z_0 = j_u, Z_1 = j_v \}. \end{aligned}$$

Given $n_s(u, v)$, $0 \leq u, v \leq r$, $\sum_{i=1}^{s(y, \lambda)} W_i$ is the sum of $\sum_{0 \leq u, v \leq r} n_s(u, v)$ non-negative random variables. These random variables are *independent* of W_0 ,

each other and of X_0, X_1, \dots . Exactly $n_s(u, v)$ of these random variables have distribution function $H_{u,v}(x)$. Let $\bar{n}(u, v)$ be integers such that

$$|\bar{n}(u, v) - sq(u)Q_{u,v}| \leq q(u)Q_{u,v}A(s)s^{1/2}.$$

Then, by (3.21), if

$$t = t(y, \lambda) = s(y, \lambda) - 3A(s(y, \lambda))(s(y, \lambda))^{1/2} \geq k_0,$$

$$P\{n_i(u, v) \leq \bar{n}(u, v), 0 \leq u, v \leq r\} \geq 1 - \epsilon$$

and hence, by the above remarks

$$\begin{aligned} P\left\{\sum_{i=0}^t W_i \geq \lambda\right\} &\leq \epsilon + P\left\{\sum_{i=0}^t W_i \geq \lambda \mid n_i(u, v) \leq \bar{n}(u, v), 0 \leq u, v \leq r\right\} \\ (3.26) \qquad &\leq \epsilon + P\left\{\sum_{i=0}^t W_i \geq \lambda \mid n_s(u, v) = \bar{n}(u, v), 0 \leq u, v \leq r\right\}. \end{aligned}$$

Since, by (3.18)

$$A(s(y, \lambda))(s(y, \lambda))^{1/2} = o(h(\lambda)), \qquad \lambda \rightarrow \infty,$$

$$t(y, \lambda) \geq c\lambda + (y - \epsilon)h(\lambda) + 1$$

for sufficiently large λ .

Combining (3.21)–(3.26) we have

$$\begin{aligned} P\left\{\sum_{i=0}^{s(y, \lambda)} W_i \geq \lambda \mid E(x, y, k, \lambda)\right\} &\geq P\left\{\sum_{i=0}^{t(y, \lambda)} W_i \geq \lambda\right\} - \epsilon \\ (3.27) \qquad &\geq P\{K(\lambda) \leq t(y, \lambda) - 1\} - \epsilon \\ &\geq P\left\{\frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y - \epsilon\right\} - \epsilon. \end{aligned}$$

Together with (3.24) this implies

$$\begin{aligned} P\left\{\frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y\right\} \\ (3.28) \qquad &\geq P\left\{\frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x\right\} \cdot P\left\{\frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y - \epsilon\right\} - 3\epsilon. \end{aligned}$$

If x, y is a continuity point of $F(x) \cdot G(y)$ we get as $k \rightarrow \infty, \lambda \rightarrow \infty, \epsilon \downarrow 0$,

$$\liminf_{k \rightarrow \infty; \lambda \rightarrow \infty} P\left\{\frac{\sum_{i=0}^k X_i - ak}{f(k)} \leq x \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y\right\} \geq F(x) \cdot G(y).$$

In a similar way one obtains the same upper bound for

$$\limsup_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - ak}{f(x)} \leq x \text{ \& } \frac{K(\lambda) - c\lambda}{h(\lambda)} \leq y \right\},$$

thus completing the proof.

In the next section we shall make use of the law of large numbers and a theorem concerning limit distributions for Markov chains in a form slightly more general than usually given. We formulate it below for finite Markov chains. As in §2 let Z_0, Z_1, \dots be a finite, irreducible Markov chain with state space $\{0, 1, \dots, r\}$ and stationary probability $q(j)$ of state j , and let X_0, X_1, \dots be a sequence of random variables satisfying (2.1). We introduce

$$(3.29) \quad \mu_j = \int_{-\infty}^{+\infty} x dF_j(x),$$

$$(3.30) \quad C = \sum_{j=0}^r q(j) \mu_j,$$

$$(3.31) \quad \sigma_j^2 = \int_{-\infty}^{+\infty} (x - C)^2 dF_j(x)$$

and

$$(3.32) \quad \begin{aligned} \sigma^2 &= \sum_{j=0}^r q(j) \sigma_j^2 + 2 \sum_{j=1}^r \sum_{k=0}^r (m_{j0} + m_{0k} - m_{jk}) q(j) q(k) (\mu_j - C) (\mu_k - C) \\ &= \sum_{j=0}^r q(j) \sigma_j^2 - 2 \sum_{j=0}^r \sum_{k=0}^r q(j) (\mu_j - C) m_{jk} q(k) (\mu_k - C) \end{aligned}$$

where

$$(3.33) \quad \begin{aligned} m_{i,j} &= \sum_{n=1}^{\infty} n P\{Z_n = j, Z_k \neq j \text{ for } 1 \leq k \leq n-1 \mid Z_0 = i\} \\ &= \text{Expected first passage time from } i \text{ to } j. \end{aligned}$$

The equality of the two expressions for σ^2 in (3.32) follows from the fact $\sum_{j=0}^r q(j) (\mu_j - C) = 0$. Notice that $m_{i,j}$ can be computed by means of sections 4.4.7 and 5.1 (pp. 79 and 102) of [10].

LEMMA 3. *If the integrals (3.29) converge absolutely, then with probability 1,*

$$(3.34) \quad \lim_{k \rightarrow \infty} \frac{\sum_{i=0}^k X_i}{k} = C.$$

If in addition $\sigma_j^2 < \infty$, $j=0, \dots, r$, then

$$(3.35) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - kC}{\sigma k^{1/2}} \leq x \right\} = \phi(x) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^x e^{-t^2/2} dt.$$

(If $\sigma = 0$, (3.35) means that its left hand side without the factor σ tends to zero in probability.)

