Z. Wahrscheinlichkeitstheorie verw. Geb. 27, 195–214 (1973)

On Denumerable Chains of Infinite Order

M. Iosifescu and A. Spătaru

Introduction

In [5] Harris studied strictly stationary doubly infinite sequences $(\xi_t)_{t\in Z}$, where ξ_t can take $D\geqq 2$ distinct values $0,1,\ldots,D-1$. The transitions of such a process are described by the function

$$P(\ldots,i_2,i_1;i)=\mathsf{P}(\xi_t=i\,|\,\xi_{t-1}=i_1,\xi_{t-2}=i_2,\ldots),$$

$0\leqq i,i_1,i_2,\ldots\leqq D-1$, $t\in Z$. Because of the fact that the future behaviour of ξ_t depends in general on its complete past history, these sequences are referred to as homogeneous chains of infinite order with a finite number of states. Harris' aim was to relate stochastic properties of $(\xi_t)_{t\in Z}$ to functional properties of its transition function. His technique was to map the half-infinite sequence $\ldots,\xi_{t-2},\xi_{t-1}$ onto the unit interval by means of the correspondence $\eta_t=\sum_{r\in N^*}\xi_{t-r}/D^r$, $t\in Z$. In other words, η_t is the number whose D-adic expansion is $.\xi_{t-1}\,\xi_{t-2}\ldots$. The η_t then form a Markov chain whose transition probabilities are given by

$$\mathsf{P}\left(\eta_{t+1}=\frac{i+\eta_t}{D}\,\middle|\,\eta_t=\sum_{r\in N^*}i_{t-r}/D^r\right)=P(\ldots,i_{t-2},i_{t-1};i).$$

In what follows we consider chains of infinite order with a denumerable set of states taken to be the natural numbers. Our device is to associate with the half-infinite sequence $\ldots,\xi_{t-2},\xi_{t-1}$ the irrational number η_t whose continued fraction expansion[1] is $(\xi_{t-1},\xi_{t-2},\ldots)$. Thus, we are led to study in Section 2 certain Y-valued Markov chains with transition probabilities of the form

$$\mathsf{P}\left(\eta_{n+1}=\frac{1}{i+y}\,\middle|\,\eta_n=y\right)=p_i(y),\qquad i\in N^*,\ y\in Y,$$

where $(p_i(y))_{i\in N^*}$ is a given probability distribution on the natural numbers for any $y\in Y$. Next, by making use of the properties of such Y-valued Markov chains, we establish in Section 3 the existence of denumerable chains of infinite order under conditions different from those given in [7], p. 188. The results obtained can be viewed as properties of the continued fraction expansion as well.

Our treatment follows closely that of Harris. Nevertheless, it is necessary to note that his proofs are just sketched and even untrue in full generality under the conditions assumed. (See further Footnote 2.)

[1] Other f-expansions (see, e.g., [9]) could be employed as well.

Notations

$N^* = \{1, 2, 3, \ldots\}$,
$N = \{0, 1, 2, \ldots\}$,
$-N = \{\ldots, -2, -1, 0\}$,
$Z = (-N) \cup N^*$,
$Y =$ the set of irrationals in $[0, 1]$,
$R =$ the set of real numbers,
$a^+ = \max(a, 0)$, $a \in R$,
$[a] =$ integral part of $a \in R$,
$c_m = m$-th Fibonacci number, $m \in N$, defined recursively by
$c_0 = c_1 = 1$, $c_m = c_{m-1} + c_{m-2}$, $m \geqq 2$.

1. Preliminaries

1.1. For any probability distribution $\mathbf{p} = (p_i)_{i \in N^*}$ on N^* let $s_0(\mathbf{p}) = 0$, $s_i(\mathbf{p}) = \sum_{l=1}^{i} p_l$, $i \in N^*$. Let θ be a random variable uniformly distributed on $[0, 1]$. For any two probability distributions $\mathbf{p} = (p_i)_{i \in N^*}$ and $\mathbf{q} = (q_i)_{i \in N^*}$ on N^* define N^*-valued random variables σ and τ by

$$\sigma = i \quad \text{iff} \quad s_{i-1}(\mathbf{p}) < \theta \leqq s_i(\mathbf{p}), \qquad i \in N^*,$$

and

$$\tau = i \quad \text{iff} \quad s_{i-1}(\mathbf{q}) < \theta \leqq s_i(\mathbf{q}), \qquad i \in N^*,$$

and put

$$\begin{aligned} d(\mathbf{p}, \mathbf{q}) &= Pr(\sigma \neq \tau) \\ &= \sum_{i \in N^*} \big(Pr(\sigma = i, \tau > i) + Pr(\sigma > i, \tau = i)\big) \\ &= \sum_{i \in N^*} \big\{(s_i(\mathbf{p}) - \max(s_{i-1}(\mathbf{p}), s_i(\mathbf{q})))^+ + (s_i(\mathbf{q}) - \max(s_{i-1}(\mathbf{q}), s_i(\mathbf{p})))^+\big\}. \end{aligned}$$

Let us notice that for any $A \subset N^*$ we may write

$$Pr(\sigma \in A, \tau \notin A) \leqq d(\mathbf{p}, \mathbf{q}), \qquad Pr(\sigma \notin A, \tau \in A) \leqq d(\mathbf{p}, \mathbf{q}),$$

whence

$$d(\mathbf{p}, \mathbf{q}) \geqq |Pr(\sigma \in A, \tau \notin A) - Pr(\sigma \notin A, \tau \in A)|.$$

On the other hand

$$Pr(\sigma \in A, \tau \notin A) - Pr(\sigma \notin A, \tau \in A) = Pr(\sigma \in A) - Pr(\tau \in A) = \sum_{i \in A} p_i - \sum_{i \in A} q_i.$$

Therefore

$$d(\mathbf{p}, \mathbf{q}) \geqq \sup_{A \subset N^*} \Big| \sum_{i \in A} p_i - \sum_{i \in A} q_i \Big| = 2^{-1} \sum_{i \in N^*} |p_i - q_i|. \tag{1}$$

(The last equality is well known. See, e.g., [1], p. 224.)

It should be noted that, in general, the sign "$\geqq$" in (1) cannot be replaced by "$=$". A special case for which "$=$" holds instead of "$\geqq$" is that of distributions $\mathbf{p}$ and $\mathbf{q}$ such that $p_i = q_i = 0$ for $i > 2$.[2]

Remark. It is easily seen that $d(.,.)$ is a metric on the set of all probability distributions on N^*.

---

[2] It seems that Harris, who considered in some detail this case only, overlooked the general situation. Consequently, the definition of his ε_m, $m \in N$, and the statement of his Condition B (see [5], p. 712) have to be altered for $D \geqq 3$.

Next, consider a family $(\mathbf{p}(\gamma))_{\gamma\in\Gamma}$ of probability distributions on N^*, where Γ is an arbitrary set. Put

$$\varepsilon_0=\sup_{\gamma,\gamma'\in\Gamma} d(\mathbf{p}(\gamma),\mathbf{p}(\gamma')).$$

Proposition 1. *Assume that* $\varepsilon_0<\delta<1$. *Then either*

(a) $p_i(\gamma)\leqq\delta$ *for any* $i\in N^*$, $\gamma\in\Gamma$, *or*

(b_i) *there exists* $i\in N^*$ *such that* $p_i(\gamma)>\delta-\varepsilon_0$ *for any* $\gamma\in\Gamma$ (*consequently,* $p_j(\gamma)\leqq 1-\delta+\varepsilon_0$ *for any* $j\neq i$, $\gamma\in\Gamma$).

Proof. If (a) does not hold, then there exist $i\in N^*$ and $\gamma\in\Gamma$ such that $p_i(\gamma)>\delta$. Let us prove that then (b_i) holds. Indeed, on account of (1) with $A=\{i\}$, the existence of a $\gamma'\in\Gamma$ such that $p_i(\gamma')\leqq\delta-\varepsilon_0$ would lead to

$$d(\mathbf{p}(\gamma),\mathbf{p}(\gamma'))\geqq|p_i(\gamma)-p_i(\gamma')|>\delta-(\delta-\varepsilon_0)=\varepsilon_0,$$

thus contradicting the definition of ε_0, q.e.d.

Proposition 2. *Assume that* $\varepsilon_0<1$. *Then there exists at most one* $i\in N^*$ *such that* $p_i(\gamma_i)=1$ *for some* $\gamma_i\in\Gamma$.

Proof. Suppose on the contrary there are $i\neq j\in N^*$ and $\gamma_i,\gamma_j\in\Gamma$ such that $p_i(\gamma_i)=p_j(\gamma_j)=1$. Then $p_i(\gamma_j)=0$ and on account of (1) with $A=\{i\}$ we can write

$$d(\mathbf{p}(\gamma_i),\mathbf{p}(\gamma_j))\geqq|p_i(\gamma_i)-p_i(\gamma_j)|=1.$$

It follows that $\varepsilon_0=1$, thus contradicting the hypothesis made, q.e.d.