Let $U_0^{(j)}, U_1^{(j)}, \dots$ be independent random variables, each with distribution function $F_j(x)$. If there exist $\alpha_j > 0$, constants C_j and distribution functions $G_j(x)$ such that for $0 \leq j \leq r$,

$$(3.36) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k U_i^{(j)} - kC_j}{k^{\alpha_j}} \leq x \right\} = G_j(x)$$

with

$$\alpha_{j_1} = \alpha_{j_2} = \dots = \alpha_{j_s} = \max_{0 \leq j \leq r} \alpha_j = \alpha$$

(but $\alpha_j < \alpha$ if j not one of j_1, \dots, j_s) and if

$$(3.37) \quad \alpha > \frac{1}{2} \text{ in case at least one } C_j \neq 0,$$

then⁽⁸⁾

$$(3.38) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k X_i - k \sum_{j=0}^r q(j)C_j}{k^\alpha} \leq x \right\} = G_{j_1}(x(q(j_1))^{-\alpha}) * \dots * G_{j_s}(x(q(j_s))^{-\alpha}).$$

Proof. (3.34) is proved exactly as Theorem 2 or 6 in [1]. Similarly (3.35) is proved as the central limit theorem in [4], [1] or [11]. The expression for σ^2 in (3.32) was given in [1]. (3.38) we prove again only if some $C_j \neq 0$. Let $n_k(j)$ = number of indices i , $0 \leq i \leq k$, for which $Z_i = j$. Then, with

$$(3.39) \quad \frac{1}{2} < \delta = \min \left(\frac{3}{4}, \frac{\alpha}{2} + \frac{1}{4} \right) < \min(1, \alpha),$$

$$\lim_{k \rightarrow \infty} P \{ |n_k(j) - q(j)k| > k^\delta \} = 0$$

and consequently, with probability close to one,

$$\sum_{i=0}^k X_i - k \sum_{j=0}^r q(j)C_j = \sum_{j=0}^r \sum_{(j)} (X_i - C_j) + \theta k^\delta$$

where $\sum_{(j)}$ runs over those $i \leq k$ for which $Z_i = j$ and $|\theta| \leq 1$. By (3.39) however $k^{-\alpha} \sum_{(j)} (X_i - C_i)$ has the same limiting distribution as

$$(3.40) \quad k^{-\alpha} \left\{ \sum_{i=0}^{q(j)k+k^\delta} U_i^{(j)} - (q(j)k + k^\delta) C_j \right\}$$

(compare the argument leading to (2.13) from (2.4)). Since the limiting distribution of (3.40) is $G_j(x(q(j))^{-\alpha})$ if $\alpha_j = \alpha$ and $\bar{F}(x)$ (cf. (2.16)) if $\alpha_j < \alpha$, (3.38) follows.

REMARK 1. For applications it is useful to have the following generalization. Instead of (2.1) assume there exist distribution functions $F_{i,j}(x)$ such that

$$(3.41) \quad P\{X_0 \leq x_0, X_1 \leq x_1, \dots, X_k \leq x_k \mid Z_0 = j_0, Z_1 = j_1, Z_2 = j_2, \dots\} \\ = \prod_{s=0}^k F_{j_s, j_{s+1}}(x_s).$$

I.e., given the Z 's the X 's are independent but now the distribution of X_i depends on Z_i and Z_{i+1} ,

$$P\{X_i \leq x \mid Z_i = j_i, Z_{i+1} = j_{i+1}\} = F_{j_i, j_{i+1}}(x).$$

Put

$$F_j(x) = P\{X_i \leq x \mid Z_i = j\} = \sum_{k=0}^r P\{Z_{i+1} = k \mid Z_i = j\} F_{j,k}(x).$$

With this definition one can prove by a somewhat more complicated argument that (3.36) implies (3.38) whenever $\alpha > 1/2$.

REMARK 2. Obviously only slight modifications are necessary to allow normalizing functions $k^{\alpha_j} L_j(k)$, L_j slowly varying, in (3.36).

4. Applications to Markov and semi-Markov chains. We assume as in §3 that Y_0, Y_1, \dots is an irreducible, recurrent Markov chain with state space the integers and n step transition probabilities $P_{ij}^{(n)}$. $Y(t)$ is defined by (3.3) and (3.1)–(3.4) are assumed to hold. It was proved in [4] (cf. also [1] or [3]) that there exists a set of finite, positive numbers $\{p_i\}$ such that

$$(4.1) \quad \lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N p_{ii}^{(n)}}{\sum_{n=1}^N p_{jj}^{(n)}} = \frac{p_i}{p_j}.$$

The stationary probabilities of the imbedded Markov chain Z_0, Z_1, \dots (cf. (3.8)–(3.12)) are then given by [3]

$$(4.2) \quad q(u) = \frac{p_{j_u}}{\sum_{v=0}^r p_{j_v}}.$$

We first discuss occupation times for semi-Markov chains and specialize

later to obtain results for ordinary Markov chains. Put

$$(4.3) \quad U(j) = \begin{cases} 1 & \text{if } j \in J = \{j_0, j_1, \dots, j_r\}, \\ 0 & \text{if } j \notin J, \end{cases}$$

and consider

$$(4.4) \quad \beta(\lambda) = \int_0^\lambda U(Y(t))dt = \text{time spent in } J \text{ up till } \lambda.$$

An ingenious argument of Takács makes Lemma 2 applicable to this problem. The following lemma is Lemma 2 in [16]. Notice that it has nothing to do with the semimarkovian character of the Y process. Takács simply proved the events between braces in the left and right hand side of the equation in Lemma 4 to be identical (cf. (3.1), (3.8) and (3.15) for the definitions of T_{r_i} and $K(\lambda)$).

LEMMA 4 (TAKÁCS).

$$\{\beta(\lambda) \leq x\} = \left\{ \sum_{i=0}^{K(\lambda-x)} T_{r_i} \leq x \right\}.$$

Before formulating the first theorem we want to indicate how Lemma 3 can be used to verify conditions (4.6) and (4.7) of this theorem. T_{r_i} is the time spent in J during the $(i+1)$ st visit to J . If Z_0, Z_1, \dots is the imbedded Markov chain with state space J (cf. (3.9)) then the distribution of T_{r_i} depends on the value of Z_i only. Given $Z_i = j_u$ the distribution function of T_{r_i} is $H_{j_u}(t)$ and T_{r_i} is independent of $\{T_{r_j}; j \neq i\}$. Hence by Lemma 3 it suffices for (4.6) that

$$\lim_{k \rightarrow \infty} \mathbb{P} \left\{ \frac{\sum_{i=0}^k U_i^{(u)}}{k^{\alpha_u}} \leq x \right\}$$

exists for $u=0, 1, \dots, r$ where $U_i^{(u)}, i=0, 1, \dots$, are independent random variables each with distribution function $H_{j_u}(t)$. In other words (4.6) will be satisfied if the $H_{j_u}(t)$ belong to a suitable domain of attraction. Similarly W_i is the time spent outside J between the i th and $(i+1)$ st visit to J . Its distribution depends on the values of Z_{i-1} and Z_i only. Given that $Z_{i-1} = j_u$ and $Z_i = j_v$, W_i is independent of $\{W_j; j \neq i\}$ and its distribution function, $F_{u,v}(x)$ say, does not depend on i . Since (4.7) is equivalent to

$$\lim_{\lambda \rightarrow \infty} P \left\{ \sum_{i=0}^{[\nu h(\lambda)]+1} W_i \geq \lambda \right\} = G(y),$$

we see that (4.7) will follow from Lemma 3 (applied to the chain of pairs (Z_{i-1}, Z_i)), if

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k W_i^{(u,v)}}{k^{\alpha_{u,v}}} \leq x \right\}$$

exists for $u, v = 0, 1, \dots, r$, where the $W_i^{(u,v)}$ $i = 0, 1, \dots$, are independent random variables with distribution function $F_{u,v}(x)$. By Remark 1 to Lemma 3 it even suffices, in case $\max \alpha_u > 1/2$, if for $u = 0, 1, \dots, r$

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k W_i^{(u)}}{k^{\alpha_u}} \leq x \right\}$$

exists, where $W_i^{(u)}$, $i = 0, 1, \dots$ are independent random variables with

$$P\{W_i^{(u)} \leq x\} = P\{W_i \leq x \mid Z_{i-1} = j_u\}.$$

A slightly different method for finding the limiting distribution of $\sum W_i$ is used in the example at the end of this section. That method can be used whenever the Y process has a finite state space.

THEOREM 1. *If there exist functions $f(k)$ and $h(\lambda)$ of the form (2.2), (2.3), satisfying*

$$(4.5) \quad \lim_{\lambda \rightarrow \infty} \frac{f(h(\lambda))}{\lambda} = D \leq \infty$$

and if

$$(4.6) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k T_{r_i}}{f(k)} \leq x \right\} = F(x)$$

and

$$(4.7) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{K(\lambda)}{h(\lambda)} \leq y \right\} = G(y)$$

then, for $D < \infty$

$$(4.8) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\beta(\lambda)}{f(h(\lambda))} \leq x \right\} = \begin{cases} \int_{0-}^{\infty} F(xy^{-\alpha}(1-xD)^{-\alpha\beta}) dG(y) & \text{if } x < D^{-1}, \\ 1 & \text{if } x > D^{-1}, \end{cases}$$

($D^{-1} = \infty$ if $D = 0$). If $D = \infty$ and ^(*) $F(0+) = G(0+) = 0$, then

(*) If $F(0+)$ and $G(0+)$ are strictly less than one, it follows already from (4.6) and (4.7) that $F(0+) = G(0+) = 0$. This is therefore only a minor restriction.

$$(4.9) \quad \frac{\beta(\lambda)}{\lambda} \rightarrow 1 \text{ in probability } (\lambda \rightarrow \infty).$$

Proof. By Lemma 4

$$(4.10) \quad P\left\{\frac{\beta(\lambda)}{f(h(\lambda))} \leq x\right\} = P\left\{\frac{\sum_{i=0}^{K(\lambda - xf(h(\lambda)))} T_{r_i}}{f(h(\lambda))} \leq x\right\}.$$

Let us call $N(\lambda) = K(\lambda - xf(h(\lambda)))$. By (4.7), the fact that $K(\lambda)$ is a monotonic function of λ , and (2.3) one has for $x < D^{-1}$.

$$\lim_{\lambda \rightarrow \infty} P\left\{\frac{N(\lambda)}{h(\lambda)} \leq y\right\} = G(y(1 - xD)^{-\beta})$$

and, by Lemma 2

$$\lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P\left\{\frac{\sum_{i=0}^k T_{r_i}}{f(k)} \leq z \text{ \& } \frac{N(\lambda)}{h(\lambda)} \leq y\right\} = F(z) \cdot G(y(1 - xD)^{-\beta}).$$

Using Lemma 1 one obtains

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} P\left\{\frac{\sum_{i=0}^{N(\lambda)} T_{r_i}}{f(h(\lambda))} \leq z\right\} &= \int_{0-}^{\infty} F(zy^{-\alpha}) d_v G(y(1 - xD)^{-\beta}) \\ &= \int_{0-}^{\infty} F(zy^{-\alpha}(1 - xD)^{-\alpha\beta}) dG(y). \end{aligned}$$

Putting $z=x$, the first part of (4.8) follows while the second part follows if $x \uparrow D^{-1}$.

To prove (4.9), we note that always

$$0 \leq \frac{\beta(\lambda)}{\lambda} \leq 1,$$

while for $0 < x < 1$

$$P\left\{\frac{\beta(\lambda)}{\lambda} \leq x\right\} = P\left\{\frac{\sum_{i=0}^{K(\lambda(1-x))} T_{r_i}}{f(h(\lambda(1-x)))} \leq x \frac{\lambda}{f(h(\lambda(1-x)))}\right\} \rightarrow 0$$

($\lambda \rightarrow \infty$) because $\lambda \{f(h(\lambda(1-x)))\}^{-1} \rightarrow 0$ and $F(0+) = G(0+) = 0$.

REMARK. Clearly, Lemmas 1 and 2 allow us to treat also the case where (4.6) and (4.7) are replaced by

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k T_{r_i} - ak}{bk^\alpha} \leq x \right\} = F(x)$$

and

$$\lim_{\lambda \rightarrow \infty} P \left\{ \frac{K(\lambda) - c\lambda}{d\lambda^\beta} \leq y \right\} = G(y)$$

with $ac \neq 0$, $0 < \alpha < 1$, $1/2 < \beta < 1$. We shall not write out the results explicitly. The reader may easily find the appropriate limit distributions for such a case by using the tables on pp. 536, 537 in [16]. Theorem 1 was proved by Takács [16] if $Y(t)$ can take only two values and if

$$\|P_{i,j}\| = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

i.e. if the two states alternate. Lamperti [13] considered also 2 states only, with

$$\|P_{i,j}\| = \begin{vmatrix} p & q \\ p & q \end{vmatrix}.$$

One may consider the random variables W_i somewhat unpleasant to work with, since they only represent the time spent *outside* J . If, however, $\beta(\lambda)/\lambda$ tends to some constant a , $0 \leq a < 1$, in probability then one may replace W_i by $T_{i-1} + W_i$. Instead of $\beta(\lambda)$ one may even consider more general random variables in this case. This will be done in the next theorem.