1.2. The following lemma (which we shall need in the proof of Lemma 4 below) is only slightly different from Lemma 1 in [5], p. 713 given there without proof.

Lemma 3. *Let* $(\zeta_i)_{i\in N^*}$ *be a sequence of* $N^*\cup\{\infty\}$*-valued random variables. Let* $\sigma_0=0$, $\sigma_n=\sum_{i=1}^{n}\zeta_i$, $n\in N^*$, $u_m=Pr\left(\bigcup_{j=1}^{m}(\sigma_j=m)\right)$, $m\in N^*$. *Assume that*

$$Pr(\zeta_n>k\,|\,\sigma_{n-1})\geqq r_k, \tag{2}$$

whatever $n,k\in N^*$, *where the* r_k *are nonnegative numbers such that*

$$\sum_{k\in N^*} r_k=\infty. \tag{3}$$

Then

$$\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} u_i=0. \tag{4}$$

If in addition the r_k *satisfy* $r_k\geqq a>0$, $k\in N^*$, *then* (4) *can be replaced by*

$$\sum_{i\in N^*} u_i<\infty. \tag{1'}$$

Proof. It is obvious that

$$u_m=\sum_{j=1}^{m} Pr(\sigma_j=m),\qquad m\in N^*. \tag{5}$$

Let us consider

$$v_m=1-u_m=Pr\left(\bigcap_{j=1}^{m}(\sigma_j\neq m)\right),\qquad m\in N^*.$$

It is easily seen that

$$v_m = \sum_{k=1}^{m-1} \sum_{j=1}^{k} Pr(\sigma_j = k, \zeta_{j+1} > m-k) + Pr(\zeta_1 > m)$$

for any $m \in N^*$, with the usual convention about empty summation. On account of (2) and (5) we may write

$$v_m \geqq \sum_{k=1}^{m-1} \sum_{j=1}^{k} r_{m-k} Pr(\sigma_j = k) + r_m = \sum_{k=0}^{m-1} u_k r_{m-k}$$

(with $u_0 = 1$). Since $u_m + v_m = 1$, $m \in N^*$, we obtain

$$\sum_{k=0}^{m} u_k r_{m-k} \leqq 1, \qquad m \in N^* \tag{6}$$

(with $r_0 = 1$). Now, (6) implies that

$$\Big(\sum_{k \in N} u_k x^k\Big)\Big(\sum_{k \in N} r_k x^k\Big) \leqq (1-x)^{-1}$$

for any $0 \leqq x < 1$, whence according to (3),

$$\lim_{x \uparrow 1} (1-x) \sum_{k \in N} u_k x^k = 0.$$

Then, (4) follows on account of Theorem 96 in [4], p. 155.

(A simple proof of the fact that the last equation implies (4) (for which we are indebted to Prof. Ciprian Foiaş) is as follows: For any $\varepsilon > 0$ there exists $0 \leqq x_\varepsilon < 1$ such that $(1-x) \sum_{k \in N} u_k x^k < \varepsilon$ for any x in the interval $[x_\varepsilon, 1)$. It follows that for any natural number $n > (1-x_\varepsilon)^{-1}$ we may successively write

$$\varepsilon > n^{-1} \sum_{k \in N} u_k \left(1 - \frac{1}{n}\right)^k \geqq n^{-1} \sum_{k=0}^{[n/2]} u_k \left(1 - \frac{k}{n}\right) \geqq (2n)^{-1} \sum_{k=0}^{[n/2]} u_k,$$

so that (4) holds.)

Finally, if $r_k \geqq a > 0$, $k \in N^*$, then (6) implies that

$$\sum_{k=0}^{m} u_k \leqq a^{-1}, \qquad m \in N^*$$

and (4′) follows, q.e.d.

1.3. Let Y be the set of irrationals in $[0, 1]$. For each $y \in Y$ let $(i_1, i_2, \dots)$ denote the infinite continued fraction expansion

$$\frac{1|}{|i_1} + \frac{1|}{|i_2} + \dots + \frac{1|}{|i_n} + \dots$$

of y. We shall use the notation $(y \equiv y')_m$, $y, y' \in Y$, $m \in N$, to mean that the first m digits in the continued fraction expansion of y are the same as the first m in the expansion of y'. (Of course, $(y \equiv y')_0$ means no restriction is imposed on y and y'.) It is easily proved that if $(y \equiv y')_m$, $y, y' \in Y$, $m \in N$, then $|y - y'| < (c_m c_{m+1})^{-1}$, where

$$c_m = 2^{-m-1}((1+\sqrt{5})^{m+1} - (1-\sqrt{5})^{m+1})/\sqrt{5}, \qquad m \in N,$$

i.e. the Fibonacci numbers defined by $c_0 = c_1 = 1$, $c_m = c_{m-1} + c_{m-2}$, $m \geqq 2$.

2. A Class of Y-Valued Markov Chains

2.1. Assume we are given a sequence $\mathbf{p}(.)=(p_i(.))_{i\in N^*}$ of functions defined on Y such that $\mathbf{p}(y)$ is a probability distribution on N^* for any $y\in Y$. Put $s_0(y)=0$, $s_i(y)=s_i(\mathbf{p}(y))=\sum_{l=1}^{i} p_l(y)$, $i\in N^*$, $y\in Y$, and define

$$\varepsilon_m=\sup_{(y\equiv y')_m} d(\mathbf{p}(y),\mathbf{p}(y')),\qquad m\in N. \tag{7}$$

Clearly, $\varepsilon_m\leqq 1$, $m\in N$, and the sequence $(\varepsilon_m)_{m\in N}$ is nonincreasing.

In what follows we shall use

Condition H: $\sum_{k\in N}\prod_{m=0}^{k}(1-\varepsilon_m)=\infty$.

We notice that Condition H implies that both $\varepsilon_0<1$ and $\lim_{m\to\infty}\varepsilon_m=0$. On account of (1) this last equation implies that the functions p_i, $i\in N^*$, are continuous on Y.

Next, by virtue of Propositions 1 and 2, the inequality $\varepsilon_0<1$ implies the following alternatives. Either

(a) $\sup_{i\in N^*,\, y\in Y} p_i(y)<1$, or

(b_i) $\inf_{y\in Y} p_i(y)>0$ (consequently, $\sup_{j\neq i,\, y\in Y} p_j(y)<1$).

Also, either

(a_i) $p_i(y_i)=1$, where $y_i\in Y$ satisfies $y_i=(i+y_i)^{-1}$, or

(b) $p_i(y_i)<1$ for any $i\in N^*$.

Let $(t_n)_{n\in N}$ be a sequence of independent random variables uniformly distributed on $[0,1]$. Define Y-valued random sequences $(\eta_n)_{n\in N}$ and $(\eta'_n)_{n\in N}$ as follows. Set $\eta_0=y_0$, $\eta'_0=y'_0$, $y_0, y'_0\in Y$, and, assuming that η_n and η'_n are determined, then

$$\eta_{n+1}=\frac{1}{i+\eta_n}\quad\text{iff}\quad s_{i-1}(\eta_n)<t_n\leqq s_i(\eta_n),\qquad i\in N^*,$$

while

$$\eta'_{n+1}=\frac{1}{i+\eta'_n}\quad\text{iff}\quad s_{i-1}(\eta'_n)<t_n\leqq s_i(\eta'_n),\qquad i\in N^*.$$

Consider also N^*-valued random sequences $(\alpha_n)_{n\in N}$ and $(\alpha'_n)_{n\in N}$ defined by

$$\alpha_n=i\quad\text{iff}\quad s_{i-1}(\eta_n)<t_n\leqq s_i(\eta_n),\qquad i\in N^*,$$

and

$$\alpha'_n=i\quad\text{iff}\quad s_{i-1}(\eta'_n)<t_n\leqq s_i(\eta'_n),\qquad i\in N^*.$$

Therefore

$$\eta_{n+1}=\frac{1}{\alpha_n+\eta_n},\qquad \eta'_{n+1}=\frac{1}{\alpha'_n+\eta'_n},\qquad n\in N. \tag{8}$$

Clearly, $(\eta_n)_{n\in N}$ and $(\eta'_n)_{n\in N}$ are Y-valued homogeneous Markov chains with the same transition law given by [3]

$$\mathbf{P}\left(\eta_{n+1}(\eta'_{n+1})=\frac{1}{i+y}\,\middle|\,\eta_n(\eta'_n)=y\right)=p_i(y),\qquad i\in N^*,\ y\in Y.$$

[3] Throughout this section $\mathbf{P}$ is to be understood as depending on y_0 and y'_0. We avoid the notation $\mathbf{P}_{y_0,y'_0}$ in order to simplify the writing.

Notice that by the very construction of the sequences considered, for any event E belonging to the σ-algebra generated by the α_n and α'_n, $n \in N$, we have

$$\mathbf{P}(E \mid \eta_m, \eta'_m, 0 \leqq m \leqq n) = \mathbf{P}(E \mid \eta_n, \eta'_n), \qquad n \in N. \tag{9}$$

Further, for any $n \in N$

$$\mathbf{P}(\alpha_n \neq \alpha'_n \mid \eta_n, \eta'_n) = d(\mathbf{p}(\eta_n), \mathbf{p}(\eta'_n)) \leqq \varepsilon_0, \tag{10}$$

whence $\mathbf{P}(\alpha_n \neq \alpha'_n) \leqq \varepsilon_0$. Then, by making use of (10) it is not difficult to prove that

$$\mathbf{P}(\alpha_{n+j} = \alpha'_{n+j}, 0 \leqq j \leqq k \mid \eta_n, \eta'_n) \geqq \prod_{m=0}^{k} (1 - \varepsilon_m), \tag{11}$$

$$\mathbf{P}(\alpha_{n+j} \neq \alpha'_{n+j}, 0 \leqq j \leqq k) \leqq \varepsilon_0^{k+1}, \tag{12}$$

whatever $n, k \in N$, and that

$$\mathbf{P}(\alpha_{n+j} = \alpha'_{n+j}, \alpha_{n+l} \neq \alpha'_{n+l}, 0 \leqq j < k \leqq l \leqq m) \leqq \varepsilon_k \, \varepsilon_0^{m-k} \tag{13}$$

whatever $n \in N$, $k \in N^*$ and $m \geqq k$.

2.2. Now, we are prepared to prove

Lemma 4. *Assume that Condition H holds. Then for any $\varepsilon > 0$ one has*

$$\lim_{n \to \infty} n^{-1} \sum_{m=1}^{n} \mathbf{P}(|\eta_m - \eta'_m| < \varepsilon) = 1$$

uniformly with respect to $y_0, y'_0 \in Y$.

Proof. The first step of the proof, related to Doeblin's "two-particle" method, will be the construction of a sequence of random variables to which Lemma 3 applies.

Following Harris ([5], p. 714) let us say that an "engagement" occurs on the n-th step if $\alpha_{n-1} \neq \alpha'_{n-1}$ and $\alpha_n = \alpha'_n$. For any $s \leqq t \in N$, let $\rho_{s,t}$ denote the number of engagements occurring in the interval $[s, t]$.

Let τ_n, $n \in N^*$, be the moment of occurrence of the n-th engagement. Since even τ_1 may take the value ∞ some more explanation is needed. In fact τ_2 will be defined only on $\Omega_1 = (\tau_1 < \infty)$ and, in general, if τ_n is defined on Ω_{n-1}, $n \geqq 2$, then τ_{n+1} will be defined on $\Omega_{n-1} \cap (\tau_n < \infty)$.

Consider the sequence $(\zeta_i)_{i \in N^*}$ defined by $\zeta_1 = \tau_1$ and for $n \geqq 2$

$$\zeta_n = \begin{cases} \tau_n - \tau_{n-1} & \text{on } \Omega_n \\ \infty & \text{on } \Omega_n^c. \end{cases}$$

Let us show that this sequence fulfils the assumptions of Lemma 3. Clearly, $\sigma_n = \sum_{i=1}^{n} \zeta_i$ equals τ_n on Ω_n and ∞ on Ω_n^c, $n \in N^*$. Then

$$\mathbf{P}(\zeta_n > k \mid \sigma_{n-1} = \infty) = 1, \qquad n \geqq 2,$$

and on account of (9) and (11)

$$\mathbf{P}(\zeta_n > k \mid \sigma_{n-1} = t < \infty) \geqq \mathbf{P}(\alpha_{t+j} = \alpha'_{t+j}, 0 \leqq j \leqq k \mid \sigma_{n-1} = t) \geqq \prod_{m=0}^{k} (1 - \varepsilon_m), \qquad n \in N^*.$$

Since we have assumed Condition H the verification is complete.

Now, let us notice that, in the present context, u_m of Lemma 3 is in fact the probability that an engagement occurs on the m-th step. Hence $u_1+\cdots+u_n$ is the expected number of engagements in the first n steps. Therefore, on account of Lemma 3, we may assert that

$$\lim_{n\to\infty}\frac{\mathbf{E}\,\rho_{0,n}}{n}=0, \tag{14}$$

uniformly with respect to $y_0, y_0'\in Y$.

Next, for any natural number $k<r$ we have

$$\begin{aligned}\mathbf{P}(\rho_{r-k,r-1}=0)=&\mathbf{P}(\alpha_j=\alpha_j', r-k-1\leqq j\leqq r-1)\\&+\sum_{m=1}^{k}\mathbf{P}(\alpha_j=\alpha_j', \alpha_l\neq\alpha_l', r-k-1\leqq j<r-m\leqq l\leqq r-1)\\&+\mathbf{P}(\alpha_j\neq\alpha_j', r-k-1\leqq j\leqq r-1).\end{aligned}$$

It follows on account of (12) and (13) that

$$\mathbf{P}(\rho_{r-k,r-1}=0)=1-\mathbf{P}(\rho_{r-k,r-1}\geqq 1)\leqq \mathbf{P}(\alpha_j=\alpha_j', r-k-1\leqq j\leqq r-1)+a_k,$$

where $a_k=\sum_{m=0}^{k}\varepsilon_{k-m}\varepsilon_0^m$. By making use of Markov's inequality (see e.g. [8], p. 158) we deduce that

$$\mathbf{P}(\alpha_j=\alpha_j', r-k-1\leqq j\leqq r-1)\geqq 1-\mathbf{E}\,\rho_{r-k,r-1}-a_k. \tag{15}$$

Now, let $\varepsilon>0$ and take $k\geqq\min\{m: (c_m c_{m+1})^{-1}\leqq\varepsilon\}$. Since the event $(\alpha_j=\alpha_j', r-k-1\leqq j\leqq r-1)$ implies $(\eta_r\equiv\eta_r')_{k+1}$ and since $\sum_{r=k}^{n}\rho_{r-k,r-1}\leqq k\,\rho_{0,n-1}$, on account of (15) we may successively write

$$\begin{aligned}n^{-1}\sum_{m=1}^{n}\mathbf{P}(|\eta_m-\eta_m'|<\varepsilon)\geqq& n^{-1}\sum_{r=k+1}^{n}\mathbf{P}(\alpha_j=\alpha_j', r-k-1\leqq j\leqq r-1)\\ &\geqq (n-k)\,n^{-1}-k\,n^{-1}\,\mathbf{E}\,\rho_{0,n-1}-(n-k)\,n^{-1}\,a_k,\end{aligned}$$

whence on letting $n\to\infty$ and making use of (14), we obtain

$$\liminf_{n\to\infty} n^{-1}\sum_{m=1}^{n}\mathbf{P}(|\eta_m-\eta_m'|<\varepsilon)\geqq 1-a_k$$

uniformly with respect to $y_0, y_0'\in Y$.

To complete the proof we need only show that $\lim_{k\to\infty}a_k=0$. As we have already noted, Condition H implies that both $\varepsilon_0<1$ and $\lim_{m\to\infty}\varepsilon_m=0$. Next, for any $0<k'<k$ we can write $a_k\leqq\varepsilon_{k'}/(1-\varepsilon_0)+k'\,\varepsilon_0^{k-k'+1}$. Therefore, the convergence of a_k to zero follows upon letting first $k\to\infty$ and then $k'\to\infty$, q.e.d.

Corollary. *Let* $F_n(y_0; x) = n^{-1} \sum_{m=1}^{n} \mathbf{P}(\eta_m \leqq x)$ $\left(\text{so that } n^{-1} \sum_{m=1}^{n} \mathbf{P}(\eta'_m \leqq x) = F_n(y'_0; x)\right)$, $x \in R$. *Then*

$$\lim_{n\to\infty} d_L(F_n(y_0; .), F_n(y'_0; .)) = 0,$$

uniformly with respect to $y_0, y'_0 \in Y$, *where* d_L *denotes Paul Lévy's distance*[4].

Proof. Let $\varepsilon > 0$ be arbitrarily given. By Lemma 4 we can find an $n_\varepsilon \in N^*$ such that $n^{-1} \sum_{m=1}^{n} \mathbf{P}(|\eta_m - \eta'_m| \geqq \varepsilon) < \varepsilon$ for any $n \geqq n_\varepsilon$, $y_0, y'_0 \in Y$.

Now, $(\eta_m \leqq x) \cap (\eta'_m > x + \varepsilon) \subset (|\eta_m - \eta'_m| \geqq \varepsilon)$, for any $m \in N^*$, $x \in R$, and on account of the elementary inequality $\mathbf{P}(A \cap B) \geqq \mathbf{P}(A) - \mathbf{P}(B^c)$, we deduce that

$$F_n(y_0; x) \leqq F_n(y'_0; x + \varepsilon) + \varepsilon$$

for any $n \geqq n_\varepsilon$, $x \in R$, $y_0, y'_0 \in Y$. By interchanging y_0 and y'_0 and replacing x by $x - \varepsilon$ we finally deduce that

$$F_n(y'_0; x - \varepsilon) - \varepsilon \leqq F_n(y_0; x) \leqq F_n(y'_0; x + \varepsilon) + \varepsilon$$

for any $n \geqq n_\varepsilon$, $x \in R$, $y_0, y'_0 \in Y$. But this amounts to

$$d_L(F_n(y_0; .), F_n(y'_0; .)) \leqq \varepsilon$$

for any $n \geqq n_\varepsilon$, $y_0, y'_0 \in Y$, thus completing the proof, q.e.d.