Let $V(j)$ be a function which equals zero whenever j outside $J = \{j_0, \dots, j_r\}$ and define the following quantities (if the appropriate integrals exist):

$$(4.11) \quad C = \sum_{u=0}^r q(u)V(j_u), \quad (q(u) \text{ as in (4.2)}),$$

$$(4.12) \quad \sigma^2 = \sum_{u=0}^r q(u)(V(j_u) - C)^2 - 2 \sum_{u=0}^r \sum_{v=0}^r q(u)(V(j_u) - C)m_{j_u, j_v}(V(j_v) - C)$$

with $m_{i,j}$ as in (3.33),

$$(4.13) \quad \mu_u = V(j_u) \int_0^\infty x dH_{j_u}(x), \quad 0 \leq u \leq r,$$

$$(4.14) \quad B = \sum_{u=0}^r q(u)\mu_u,$$

$$(4.15) \quad \sigma_u^2 = \int_0^\infty (V(j_u)x - B)^2 dH_{j_u}(x), \quad 0 \leq u \leq r,$$

and

$$(4.16) \quad \rho^2 = \sum_{u=0}^r q(u) \sigma_u^2 - 2 \sum_{u=0}^r \sum_{v=0}^r q(u) (\mu_u - B) m_{j_u, j_v} q(v) (\mu_v - B)$$

with $m_{i,j}$ as in (3.33).

Instead of W_i we now introduce

$$(4.17) \quad \bar{W}_0 = W_0 \quad \text{and} \quad \bar{W}_i = \sum_{j=r_i-1}^{r_i-1} T_j, \quad i \geq 1,$$

and define $M(\lambda)$ by

$$(4.18) \quad \sum_{i=0}^{M(\lambda)} \bar{W}_i < \lambda \leq \sum_{i=0}^{M(\lambda)+1} \bar{W}_i.$$

$M(\lambda)+1$ is the number of visits to J of the $Y(t)$ process up till time λ .

As for checking condition (4.20) of the next theorem, the remark which was made before Theorem 1 concerning (4.6) and (4.7) applies almost word for word. Sometimes one can also find the limiting distribution of $M(\lambda)$ from that of $\beta(\lambda)$. In fact if the integrals in (4.13) converge and

$$B' = \sum_{u=0}^r q(u) \int_0^\infty x dH_{j_u}(x), \quad 0 \leq u \leq r,$$

then, by (3.34), with probability one

$$\lim_{k \rightarrow \infty} \frac{\sum_{i=0}^k U(Y_{r_i}) T_{r_i}}{k} \rightarrow B'.$$

Consequently, also (compare [1, p. 405])

$$\lim_{\lambda \rightarrow \infty} \frac{\beta(\lambda)}{M(\lambda)} = B' \text{ with probability 1}$$

and

$$(4.19) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{M(\lambda)}{h(\lambda)} \leq y \right\} = \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\beta(\lambda)}{h(\lambda)} \leq B'y \right\}.$$

THEOREM 2. *If $V(j)$ vanishes outside $J = \{j_0, \dots, j_r\}$ and*

$$(4.20) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{M(\lambda)}{h(\lambda)} \leq y \right\} = G(y)$$

for some function $h(\lambda)$ of the form (2.3), then for $C \neq 0$

$$(4.21) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} V(Y_{r_i})}{h(\lambda)} \leq x \right\} = \begin{cases} G(xC^{-1}) & \text{if } C > 0, \\ 1 - G(xC^{-1}) & \text{if } C < 0, \end{cases}$$

and if the integrals in (4.13) converge, for $B \neq 0$

$$(4.22) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\int_0^\lambda V(Y(t)) dt}{h(\lambda)} \leq x \right\} = \begin{cases} G(xB^{-1}) & \text{if } B > 0, \\ 1 - G(xB^{-1}) & \text{if } B < 0. \end{cases}$$

If (4.20) holds and if for some $0 \leq a < 1$

$$(4.23) \quad \frac{\beta(\lambda)}{\lambda} \rightarrow a \text{ in probability } (\lambda \rightarrow \infty)$$

then

$$(4.24) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} V(Y_{r_i}) - M(\lambda)C}{\sigma[h(\lambda)]^{1/2}} \leq x \right\} = \int_{0-}^{\infty} \phi(xy^{-1/2}) dG(y)$$

where σ^2 is given by (4.12) (with the same convention as in (3.35) if $\sigma = 0$).

If in addition to (4.20) and (4.23) the integrals in (4.13) and (4.15) converge, then

$$(4.25) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\int_0^\lambda V(Y(t)) dt - M(\lambda)B}{\rho[h(\lambda)]^{1/2}} \leq x \right\} = \int_{0-}^{\infty} \phi(xy^{-1/2}) dG(y)$$

where ρ^2 is given by (4.16) (with the same convention as in (3.35) if $\rho = 0$).