Remark. If one assumes that $\varepsilon_0 < 1$ and $\sum_{m \in N} \varepsilon_m < \infty$ (a stronger hypothesis than Condition H) the conclusion of Lemma 4 can be strenghtened in a certain respect. We have namely

Lemma 4'. *If* $\varepsilon_0 < 1$ *and* $\sum_{m \in N} \varepsilon_m < \infty$ *then*

$$\mathbf{P}(\lim_{n\to\infty} (\eta_n - \eta'_n) = 0) = 1.$$

Proof. It is sufficient to prove that

$$\mathbf{P}\Big(\bigcup_{n \in N} \bigcap_{m \geqq n} (\alpha_m = \alpha'_m)\Big) = 1.$$

On account of (12) we have

$$\mathbf{P}\Big(\bigcup_{n \in N} \bigcap_{m \geqq n} (\alpha_m = \alpha'_m)\Big) = \lim_{n\to\infty} \mathbf{P}\Big(\bigcap_{m \geqq n} (\alpha_m = \alpha'_m)\Big) = \lim_{n\to\infty} \mathbf{P}(\rho_{n+1} = 0),$$

where ρ_{n+1} is the number of engagements occurring after the n-th step. As already noticed in the proof of Lemma 4, in the present context u_m of Lemma 3 is the probability that an engagement occurs on the m-th step. Hence, $\sum_{m \geqq n+1} u_m$ equals $\mathbf{E}\rho_{n+1}$. Further, r_k of Lemma 3 can be taken as $\prod_{m=0}^{k} (1 - \varepsilon_m)$, and then $\sum_{m \in N} \varepsilon_m < \infty$

[4] As well known d_L defined by

$$d_L(F, G) = \inf\{\varepsilon > 0:\ F(x - \varepsilon) - \varepsilon \leqq G(x) \leqq F(x + \varepsilon) + \varepsilon, x \in R\},$$

where F and G are distribution functions, is a distance on the set of all distribution functions. Weak convergence of distribution functions and convergence in the metric d_L are equivalent. See, e.g., [11], p. 133.

implies the existence of an $a>0$ such that $r_k \geqq a$, $k \in N$. As by Markov's inequality $\mathbf{P}(\rho_{n+1}=0) \geqq 1-\mathbf{E}\rho_{n+1}$, we need only apply the second half of Lemma 3, q.e.d.

Clearly, Lemma 4′ implies that whatever $\varepsilon>0$

$$\lim_{n\to\infty} \mathbf{P}(|\eta_n-\eta'_n|<\varepsilon)=1$$

for *any* y_0, $y'_0 \in Y$, but uniformity of convergence with respect to y_0, $y'_0 \in Y$ seems not to follow.

Actually, on account of (15) for any $\varepsilon>0$ and any $n>k \geqq \min\{m\colon (c_m c_{m+1})^{-1} \leqq \varepsilon\}$ we have

$$\mathbf{P}(|\eta_n-\eta'_n|<\varepsilon) \geqq \mathbf{P}(\alpha_j=\alpha'_j,\, n-k-1 \leqq j \leqq n-1)$$

$$\geqq 1-\mathbf{E}\rho_{n-k,n-1}-a_k=1-\sum_{m=n-k}^{n-1} u_m-a_k,$$

so that uniformity of convergence with respect to y_0, $y'_0 \in Y$ of $\mathbf{P}(|\eta_n-\eta'_n|<\varepsilon)$ to 1 as $n\to\infty$ would be ensured by uniformity of convergence with respect to y_0, $y'_0 \in Y$ of u_m to zero as $m\to\infty$. Notice that from Lemma 3 we only know that $\sum_{k=0}^{m} u_k \leqq a^{-1}$, $m \in N^*$.

2.3. Let $\mathscr{F}$ denote the set of (right-continuous) distribution functions F such that $F(0-)=0$ and $F(1)=1$. It is well known that $(\mathscr{F}, d_L)$ is a compact metric space. (See [8], p. 180.)

Proposition 5. *Assume that Condition H holds. Then in case* $(\mathfrak{a}_i)$

$$\lim_{n\to\infty} d_L(F_n(y_0;.), F)=0 \tag{16}$$

uniformly with respect to $y_0 \in Y$, *where*

$$F(x)=\begin{cases}0 & \text{for } x<y_i\\ 1 & \text{for } x\geqq y_i.\end{cases}$$

Proof. We begin by noticing that on account of Corollary to Lemma 4, Eq. (16) for a fixed $y_0 \in Y$, implies its validity for any $y_0 \in Y$ (i.e. F is independent of y_0) and, moreover, convergence is uniform with respect to y_0. Then, in case $(\mathfrak{a}_i)$ we have $\mathbf{P}(\eta_n=y_i|\eta_0=y_i)=1$ for any $n \in N^*$, so that

$$F_n(y_i;x)=\begin{cases}0 & \text{for } x<y_i\\ 1 & \text{for } x\geqq y_i,\end{cases}$$

whence, on account of the remark above, (16) holds and F has the form stated, q.e.d.

To state the next result we need

Condition U. The series $\sum_{i\in N^*} p_i(y)$ is uniformly convergent with respect to $y \in Y$.

Notice that Condition U is automatically satisfied (by virtue of Dini theorem) in the case where $(p_i(y))_{i\in N^*}$ is a probability distribution for *any* $y \in [0, 1]$ and the p_i are continuous functions on $[0, 1]$. (Actually, continuity can be replaced by a weaker assumption, namely lower semi-continuity. See [1], p. 218.)

Remark. Condition H does not imply Condition U as it might seem. This is shown by the following example. Take $p_1(y)=(1+y)/2$ and for $i\geqq 2$

$$p_i(y)=\begin{cases} y/2 & \text{if } 0<y<1/i \\ (1-(i-1)y)/2 & \text{if } 1/i<y<1/(i-1) \\ 0 & \text{if } 1/(i-1)<y<1. \end{cases}$$

This leads to

$$s_i(y)=\left\{\begin{matrix} (1+iy)/2 & \text{if } 0<y<1/i \\ 1 & \text{if } 1/i<y<1 \end{matrix}\right\}, \quad i\in N^*,$$

so that Condition U is not satisfied. On the other hand $\varepsilon_0=\frac{1}{2}$, $\varepsilon_m=(4c_{m-1}c_m)^{-1}$, $m\in N^*$, that is Condition H holds. (Actually, the stronger condition $\varepsilon_0<1$ and $\sum_{m\in N}\varepsilon_m<\infty$ is satisfied.) The ε_m are deduced as follows. Set $\Phi(x,x')=d(\mathbf{p}(x),\mathbf{p}(x'))$ for $0\leqq x, x'\leqq 1$. As $\Phi(x,x')=\Phi(x',x)$ we need only study the case $x>x'$. We have

$$\Phi(x,x')=\begin{cases} 1-s_1(x') & \text{if } (x,x')\in D_1 \\ \sum_{j=1}^{k-1}\left(s_j(x)-s_j(x')\right)+1-s_k(x') & \text{if } (x,x')\in D_k,\ k\geqq 2, \end{cases}$$

where $D_1=\{(x,x')\colon x'\leqq x/2\}$, $D_k=\{(x,x')\colon (k-1)x/k<x'\leqq kx/(k+1)\}$, $k\geqq 2$. This shows that Φ is continuous, $\Phi(.,x')$ is increasing on $[x',1]$ and $\Phi(x,.)$ decreasing on $[0,x]$. Thus $\varepsilon_0=\Phi(1,0)=1-s_1(0)=\frac{1}{2}$. Next, in the region $\{(x,x')\colon x>x',\ 1/(i+1)\leqq x,x'\leqq 1/i\}$ the function Φ takes the value $i(i+1)(x-x')/4$ at (x,x'). A moment's reflection then shows that

$$\varepsilon_1=\sup_{i\in N^*}\Phi\left(\frac{1}{i},\frac{1}{i+1}\right)=\sup_{i\in N^*}\frac{i(i+1)}{4}\left(\frac{1}{i}-\frac{1}{i+1}\right)=\frac{1}{4},$$

and, for $m\geqq 2$,

$$\varepsilon_m=\sup_{i\in N^*}\frac{i(i+1)}{4}\left|(i+c_{m-2}/c_{m-1})^{-1}-(i+c_{m-1}/c_m)^{-1}\right|=(4c_{m-1}c_m)^{-1}.$$

Theorem 6. *Assume that Conditions H and U hold. Then in case* (b) *there exists a continuous* $F\in\mathscr{F}$ *such that*

$$\lim_{n\to\infty} d_L(F_n(y_0;.),F)=0, \tag{16'}$$

uniformly with respect to $y_0\in Y$.

Proof. Uniformity of convergence with respect to $y_0\in Y$ in (16') can be justified as in the proof of Proposition 5, so that we shall no longer deal with it.

Since $(\mathscr{F}, d_L)$ is a compact metric space, and $F_n(y_0;.)\in\mathscr{F}$ for any $n\in N^*$, $y_0\in Y$, to prove (16') it is sufficient to show that all weak-convergent subsequences of the sequence $(F_n(y_0;.))_{n\in N^*}$ have the same limit. Thus, let us suppose there are G_1, $G_2\in\mathscr{F}$ and two increasing sequences of natural numbers $(n_k)_{k\in N^*}$ and $(n'_k)_{k\in N^*}$ such that

$$\lim_{k\to\infty} d_L(F_n(y_0;.),G_1)=\lim_{k\to\infty} d_L(F_{n'_k}(y_0;.),G_2)=0.$$

We have to prove that $G_1=G_2$.

On account of Chapman-Kolmogorov equation we can write[5] for $x>0$, $m\in N^*$,

$$\mathbf{P}(\eta_{m+1}\leqq x)=\int_0^1 \mathbf{P}(\eta_{m+1}\leqq x|\eta_m=z)\,\mathrm{d}_z\mathbf{P}(\eta_m\leqq z)$$

$$=\int_0^1\Big(\sum_{i\geqq x^{-1}-z} p_i(z)\Big)\,\mathrm{d}_z\mathbf{P}(\eta_m\leqq z),$$

whence

$$n_k^{-1}\sum_{m=2}^{n_k+1}\mathbf{P}(\eta_m\leqq x)=\int_0^1\Big(\sum_{i\geqq x^{-1}-z} p_i(z)\Big)\,\mathrm{d}_z F_{n_k}(y_0;z).$$

Assume for the moment that G_1 and G_2 are continuous. Therefore

$$\lim_{k\to\infty}F_{n_k}(y_0;x)=G_1(x),\qquad \lim_{k\to\infty}F_{n'_k}(y_0;x)=G_2(x),\qquad x\in R.$$

Then letting $k\to\infty$ in the preceding equation we obtain that

$$G_1(x)=\int_0^1\mathbf{P}(\eta_1\leqq x|\eta_0=y_0)\,\mathrm{d}G_1(y_0),\qquad x\in R. \tag{17}$$

(We are justified to pass to the limit under the integral sign since G_1 has been assumed to be continuous and the integrand has at most a countable set of points of discontinuity.) The above equation leads to

$$G_1(x)=\int_0^1\mathbf{P}(\eta_n\leqq x|\eta_0=y_0)\,\mathrm{d}G_1(y_0),\qquad x\in R,$$

for any $n\in N^*$, implying

$$G_1(x)=\int_0^1 F_{n'_k}(y_0;x)\,\mathrm{d}G_1(y_0),\qquad x\in R,\ k\in N^*,$$

whence, upon letting $k\to\infty$ we obtain that $G_1=G_2$.

Thus, to complete the proof we need only show that, if

$$\lim_{k\to\infty}d_L(F_{n_k}(y_0;.),G)=0,$$

then G is continuous. Let $C(G)$ be the set of points of continuity of G in $[0,1]$. We have to prove that $C(G)=[0,1]$. Consider first continuity at irrational points. Let $y=(i_1,i_2,\ldots)\in Y$. Whatever $0<m\leqq n$ we can write

$$\begin{aligned}\mathbf{P}(\alpha_{n-l}=i_l,\,1\leqq l\leqq m|\eta_{n-m})&=p_{i_m}(\eta_{n-m})\\ &\cdot p_{i_{m-1}}\left(\frac{1}{i_m+\eta_{n-m}}\right)\cdots p_{i_2}\left(\frac{1|}{|i_3}+\cdots+\frac{1|}{|i_{m-1}}+\frac{1|}{|i_m+\eta_{n-m}}\right)\\ &\cdot p_{i_1}\left(\frac{1|}{|i_2}+\cdots+\frac{1|}{|i_{m-1}}+\frac{1|}{|i_m+\eta_{n-m}}\right).\end{aligned}\tag{18}$$

[5] Although the p_i are defined only on Y, the integrals we write make sense since for any $m\in N^*$ the distribution function $\mathbf{P}(\eta_m\leqq z)$ is continuous at rational points, so that it assigns probability zero to the rationals in [0, 1]. This would enable us to give the p_i arbitrary values at rational points. In order to preserve continuity (in the topology of [0, 1]) at irrational points we can and do define p_i at a rational point as its minimum (with respect to Y) at that point. (See [6], p. 300.)

We want to show by making use of (18) that for any $\varepsilon>0$ there exists $m(\varepsilon)\in N^*$ such that

$$\mathbf{P}(\alpha_{n-l}=i_l,\ 1\leqq l\leqq m(\varepsilon)|\eta_{n-m(\varepsilon)})\leqq\varepsilon \tag{19}$$

for any $n\geqq m(\varepsilon)$. Clearly, (19) holds in case ($\mathfrak{a}$). The same is true in case ($\mathfrak{b}_i$) if for any $r\in N^*$ there exists a natural number $n\geqq r$ such that $i_n\neq i$. It remains to consider the case in which ($\mathfrak{b}_i$) holds and there exists an $r\in N^*$ such that $i_n=i$ for any $n\geqq r$. In this case (19) also holds since $p_i(y_i)<1$ and p_i is continuous at y_i. (Notice that $y_i=(i,i,\ldots)$.)

Inequality (19) implies that

$$\mathbf{P}\big(\alpha_{n-l}=i_l,\ 1\leqq l\leqq m(\varepsilon)\big)\leqq\varepsilon \tag{20}$$

for any $n\geqq m(\varepsilon)$.

Now, let us put $a_1(y)=1/(i_1+1)$ and

$$a_m(y)=\frac{1|}{|i_1}+\cdots+\frac{1|}{|i_{m-1}}+\frac{1|}{|i_m+1}, \qquad m\geqq 2,$$

$$b_m(y)=\frac{1|}{|i_1}+\cdots+\frac{1|}{|i_m}, \qquad m\in N^*,$$

and define

$$\delta(\varepsilon,y)=\min\big(|y-a_{m(\varepsilon)}(y)|,|y-b_{m(\varepsilon)}(y)|\big).$$

It is easily seen that for $n\geqq m(\varepsilon)$ the event $\big(|\eta_n-y|<\delta(\varepsilon,y)\big)$ implies the event $\big(\alpha_{n-l}=i_l,\ 1\leqq l\leqq m(\varepsilon)\big)$. Then, for any $x\in[0,1]$ such that $|y-x|<\delta(\varepsilon,y)$ and any $n\geqq m(\varepsilon)$ we can successively write

$$\begin{aligned}|F_n(y_0;y)-F_n(y_0;x)|&\leqq n^{-1}\sum_{m=1}^{n}|\mathbf{P}(\eta_m\leqq y)-\mathbf{P}(\eta_m\leqq x)|\\&\leqq n^{-1}\sum_{m=1}^{n}\mathbf{P}(|\eta_m-y|<|y-x|)\leqq n^{-1}\sum_{m=1}^{n}\mathbf{P}\big(|\eta_m-y|<\delta(\varepsilon,y)\big)\\&\leqq\big(m(\varepsilon)-1\big)n^{-1}+n^{-1}\sum_{m=m(\varepsilon)}^{n}\mathbf{P}\big(\alpha_{m-l}=i_l,\ 1\leqq l\leqq m(\varepsilon)\big).\end{aligned}$$

On account of (20) we deduce that

$$|F_n(y_0;y)-F_n(y_0;x)|\leqq m(\varepsilon)\,n^{-1}+\varepsilon. \tag{21}$$

Let $x_1<y<x_2$, $x_1,x_2\in C(G)$, such that $y-x_1<\delta(\varepsilon,y)$, $x_2-y<\delta(\varepsilon,y)$. It follows from (21) that

$$F_n(y_0;x_2)-F_n(y_0;x_1)\leqq 2\big(m(\varepsilon)\,n^{-1}+\varepsilon\big),$$

whence, upon letting $n\to\infty$,

$$G(x_2)-G(x_1)\leqq 2\varepsilon,$$

proving that G is continuous at y.