Proof. By Lemma 3

$$\frac{\sum_{i=0}^k V(Y_{r_i})}{k} \rightarrow C$$

with probability one. This, together with (4.20) implies (4.21) by the corollary to Lemma 1. Similarly, if the integrals in (4.13) converge

$$\frac{\sum_{i=0}^k V(Y_{r_i}) T_{r_i}}{k} \rightarrow B$$

with probability one. Consequently

$$\frac{\sum_{i=0}^{M(\lambda)-1} V(Y_{r_i}) T_{r_i}}{M(\lambda)} \rightarrow B$$

and

$$\frac{\sum_{i=0}^{M(\lambda)} V(Y_{r_i}) T_{r_i}}{M(\lambda)} \rightarrow B.$$

It follows that^(*)

$$\frac{V(Y_{M(\lambda)}) T_{M(\lambda)}}{M(\lambda)} \rightarrow 0 \quad \text{and} \quad \frac{V(Y_{M(\lambda)}) T_{M(\lambda)}}{h(\lambda)} \rightarrow 0$$

in probability. Since $(|\theta| \leq 1)$

$$(4.26) \quad \int_0^\lambda V(Y(t)) dt = \sum_{i=0}^{M(\lambda)-1} V(Y_{r_i}) T_{r_i} + \theta V(Y_{M(\lambda)}) T_{M(\lambda)},$$

(4.22) follows again from (4.20) and the corollary to Lemma 1. Let us finally prove (4.25). The proof of (4.24) will then be clear. First we show that conditions on $M(\lambda)$ can be replaced by conditions on $K(\lambda)$. From (4.18) and the definition of $\beta(\lambda)$ we have

$$(4.27) \quad \sum_{i=0}^{M(\lambda)} W_i + \beta(\lambda) \leq \lambda$$

and thus for any $\epsilon > 0$

$$M(\lambda) \leq K(\lambda - \beta(\lambda) + \epsilon).$$

On the other hand, again from (4.18)

$$(4.28) \quad \sum_{i=0}^{M(\lambda)+1} W_i \geq \lambda - \beta(\lambda)$$

which combined with the above gives

$$(4.29) \quad K(\lambda - \beta(\lambda)) \leq M(\lambda) \leq K(\lambda - \beta(\lambda) + \epsilon).$$

Since $\beta(\lambda)/\lambda \rightarrow a$ in probability we have for any $\epsilon > 0$

$$(4.30) \quad \lim_{\lambda \rightarrow \infty} P\{K(\lambda(1-a-\epsilon)) \leq M(\lambda) \leq K(\lambda(1-a+\epsilon))\} = 1$$

and this combined with (4.20) shows

^(*) Actually one can prove from renewal theory, that $\lim_{\lambda \rightarrow \infty} P\{|\beta(\lambda) - \sum_{i=0}^{M(\lambda)} V(Y_{r_i}) T_{r_i}| \geq A\} = 0$, uniformly in λ if the integrals in (4.13) converge.

$$(4.31) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{K(\lambda(1-a))}{h(\lambda)} \leq y \right\} = G(y).$$

Using Lemma 3, one has, if the integrals in (4.13) and (4.15) exist,

$$(4.32) \quad \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k V(Y_{r_i}) T_{r_i} - kB}{\rho k^{1/2}} \leq x \right\} = \phi(x),$$

and thus, by Lemma 2, (4.30), (4.31),

$$(4.33) \quad \begin{aligned} & \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k V(Y_{r_i}) T_{r_i} - kB}{\rho k^{1/2}} \leq x \text{ \& } \frac{M(\lambda)}{h(\lambda)} \leq y \right\} \\ &= \lim_{k \rightarrow \infty; \lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k V(Y_{r_i}) T_{r_i} - kB}{\rho k^{1/2}} \leq x \text{ \& } \frac{K(\lambda(1-a))}{h(\lambda)} \leq y \right\} = \phi(x) \cdot G(y). \end{aligned}$$

Lemma 1 shows then

$$(4.34) \quad \begin{aligned} & \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)-1} V(Y_{r_i}) T_{r_i} - M(\lambda)B}{\rho[h(\lambda)]^{1/2}} \leq x \right\} \\ &= \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{M(\lambda)} V(Y_{r_i}) T_{r_i} - M(\lambda)B}{\rho[h(\lambda)]^{1/2}} \leq x \right\} = \int_0^\infty \phi(xy^{-1/2}) dG(y). \end{aligned}$$

Notice that we do not know a priori the equality of the first and second member of (4.34) but obtain from Lemma 1 that both equal the third member. If we now take

$$V(j) = U(j) = \text{characteristic function of } J \text{ (cf. (4.3))}$$

then we obtain from the equality of the first and second member in (4.34) and the fact $T_i \geq 0$, that⁽⁸⁾

$$\frac{U(Y_{M(\lambda)}) T_{M(\lambda)}}{h(\lambda)} \rightarrow 0 \quad \text{and} \quad \frac{V(Y_{M(\lambda)}) T_{M(\lambda)}}{h(\lambda)} \rightarrow 0 \text{ in probability.}$$

This, together with (4.26) and (4.34) proves (4.25).

Theorem 2 extends the results of [2] to V 's which are not necessarily non-negative. The main interest of (4.24) and (4.25) is undoubtedly when C , respectively B equals zero. If we apply (4.25) to the function $W(j)$ which vanishes outside J but equals $V(j) - B/B'$ on J , we obtain

$$\lim_{\lambda \rightarrow \infty} P \left\{ \frac{\int_0^\lambda V(Y(t))dt - \beta(\lambda) \frac{B}{B'}}{\rho'[h(\lambda)]^{1/2}} \leq x \right\} = \int_0^\infty \phi(xy^{-1/2})dG(y)$$

where ρ'^2 is the expression in (4.16) computed for W .

By the techniques of this theorem one can of course also get limits like

$$\int_{0-}^\infty F(xy^{-\alpha})dG(y)$$

in (4.25) if the integrals in (4.15) and or (4.13) diverge and if instead of (4.32)

$$\lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k V(Y_{r_i})T_{r_i} - kB}{f(k)} \leq x \right\} = F(x).$$

An interesting special case of Theorem 2 arises when $Y(t)$ is actually a Markov chain with no instantaneous states. In this case, the distributions $H_j(x)$ are exponential distributions and the integrals in (4.13) and (4.15) converge automatically. Moreover, under the conditions of Theorem 1 in [2]

$$(4.35) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\beta(\lambda)}{\lambda^\alpha L(\lambda)} \leq x \right\} = G_\alpha(x)$$

for some $0 \leq \alpha \leq 1$ and slowly varying L , where

$$(4.36) \quad G_\alpha(x) = \frac{1}{\pi\alpha} \int_0^x \sum_{j=1}^\infty \frac{(-1)^{j-1}}{j!} \sin \pi\alpha j \Gamma(\alpha j + 1) t^{j-1} dt$$

is a Mittag-Leffler distribution. When $\alpha < 1$, (4.35) implies (4.23) with $\alpha = 0$, and together with (4.19)

$$\lim_{\lambda \rightarrow \infty} P \left\{ \frac{M(\lambda)}{\lambda^\alpha L(\lambda)} \leq x \right\} = G_\alpha(xB').$$