Let us now prove continuity at rational points. Let

$$x=\frac{1|}{|i_1}+\cdots+\frac{1|}{|i_k}$$

with $i_k \geqq 2$. Whatever $n \in N$ and $m \in N^*$ we have

$$\mathbf{P}(\alpha_n \geqq m \mid \eta_n) = \sum_{i \geqq m} p_i(\eta_n),$$

and then Condition U implies that for any $\varepsilon > 0$ there is an $m'(\varepsilon) \in N^*$ such that

$$\mathbf{P}(\alpha_n \geqq m'(\varepsilon)) < \varepsilon, \qquad n \in N. \tag{22}$$

Put

$$a'_m(x) = \frac{1|}{|i_1} + \cdots + \frac{1|}{|i_{k-1}} + \frac{1|}{|i_k - 1} + \frac{1|}{|1} + \frac{1|}{|m},$$

$$b'_m(x) = \frac{1|}{|i_1} + \cdots + \frac{1|}{|i_k} + \frac{1|}{|m}, \qquad m \in N^*,$$

and define

$$\delta'(\varepsilon, x) = \min\left(|x - a'_{m'(\varepsilon)}(x)|, |x - b'_{m'(\varepsilon)}(x)|\right).$$

It is easily seen that for $n \geqq k+2$ the event $(|\eta_n - x| < \delta'(\varepsilon, x))$ implies the event

$$\left(\alpha_{n-l} = i_l, 1 \leqq l \leqq k, \alpha_{n-k-1} \geqq m'(\varepsilon)\right)$$
$$\cup \left(\alpha_{n-l} = i_l, 1 \leqq l < k, \alpha_{n-k} = i_k - 1, \alpha_{n-k-1} = 1, \alpha_{n-k-2} \geqq m'(\varepsilon)\right).$$

Then for any $x' \in [0, 1]$ such that $|x - x'| < \delta'(\varepsilon, x)$ and any $n \geqq k+2$ we can write

$$|F_n(y_0; x) - F_n(y_0; x')| \leqq n^{-1} \sum_{m=1}^{n} \mathbf{P}(|\eta_m - x| < \delta'(\varepsilon, x))$$
$$\leqq (k+1) n^{-1} + n^{-1} \sum_{m=k+2}^{n} \left(\mathbf{P}(\alpha_{m-k-1} \geqq m'(\varepsilon)) + \mathbf{P}(\alpha_{m-k-2} \geqq m'(\varepsilon))\right),$$

and on account of (22) we deduce that

$$|F_n(y_0; x) - F_n(y_0; x')| \leqq (k+1) n^{-1} + 2\varepsilon.$$

Now continuity at x follows as in the previous case.

It remains to prove continuity at 0 and 1.

As to continuity at 0 we have to note that for $n \in N^*$ the event $(\eta_n < 1/m'(\varepsilon))$ implies the event $(\alpha_{n-1} \geqq m'(\varepsilon))$ so that for any $0 \leqq x < 1/m'(\varepsilon)$ we can write

$$F_n(y_0; x) \leqq n^{-1} \sum_{m=1}^{n} \mathbf{P}(\eta_m < 1/m'(\varepsilon)).$$

This leads to $G(x) \leqq \varepsilon$ if $x \in C(G)$, i.e. G is continuous at 0.

As to continuity at 1 notice that for $n \geqq 2$ the event $(\eta_n > (1 + 1/m'(\varepsilon))^{-1})$ implies the event $(\alpha_{n-1} = 1, \alpha_{n-2} \geqq m'(\varepsilon))$ so that for any $(1 + 1/m'(\varepsilon))^{-1} < x \leqq 1$ we can write

$$1 - F_n(y_0; x) \leqq n^{-1} \sum_{m=1}^{n} \mathbf{P}(\eta_m > (1 + 1/m'(\varepsilon))^{-1}) < n^{-1} + \varepsilon.$$

This leads to $G(x) \geqq 1 - \varepsilon$ if $x \in C(G)$, i.e. G is continuous at 1, q.e.d.

Proposition 7. *Under the corresponding assumptions in Proposition 5 and Theorem 6 the limiting distribution F is the only stationary distribution for $(\eta_n)_{n \in N}$.*

Proof. Actually, we have to prove that F is the only distribution function assigning probability 1 to Y which satisfies the equation

$$G(x)=\int_0^1 \mathbf{P}(\eta_1 \leqq x \mid \eta_0 = y_0)\,\mathrm{d}G(y_0), \qquad x\in R. \tag{23}$$

Clearly, the above equation is satisfied by the limiting F of case (a_i) and on account of (17) also by the limiting F of case (b). The argument of uniqueness is hidden in the proof of Theorem 6 (lines 13–17 on p. 205). Namely, if a distribution function G assigning probability 1 to Y satisfies (23), then this implies that

$$G(x)=\int_0^1 F_n(y_0; x)\,\mathrm{d}G(y_0), \qquad x\in R,$$

for any $n\in N^*$, whence upon letting $n\to\infty$ we deduce that $G(x)=F(x)$, $x\in R$ (case (b)) and $G(x)=F(x)$, $x\neq y_i$ (case (a_i)). The latter case leads immediately to $G(x)=F(x)$, $x\in R$, q.e.d.

Theorem 8. *Assume that Conditions H and U hold. Then the continuous limiting F of case* (b) *is either purely singular or identical to Gauss' absolutely continuous distribution function with density* $(\log 2)^{-1}/(1+x)$, $0\leqq x\leqq 1$. *The latter case occurs when* $p_i(y)=(y+1)/(y+i)(y+i+1)$, $i\in N^*$, $y\in Y$.

Proof. Assume for a contradiction that

$$F=c\,F_1+(1-c)\,F_2, \qquad 0<c<1,$$

where F_1 and $F_2\in\mathscr{F}$ are the absolutely continuous and the singular part of F respectively. If we write Eq. (23) in the operator form $G=VG$ with

$$(VG)(x)=\int_0^1 \sum_{i\geqq x^{-1}-z} p_i(z)\,\mathrm{d}G(z), \qquad x>0,$$

then we have

$$c(F_1-VF_1)=-(1-c)(F_2-VF_2).$$

The desired contradiction will be reached if we prove that VF_1 is absolutely continuous and VF_2 singular. For, on account of Proposition 7, neither F_1-VF_1 nor F_2-VF_2 can vanish identically.

Absolute continuity of VF_1 is immediate. Indeed, for $(j+1)^{-1}<x\leqq j^{-1}$, $j\in N^*$, we have

$$(VF_1)(j^{-1})-(VF_1)(x)=\sum_{i\geqq j}\int_0^1 F_1'(z)\,p_i(z)\,\mathrm{d}z$$
$$-\left(\sum_{i\geqq j+1}\int_0^{x^{-1}-j} F_1'(z)\,p_i(z)\,\mathrm{d}z+\sum_{i\geqq j}\int_{x^{-1}-j}^{1} F_1'(z)\,p_i(z)\,\mathrm{d}z\right)=\int_0^{x^{-1}-j} F_1'(z)\,p_j(z)\,\mathrm{d}z.$$

Therefore VF_1 is absolutely continuous, its density being, e.g.,

$$\begin{array}{ll} 0 & \text{for } x\leqq 0 \text{ or } x>1 \\ x^{-2}F_1'(x^{-1}-j)\,p_j(x^{-1}-j) & \text{for } (j+1)^{-1}<x\leqq j^{-1}, \quad j\in N^*. \end{array}$$

Let us now prove that VF_2 is singular. Let λ denote Lebesgue measure and μ the measure generated by F_2 (i.e. $\mu((-\infty, x])=F_2(x)$). As F_2 is singular, there are two disjoint sets A and B such that $A \cup B = [0, 1)$ and $\lambda(A) = \mu(B) = 0$. In order to simplify the writing, for any set Δ of real numbers put

$$\Delta + j = \{x : x - j \in \Delta\}, \quad \Delta - j = \{x : x + j \in \Delta\}, \quad j \in N^*,$$

and, if $0 \notin \Delta$,

$$\Delta^{-1} = \{x : x^{-1} \in \Delta\}.$$

As before for $(j+1)^{-1} < x \leqq j^{-1}, j \in N^*$, we have

$$(VF_2)(j^{-1}) - (VF_2)(x) = \int_0^{x^{-1}-j} p_j(z)\, dF_2(z) = \int_{[x, j^{-1}]^{-1} - j} p_j(z)\, dF_2(z),$$

implying

$$(V\mu)(\Delta) = \int_{\Delta^{-1} - j} p_j(z)\, dF_2(z) \tag{24}$$

for any Borel set Δ in $((j+1)^{-1}, j^{-1}]$. Now define $A_j = (A+j)^{-1}$, $B_j = (B+j)^{-1}$, $j \in N^*$, $A' = \bigcup_{j \in N^*} A_j$, $B' = \bigcup_{j \in N^*} B_j$. Clearly, $A_j, B_j \subset ((j+1)^{-1}, j^{-1}]$, $j \in N^*$, $A' \cap B' = \emptyset$, $A' \cup B' = (0, 1]$. It is easily seen that $\lambda(A) = 0$ implies that $\lambda(A_j) = 0$, $j \in N^*$. Thus $\lambda(A') = 0$. Next, taking $\Delta = B_j$ in (24) yields $(V\mu)(B_j) = 0$, $j \in N^*$ (on account of the fact that $\mu(B) = 0$), whence $(V\mu)(B') = 0$. But $\lambda(A') = (V\mu)(B') = 0$, means that VF_2 is singular.