This makes Theorem 2 directly applicable. E.g., one obtains

$$(4.37) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\int_0^\lambda V(Y(t))dt - M(\lambda)B}{\rho[\lambda^\alpha L(\lambda)]^{1/2}} \leq x \right\} = \int_{0-}^\infty \phi(x(B')^{1/2}y^{-1/2})dG_\alpha(y).$$

Even if $\alpha = 1$, $L = \text{constant}$ (4.35) is useful to find the distribution of $K(\lambda)$. In this case

$$\frac{\beta(\lambda)}{\lambda L} \rightarrow 1 \quad \text{and} \quad \frac{M(\lambda)}{\lambda L} \rightarrow \frac{1}{B'} \text{ in probability.}$$

This implies, by (4.30),

$$\frac{K(\lambda)}{\lambda} \rightarrow \frac{L}{B'(1-L)} \text{ in probability.}$$

Another, very interesting special case is the case where $H_j(t)$ assigns probability one to $t=1$ for all j . This means that

$$(4.38) \quad 1 = T_0 = T_1 = T_2 = \cdots \text{ with probability one,}$$

i.e., the $Y(t)$ process is equivalent to a discrete time Markov chain. We formulate the consequences for this case in the next theorem. We shall assume the condition of [2]

$$(4.39) \quad \lim_{z \uparrow 1} \frac{(1-z)^\alpha}{L(1/(1-z))} \sum_{n=0}^{\infty} P_{00}^{(n)} z^n = p_0$$

for some $0 \leq \alpha \leq 1$ and slowly varying function $L(x)$ (p_0 is the same as in (4.1). As long as the limit in the left-hand side of (4.39) exists and differs from zero we can always make it p_0 by multiplying L by an appropriate factor). It was shown by Lamperti [12, Lemma 6.1] that (4.39) implies

$$(4.40) \quad \lim_{z \uparrow 1} \frac{(1-z)^\alpha}{L(1/(1-z))} \sum_{n=0}^{\infty} P_{i,j}^{(n)} z^n = p_j.$$

By the results of [2] we have therefore

$$(4.41) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{M(\lambda)}{\lambda^\alpha L(\lambda) \sum_{u=0}^r p_{j_u}} \leq y \right\} = G_\alpha(y)$$

since with (4.38) $M(\lambda)+1$ represent the number of visits to J in the first λ steps of the discrete time Markov chain.

THEOREM 3. *If $V(j)$ vanishes outside $J = \{j_0, \dots, j_r\}$ and if (4.39) holds, then with C and σ^2 given by (4.11) and (4.12)*

$$(4.42) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{\lambda} V(Y_i)}{\lambda^\alpha L(\lambda) \sum_{u=0}^r p_{j_u}} \leq x \right\} = \begin{cases} G_\alpha(xC^{-1}) & \text{if } C > 0, \\ 1 - G_\alpha(xC^{-1}) & \text{if } C < 0, \end{cases}$$

and

$$(4.43) \quad \lim_{\lambda \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^{\lambda} V(Y_i) - M(\lambda)C}{\sigma \left(\lambda^\alpha L(\lambda) \sum_{u=0}^r p_{j_u} \right)^{1/2}} \leq x \right\} = \int_{0-}^{\infty} \phi(xy^{-1/2}) dG_\alpha(y)$$

(with the same convention as in (3.35) if $\sigma = 0$).

For the proof we only need observe that

$$\sum_{i=0}^{\lambda-1} V(Y_i) = \sum_{i=0}^{M(\lambda)} V(Y_{r_i})$$

and that (4.41) implies (4.23) with $a = 0$ if $\alpha < 1$, since $\beta(\lambda) = M(\lambda)$ under the condition (4.38). If $\alpha = 1$ one can directly apply Lemma 1 since $G_1(y)$ assigns probability one to the value one. For the case where Y_0, Y_1, \dots is the simple random walk on the integers (i.e., $P\{Y_{i+1} - Y_i = +1\} = P\{Y_{i+1} - Y_i = -1\} = 1/2$) and $C = 0$ (4.43) was already proved by Dobrusin [6].

REMARK. If $V(j)$ does not vanish outside the finite set J but only outside an infinite set J' it is sometimes still possible to prove (4.42) or (4.43). $M(\lambda)$ then of course represents the number of visits to J' . In fact, if

$$(4.44) \quad \sum_{j \in J'} p_j < \infty$$

then the imbedded Markov chain Z' with respect to J' is positive recurrent (by use of §6.2.3 of [10]). Thus, if in addition to (4.44) also

$$(4.45) \quad \sum_{j \in J'} p_j |V(j)| < \infty$$

then, still

$$k^{-1} \sum_{i=0}^k V(Y_{r_i}) \rightarrow \left(\sum_{j \in J'} p_j \right)^{-1} \sum_{j \in J'} p_j V(j) = C'$$

with probability one (Theorem 6 in [1]). This easily gives (4.42). Also (4.43) (with the obvious change of the constants) can be shown to remain valid, if in addition to (4.44) and (4.45) one has

$$(4.46) \quad \sum_{j \in J'} \sum_{k \in J'} (m_{ji} + m_{ik} - m_{jk}) p_j p_k < \infty$$

and

$$(4.47) \quad \sum_{j \in J'} p_j V^2(j) + \sum_{j \in J'} \sum_{k \in J'} (m_{ji} + m_{ik} - m_{jk}) p_j p_k |V(j)| |V(k)| < \infty$$

where i is any (fixed) state in J' and m_{ij} as in (3.33) with Z' instead of Z .

The proof is based on an approximation argument. Instead of on J' , one takes V different from zero on larger and larger *finite* subsets of J' and applies (4.43). The error committed by this approximation is estimated by means of Theorem 8 in [1].

As we pointed out before, for a semi-markovian process as defined by Smith [15] the distribution of T_i may depend on Y_i and Y_{i+1} . In order to apply our results one first has to introduce the auxiliary Markov chain $\{Y'_i = (Y_i, Y_{i+1})\}$, in which case the distribution of T_i depends on Y'_i only.