Let us now deduce the form of F in the absolutely continuous case. We start from the equation

$$F\left(\frac{1}{j}\right) - F\left(\frac{1}{j+x}\right) = \int_0^x p_j(z)\, dF(z) \tag{25}$$

valid for $0 \leqq x \leqq 1, j \in N^*$. Put $F' = f$. Differentiation of (25) with respect to x yields

$$(j+x)^{-2} f\left(\frac{1}{j+x}\right) = p_j(x) f(x), \quad j \in N^*,$$

almost everywhere (with respect to Lebesgue measure) in $[0, 1]$. As $\sum_{j \in N^*} p_j(y) = 1$, $y \in Y$, we obtain the functional equation

$$f(x) = \sum_{j \in N^*} (j+x)^{-2} f\left(\frac{1}{j+x}\right) \tag{26}$$

almost everywhere in $[0, 1]$. Clearly, Gauss' probability density $f(x) = (\log 2)^{-1}/(1+x)$, $x \in [0, 1]$, satisfies (26), and this leads to the limiting F and the p_i described in the statement of the theorem. Thus, we have to prove that any probability density g on $[0, 1]$ satisfying (26) coincides almost everywhere with Gauss' one. To this end put $\nu(A) = \int_A g(x)\, dx$ for any Borel set A in $[0, 1]$, and consider the transformation T of Y onto itself defined by $Ty = y^{-1} \pmod 1$. Since

$$T^{-1}(0, x) = Y \cap \left(\bigcup_{j \in N^*} \left(\frac{1}{j+x}, \frac{1}{j}\right)\right), \quad 0 \leqq x \leqq 1,$$

and since on account of (26)

$$\sum_{j \in N^*} \nu\left(\left(\frac{1}{j+x}, \frac{1}{j}\right)\right) = \nu((0, x)), \quad 0 \leqq x \leqq 1,$$

it follows that T preserves ν. This proves uniqueness (almost everywhere) of the probability density satisfying (26). (See [10], p. 77.)

To complete the proof we need show that $\mathbf{p}=(p_i)_{i\in N^*}$ satisfies Condition H. (Condition U is obviously satisfied.) Here is a concise account of an elementary but tedious computation. Set $\Phi(x,x')=d(\mathbf{p}(x),\mathbf{p}(x'))$ for $0\leqq x, x'\leqq 1$. As $\Phi(x,x')=\Phi(x',x)$ we need only consider the case $x>x'$. We have $s_i(x)=i/(x+i+1)$, $i\in N^*$, and putting $i(x,x')=[(1+x)(x-x')^{-1}]$

$$\begin{aligned}\Phi(x,x')&=\sum_{i\leqq i(x,x')}(s_i(x')-s_i(x))+\sum_{i>i(x,x')}p_i(x')\\&=(x-x')\sum_{i\leqq i(x,x')}\frac{i}{(x+i+1)(x'+i+1)}+\frac{x'+1}{x'+i(x,x')+1}.\end{aligned}\tag{27}$$

Hence

$$\Phi(x,x')=\sum_{j=1}^{k-1}s_j(x')-\sum_{j=1}^{k}s_j(x)+1\tag{27'}$$

in the region $\{(x,x')\colon ((k-1)x-1)/k\leqq x'<(kx-1)/(k+1)\}$, $k\geqq 2$. Now, it is easily seen from (27′) that Φ is continuous[6], $\Phi(.,x')$ is increasing on $[x',1]$ and $\Phi(x,.)$ decreasing on $[0,x]$. It follows that $\varepsilon_0=\Phi(1,0)=2/3$. Next, one deduces from (27) that

$$(x-x')(C_1+|\log(x-x')|)\leqq\Phi(x,x')\leqq(x-x')(C_2+|\log(x-x')|),$$

where $-3<C_1<0$ and $0<C_2<2$ are absolute constants. Consequently, $\varepsilon_m=O(a^{-m})$, $m\in N$, for some $a>1$.[7] Therefore, even the stronger condition $\varepsilon_0<1$, $\sum_{m\in N}\varepsilon_m<\infty$, is satisfied rather than Condition H, q.e.d.

3. Chains of Infinite Order

3.1. A chain of infinite order may be viewed as a special random system with complete connections. (See [7], p. 186.) More precisely, if we consider the case of an N^*-valued homogeneous chain of infinite order, let P be a nonnegative function defined on $W\times N^*$, where $W=(N^*)^{(-N)}$, such that $\sum_{i\in N^*}P(w;i)=1$ for any $w\in W$. Under suitable conditions one proves the existence of a strictly stationary, doubly infinite sequence $(\xi_t)_{t\in Z}$ of N^*-valued random variables such that

$$\mathsf{P}(\xi_t=i\,|\,\xi_s, s<t)=P(\dots\xi_{t-2},\xi_{t-1};i),$$

i.e., P is the transition function of the process.

For any $i^{(r)}=(i_1,\dots,i_r)\in(N^*)^r$, $r\in N^*$, and $w=(\dots,i_{-n},\dots,i_{-1},i_0)\in W$, denote by $w+i^{(r)}$ the "path" $(\dots,i'_{-n},\dots,i'_{-1},i'_0)\in W$ for which $i'_{-n}=i_{-n+r}$, $n\in N$. A set of conditions ensuring the existence of an N^*-valued chain of infinite order with given transition function P is as follows. There exist a $\delta>0$ and a $j\in N^*$ such

[6] Interestingly enough, the points $(x,(kx-1)/(k+1))$, $1/k\leqq x\leqq 1$, $k\geqq 2$, are points of nondifferentiability of Φ. The double inequality that follows shows that the points (x,x), $0\leqq x\leqq 1$, are points of nondifferentiability, too.

[7] We conjecture that $\varepsilon_m=\Phi(c_{m-1}/c_m, c_m/c_{m+1})$, $m\in N^*$.

that $P(w;j)\geqq\delta$ for any $w\in W$. Next, $\sum_{n\in N^*}\varepsilon'_n<\infty$, where

$$\varepsilon'_n=2^{-1}\sup\sum_{i\in N^*}|P(w'+i^{(r)};i)-P(w''+i^{(r)};i)|,$$

the supremum being taken over all $w', w''\in W$, and all $i^{(r)}$, $r\geqq n$, containing j at least n times.

Moreover, under these conditions, the chain of infinite order is mixing in the sense that

$$|\mathsf{P}(B|A)-\mathsf{P}(B)|<r(t)$$

for any $A\in\mathscr{K}_0$, such that $\mathsf{P}(A)\neq 0$ and any $B\in\mathscr{K}^t$ where $r(t)$, that depends on δ and the ε'_n, tends to 0 as $t\to\infty$. Here $\mathscr{K}_0$ and $\mathscr{K}^t$ are the σ-algebras generated by the random variables ξ_s, $s\leqq 0$, and ξ_u, $u\geqq t$, respectively.

3.2. The results in Section 2 enable us to establish the existence of N^*-valued chains of infinite order under different conditions. A (weak) variant of mixing will be proved to hold, too. The basic device is to interpret a path

$$w=(\ldots, i_{-n}, \ldots, i_{-1}, i_0)\in W$$

as the continued fraction expansion (read inversely) of a $y\in Y$ and to define functions p_i by $p_i(y)=P(w;i)$ if y and w are connected as before (that is $y=(i_0, i_{-1}, \ldots)$). In this context the ε_m defined by (7) will be expressed as

$$\varepsilon_m=\sup_{(w,w')_m} d\big((P(w;i))_{i\in N^*}, (P(w';i))_{i\in N^*}\big), \qquad m\in N.$$

Here the notation $(w,w')_m$ means the last m components in the path w are the same as the last m components in the path w'. Next, Condition U amounts to the uniform convergence with respect to $w\in W$ of the series $\sum_{i\in N^*}P(w;i)$.

The existence theorem is as follows.

Theorem 9. *Assume that Conditions H and U hold and that* $P(\ldots, i, i; i)<1$ *for any* $i\in N^*$. *Then*

i) *There exists a strictly stationary, doubly infinite sequence* $(\xi_t)_{t\in Z}$ *on a probability space* $(\Omega, \mathscr{K}, \mathsf{P})$ *such that*

$$\mathsf{P}(\xi_t=i|\xi_s, s<t)=P(\ldots, \xi_{t-2}, \xi_{t-1}; i) \tag{28}$$

P-*almost surely for any* $t\in Z$, $i\in N^*$.

ii) *This is the only doubly infinite sequence for which* (28) *holds.*

Proof. i) On account of Theorem 6 we can construct a Y-valued strictly stationary Markov chain $(\eta_t)_{t\in Z}$ on a suitable probability space $(\Omega, \mathscr{K}, \mathsf{P})$, with stationary absolute distribution given by the limiting F, that is

$$\mathsf{P}(\eta_t\leqq x)=F(x), \qquad t\in Z,$$

and transition probability function given by

$$\mathsf{P}\left(\eta_{t+1}=\frac{1}{i+y}\,\middle|\,\eta_t=y\right)=p_i(y), \qquad y\in Y.$$

Define the function h on Y by

$$h(y) = \text{first digit in the continued fraction expansion of } y.$$

We shall prove that the random variables ξ_t defined by

$$\xi_t = h(\eta_{t+1}), \qquad t \in Z$$

satisfy (28).

First, for any $t \in Z$ we have

$$\mathsf{P}\left(\eta_{t+1} = \frac{1}{\xi_t + \eta_t}\right) = \sum_{i \in N^*} \mathsf{P}\left(\eta_{t+1} = \frac{1}{i + \eta_t}\right) = \sum_{i \in N^*} \int_0^1 p_i(y)\, \mathrm{d}F(y) = 1,$$

whence

$$\mathsf{P}(\eta_t = (\xi_{t-1}, \xi_{t-2}, \ldots)) = \mathsf{P}\left(\eta_s = \frac{1}{\xi_{s-1} + \eta_{s-1}}, s \leqq t\right) = 1. \tag{29}$$

Next, it is clear that to prove (28) it is sufficient to show that for any $l \in N^*$, $i_r \in N^*$, $1 \leqq r \leqq l$, we have

$$\mathsf{P}(\xi_t = i, \xi_{t-r} = i_r, 1 \leqq r \leqq l) = \int_{(\xi_{t-r} = i_r, 1 \leqq r \leqq l)} P(\ldots, \xi_{t-l-1}, \xi_{t-l}, \ldots, \xi_{t-1}; i)\, \mathrm{d}\mathsf{P}. \tag{30}$$

Put

$$x_l = \frac{1|}{|i_1} + \cdots + \frac{1|}{|i_l}, \qquad x_l^i = \frac{1|}{|i} + \frac{1|}{|i_1} + \cdots + \frac{1|}{|i_l}$$

and for any rational x in (0, 1)

$$\varphi(x) = \min\left(a'_\infty(x), b'_1(x)\right), \qquad \psi(x) = \max\left(a'_\infty(x), b'_1(x)\right).$$

(a'_∞ and b'_1 have been defined on p. 207.) Then by making use of (29) we can write

$$\mathsf{P}(\xi_t = i, \xi_{t-r} = i_r, 1 \leqq r \leqq l) = \mathsf{P}\left(\varphi(x_l^i) < \eta_{t+1} < \psi(x_l^i)\right)$$
$$= \int_0^1 \mathsf{P}\left(\varphi(x_l^i) < \eta_{t+1} < \psi(x_l^i) | \eta_t = y\right) \mathrm{d}F(y) = \int_{\varphi(x_l)}^{\psi(x_l)} p_i(y)\, \mathrm{d}F(y).$$

On the other hand we have

$$\int_{(\xi_{t-r} = i_r, 1 \leqq r \leqq l)} P(\ldots, \xi_{t-l-1}, \xi_{t-l}, \ldots, \xi_{t-1}; i)\, \mathrm{d}\mathsf{P}$$
$$= \int_{(\varphi(x_l) \leqq \eta_t < \psi(x_l))} p_i((i_1, \ldots, i_l, \xi_{t-l-1}, \ldots))\, \mathrm{d}\mathsf{P} = \int_{\varphi(x_l)}^{\psi(x_l)} p_i(y)\, \mathrm{d}F(y).$$

Therefore (30) holds so that the proof of i) is complete.

ii) It follows from the above that

$$\mathsf{P}(\xi_{t-r} = i_r, 1 \leqq r \leqq l) = F(\psi(x_l)) - F(\varphi(x_l)), \qquad t \in Z. \tag{31}$$

Therefore, the finite dimensional distributions of $(\xi_t)_{t \in Z}$ are completely determined by F. Thus uniqueness of $(\xi_t)_{t \in Z}$ satisfying (28) follows from that of F. (See Proposition 7.) q.e.d.

Remark. It is well known (28) implies that

$$\lim_{l\to\infty} \mathsf{P}(\xi_1=i|\xi_{-r}, 0\leqq r\leqq l)=P(\ldots,\xi_{-1},\xi_0;i)$$

P-almost surely. It is easily proved that if p_i is continuous on [0, 1] then

$$\lim_{l\to\infty} \mathsf{P}(\xi_1=i|\xi_{-r}=i_{-r}, 0\leqq r\leqq l)=P(\ldots,i_{-1},i_0;i)$$

for any $(\ldots,i_{-1},i_0)\in W$, provided the left side is defined for each $l\in N^*$.

3.3. Next, we have

Theorem 10. *Assume the conditions in Theorem 9. Let $\mathscr{K}_0$ denote the σ-algebra generated by the random variables ξ_s, $s\leqq 0$. Then*

$$\lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} \mathsf{P}(\xi_{t-r}=i_r, 1\leqq r\leqq l|A)=\mathsf{P}(\xi_{-r}=i_r, 1\leqq r\leqq l)$$

uniformly with respect to $l\in N^*$, $i_r\in N^*$, $1\leqq r\leqq l$ and $A\in\mathscr{K}_0$ such that $\mathsf{P}(A)\neq 0$.

Proof. It is sufficient to prove that

$$\lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} \mathsf{P}(\xi_{t-r}=i_r, 1\leqq r\leqq l|\xi_{-m}=i'_m, 1\leqq m\leqq k)=\mathsf{P}(\xi_{-r}=i_r, 1\leqq r\leqq l) \tag{32}$$

uniformly with respect to $l\in N^*$, $i_r\in N^*$, $1\leqq r\leqq l$, and $k\in N^*$, $i'_m\in N^*$, $1\leqq m\leqq k$, such that $\mathsf{P}(\xi_{-m}=i'_m, 1\leqq m\leqq k)\neq 0$.

We have

$$n^{-1}\sum_{t=1}^{n} \mathsf{P}(\xi_{t-r}=i_r, 1\leqq r\leqq l|\xi_{-m}=i'_m, 1\leqq m\leqq k)$$

$$=\frac{n^{-1}\sum_{t=1}^{n} \mathsf{P}(\xi_{t-r}=i_r, 1\leqq r\leqq l, \xi_{-m}=i'_m, 1\leqq m\leqq k)}{\mathsf{P}(\xi_{-m}=i'_m, 1\leqq m\leqq k)}$$

$$=\frac{n^{-1}\sum_{t=1}^{n} \mathsf{P}\big(\varphi(x_l)<\eta_t<\psi(x_l), \varphi(x'_k)<\eta_0<\psi(x'_k)\big)}{\mathsf{P}\big(\varphi(x'_k)<\eta_0<\psi(x'_k)\big)}$$

$$=\frac{n^{-1}\sum_{t=1}^{n}\int_{\varphi(x'_k)}^{\psi(x'_k)} \mathsf{P}\big(\varphi(x_l)<\eta_t<\psi(x_l)|\eta_0=y\big)\,\mathrm{d}F(y)}{\mathsf{P}\big(\varphi(x'_k)<\eta_0<\psi(x'_k)\big)}$$

$$=\frac{\int_{\varphi(x'_k)}^{\psi(x'_k)} \big(F_n(y;\psi(x_l))-F_n(y;\varphi(x_l))\big)\,\mathrm{d}F(y)}{F(\psi(x'_k))-F(\varphi(x'_k))}.$$

Now (32) follows on account of Theorem 6, continuity of F and Eq. (31), q.e.d.

3.4. To conclude, we notice that Theorems 9 and 10 can be viewed as properties of the continued fraction expansion of numbers $y\in Y$. As well known if $y=$

$(i_1(y), i_2(y), \ldots)$, then the $i_n(.)$ are random variables on the measurable space consisting of Y and the σ-algebra of Borel measurable sets in Y. The sequence $(i_n(.))_{n\in N^*}$ becomes a strictly stationary one when we choose as measure that generated by any limiting F in Theorem 6[8]. The doubly infinite sequence $(\xi_t)_{t\in Z}$ constructed in Theorem 9 is nothing but a two-sided version of such a strictly stationary $(i_n(.))_{n\in N^*}$. (See [3], p. 456.)

Acknowledgement. The authors wish to thank the referee for useful comments.

References

1. Billingsley, P.: Convergence of probability measures. New York: John Wiley 1968
2. Chatterji, S. D.: Maße, die von regelmäßigen Kettenbrüchen induziert sind. Math. Ann. **164**, 113–117 (1966)
3. Doob, J. L.: Stochastic processes. New York: John Wiley 1953.
4. Hardy, G. H.: Divergent series. Oxford: Clarendon Press 1949.
5. Harris, T. E.: On chains of infinite order. Pacific J. Math. **5**, 707–724 (1955)
6. Hobson, E. W.: The theory of functions of a real variable and the theory of Fourier's series. Vol. I, 3rd ed. Cambridge: University Press 1927
7. Iosifescu, M., Theodorescu, R.: Random processes and learning. Berlin-Heidelberg-New York: Springer 1969
8. Loève, M.: Probability theory. 3rd ed. Princeton: Van Nostrand 1963
9. Rényi, A.: Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hungar. **8**, 477–493 (1957)
10. Ryll-Nardzewski, C.: On the ergodic theorems (II) (Ergodic theory of continued fractions). Studia Math. **12**, 74–79 (1951)
11. Tortrat, A.: Calcul des probabilités. Paris: Masson 1963

[8] In particular, any nonatomic measure under which the $i_n(.)$ are independent and identically distributed is generated by a limiting purely singular F. This precises Satz 2 in [2].

M. Iosifescu
A. Spătaru
Centre of Mathematical Statistics of the
Romanian Academy
21 Calea Griviţei, Bucharest 12
Romania

(Received April 6, 1973)