However, a finite set of states J of the Y' chain is then a set of pairs (i, j) , $i \in J_1, j \in J_2$ where both J_1 and J_2 are finite. In such a case $\beta(\lambda)$ would represent only the time spent up till λ between transitions from J_1 to J_2 . If the whole state space is finite, one can of course take for J_2 the collection of all states. In all other cases, if one wants the time spent in J_1 up till λ , one needs some condition which insures that the time spent between transitions from J_1 to the complement of J_2 is small with respect to $\beta(\lambda)$, for large J_2 . One can then use an approximation argument by letting J_2 increase to the whole state space.

We shall also prove a theorem of a type different from the preceding ones. Define

$$s_0 < s_1 < s_2 \cdots$$

as the successive indices k for which $Y_k = j_0$ for some fixed state j_0 , and put

$$M(A, \lambda) = \text{number of indices } k \geq 1 \text{ for which}$$

$$s_k - s_{k-1} \geq A \quad \text{and} \quad s_k \leq \lambda.$$

$M(A, \lambda)$ is the number of recurrence times of the state j_0 which are at least A , up till λ .

THEOREM 4. *If (4.39) is satisfied, then*

$$\lim_{A \rightarrow \infty} \lim_{\lambda \rightarrow \infty} P \left\{ \frac{A^\alpha L(A) \Gamma(1 - \alpha)}{\lambda^\alpha L(\lambda)} M(A, \lambda) \leq x \right\} = G_\alpha(x).$$

Proof. Define independent, identically distributed random variables X_i by

$$X_i = \begin{cases} 1 & \text{if } s_i - s_{i-1} \geq A, \\ 0 & \text{if } s_i - s_{i-1} < A, \end{cases}$$

and let $N(\lambda)$ be the largest k with $s_k \leq \lambda$. Then

$$M(A, \lambda) = \sum_{i=1}^{N(\lambda)} X_i.$$

On the other hand, if

$$\gamma_m = \sum_{n=m}^{\infty} P \{ Y_n = j_0, Y_k \neq j_0 \text{ for } 1 \leq k \leq n-1 \mid Y_0 = j_0 \},$$

then $EX_i = \gamma_A$, while, as $z \uparrow 1$ (cf. [8, Theorem 1, p. 243])

$$\sum_{m=1}^{\infty} \gamma_m z^m = z(1-z)^{-1} \left(\sum_{n=0}^{\infty} P_{j_0, j_0}^{(n)} z^n \right)^{-1} \sim \frac{(1-z)^{\alpha-1}}{p_{j_0} L(1/(1-z))}$$

by (4.40). By Karamata's Tauberian theorem [7, p. 511]

$$\sum_{m=1}^A \gamma_m \sim A^{1-\alpha} (p_{j_0} L(A) \Gamma(2-\alpha))^{-1}$$

and, because $\gamma_m \geq \gamma_{m+1}$ one can deduce (compare Hilfssatz 3, p. 517 in [7])

$$\lim_{A \rightarrow \infty} A^\alpha L(A) \gamma_A = (p_{j_0} \Gamma(1-\alpha))^{-1}.$$

This implies of course for every $\epsilon > 0$ and sufficiently large A (say $A \geq A(\epsilon)$),

$$\frac{1-\epsilon}{p_{j_0}} \leq \lim_{k \rightarrow \infty} A^\alpha L(A) \Gamma(1-\alpha) \frac{\sum_{i=1}^k X_i}{k} \leq \frac{1+\epsilon}{p_{j_0}}$$

with probability one. The theorem follows now immediately from (4.41).

Let us end with a simple

EXAMPLE. Assume there are 4 states for the Y_i chain, say 0, 1, 2, 3. This can also serve as a model for a semi-Markov chain with 2 states where the distribution of T_i depends on Y_i and Y_{i+1} . For, if Y_i has 2 states, then there are 4 states for the process $Y'_i = (Y_i, Y_{i+1})$ of pairs. The transition probability matrix P should belong to an irreducible Markov chain such that the stationary probabilities q_0, q_1, q_2, q_3 are all positive. Let (3.1) and (3.2) hold with distribution functions satisfying

$$(4.48) \quad \begin{aligned} \lim_{t \rightarrow \infty} t^{\alpha_j} (1 - H_j(t)) &= A_j \neq 0, \\ 0 < \alpha_j < 2, \quad j &= 0, 1, 2, 3. \end{aligned}$$

If $U_1^{(j)}, U_2^{(j)}, \dots$ is a sequence of independent random variables, each with distribution function $H_j(x)$, then by (4.48) and [9, Theorem 5, p. 181, Theorem 2 and footnote, p. 175], one has

$$\begin{aligned} \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k U_i^{(j)}}{k^{1/\alpha_j}} \leq x \right\} &= F_{\alpha_j}(x) \quad \text{if } \alpha_j < 1, \\ \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k U_i^{(j)} - B_j(k) \cdot k}{k} \leq x \right\} &= F_1(x) \quad \text{if } \alpha_j = 1, \\ \lim_{k \rightarrow \infty} P \left\{ \frac{\sum_{i=0}^k U_i^{(j)} - C_j k}{k^{1/\alpha_j}} \leq x \right\} &= F_{\alpha_j}(x) \quad \text{if } \alpha_j > 1 \end{aligned}$$

where $F_{\alpha_j}(x)$ is some stable law of order α_j ,

$$\lim_{k \rightarrow \infty} (\log k)^{-1} B_j(k) \text{ exists}$$

and

$$C_j = \int_0^\infty x dH_j(x) \quad \text{if } \alpha_j > 1.$$

We consider now

$$\beta(\lambda) = \text{occupation time of } J = \{0, 1\} \text{ up till time } \lambda.$$

Define r_i , W_i , $K(\lambda)$ as in (3.8) and (3.13)–(3.15). From Lemma 3 we obtain the limiting distributions of $\sum_{i=0}^k T_{r_i}$, $\sum_{i=0}^k W_i$ and hence $K(\lambda)$. Theorem 1 then gives the limiting distribution of $\beta(\lambda)$. Take, for example

$$(4.49) \quad \frac{1}{\alpha_2} = \frac{1}{\alpha_3} > \frac{1}{\alpha_1} > \max\left(1, \frac{1}{\alpha_0}\right).$$

Then, by Lemma 3,

$$\lim_{k \rightarrow \infty} P\left\{\frac{\sum_{i=0}^k T_{r_i}}{k^{1/\alpha_1}} \leq x\right\} = F_{\alpha_1}\left(x\left(\frac{q_0 + q_1}{q_1}\right)^{1/\alpha_1}\right).$$

Also, if $U(j)$ is the characteristic function of $J = \{0, 1\}$ (cf. (4.3)) then by Lemma 3

$$(4.50) \quad \lim_{k \rightarrow \infty} P\left\{\frac{\sum_{i=0}^k T_i(1 - U(Y_i))}{k^{1/\alpha_2}} \leq x\right\} = F_{\alpha_2}(xq_2^{-1/\alpha_2}) * F_{\alpha_3}(xq_3^{-1/\alpha_3}).$$

But

$$\sum_{i=0}^{m(k)-1} W_i \leq \sum_{i=0}^k T_i(1 - U(Y_i)) \leq \sum_{i=0}^{m(k)} W_i$$

where $m(k)$ is the number of indices i , $0 \leq i \leq k$, for which $Y_i \in J$. Therefore

$$\frac{m(k)}{k} \rightarrow q_0 + q_1$$

with probability one and with (4.50)

$$(4.51) \quad \lim_{k \rightarrow \infty} P\left\{\frac{\sum_{i=0}^k W_i}{k^{1/\alpha_2}} \leq x\right\} = F_{\alpha_2}\left(x\left(\frac{q_0 + q_1}{q_2}\right)^{1/\alpha_2}\right) * F_{\alpha_3}\left(x\left(\frac{q_0 + q_1}{q_3}\right)^{1/\alpha_3}\right).$$

Consequently

$$\begin{aligned}\lim_{\lambda \rightarrow \infty} P\left\{\frac{K(\lambda)}{\lambda^{\alpha_2}} \leq y\right\} &= \lim_{\lambda \rightarrow \infty} P\left\{\frac{\sum_{i=0}^{y\lambda^{\alpha_2}} W_i}{y^{1/\alpha_2}\lambda} \geq y^{-1/\alpha_2}\right\} \\ &= 1 - F_{\alpha_2}\left(y^{-1/\alpha_2}\left(\frac{q_0 + q_1}{q_2}\right)^{1/\alpha_2}\right) * F_{\alpha_3}\left(y^{-1/\alpha_2}\left(\frac{q_0 + q_1}{q_3}\right)^{1/\alpha_2}\right) = G(y), \text{ say.}\end{aligned}$$

Putting $f(k) = k^{1/\alpha_1}$, $h(\lambda) = \lambda^{\alpha_2}$ one has

$$D = \lim_{\lambda \rightarrow \infty} \frac{f(h(\lambda))}{\lambda} = \lim_{\lambda \rightarrow \infty} \lambda^{\alpha_2/\alpha_1 - 1} = 0.$$

Thus, by Theorem 1,

$$\lim_{\lambda \rightarrow \infty} P\left\{\frac{\beta(\lambda)}{\lambda^{\alpha_2/\alpha_1}} \leq x\right\} = \int_0^\infty F_{\alpha_1}\left(x\left(\frac{q_0 + q_1}{q_1 y}\right)^{1/\alpha_1}\right) dG(y).$$

It is possible to compose a table, similar to the one in [16] for the many other relations between α_0 , α_1 , α_2 , α_3 which can come instead of (4.49).

REFERENCES

1. K. L. Chung, *Contributions to the theory of Markov chains*. II, Trans. Amer. Math. Soc. **76** (1954), 397-419.
2. D. A. Darling and M. Kac, *On occupation times for Markoff processes*, Trans. Amer. Math. Soc. **84** (1957), 444-458.
3. Cyrus Derman, *Some contributions to the theory of denumerable Markov chains*, Trans. Amer. Math. Soc. **79** (1955), 541-555.
4. W. Doeblin, *Sur deux problèmes de M. Kolmogoroff concernant les chaînes dénombrables*, Bull. Soc. Math. France **66** (1938), 210-220.
5. R. L. Dobrusin, *Lemma on the limit of compound random functions*, Uspehi Mat. Nauk **10** (64) (1955), 157-159, no. 2 (cf. Math. Rev. **17** (1956), 48).
6. ———, *Two limit theorems for the simplest random walk on a line*, Uspehi Mat. Nauk **10** (65) (1955), no. 3, 139-146 (cf. Math. Rev. **17** (1956), 166).
7. G. Doetsch, *Handbuch der Laplace-Transformation*. I, Birkhäuser, Basel, 1950.
8. W. Feller, *An introduction to probability theory and its applications*, Vol. 1, Wiley, New York, 1950.
9. B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley, Cambridge, Mass., 1954.
10. John G. Kemeny and J. Laurie Snell, *Finite Markov chains*, Van Nostrand, Princeton, N. J., 1960.
11. David G. Kendall, *A note on Doeblin's central limit theorem*, Proc. Amer. Math. Soc. **8** (1957), 1037-1039.
12. John Lamperti, *Some limit theorems for stochastic processes*, J. Math. Mech. **7** (1958), 433-448.
13. ———, *An occupation time theorem for a class of stochastic processes*, Trans. Amer. Math. Soc. **88** (1958), 380-387.
14. Ronald Pyke, *Limit theorems for Markov renewal processes*, Report CU 24-61 Nonr-

266(59) MS, Columbia University, Feb., 1961. Also: *Markov renewal processes with finitely many states*, Ann. Math. Statist. 32 (1961), 1243–1259 and several abstracts in Ann. Math. Statist. 31 (1960).

15. W. L. Smith, *Regenerative stochastic processes*, Proc. Roy. Soc. London 232 (1955), 6–31.

16. L. Takács, *On a sojourn time problem in the theory of stochastic processes*, Trans. Amer. Math. Soc. 93 (1959), 531–540.

Many of the quoted theorems on Markov chains can also be found in the recent book *Markov chains with stationary transition probabilities*, by Kai Lai Chung, to which the author, however, has had no access while writing this paper.

THE HEBREW UNIVERSITY,
JERUSALEM, ISRAEL