Probab. Theory Relat. Fields 96, 503–520 (1993)

Probability Theory and Related Fields

Regeneration for chains with infinite memory

P. Ney [1,*] **and E. Nummelin** [2]

[1] University of Wisconsin, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706, USA
[2] University of Helsinki, Department of Mathematics, Hallituskatu 15, SF-00014 Helsinki, Finland

Received December 3, 1990; in revised form April 6, 1993

Summary. A regeneration structure is established for chains with infinite memory. The memory is required to decay only along a single recurrent path. When there are many recurrent paths (e.g. under conservativity) the construction yields a decomposition into regenerative recurrent classes.

Mathematics Subject Classification (1980): 60K05, 60J05

1 Introduction

Regeneration methods have a long history in the study of Markov chains, starting with Doeblin [D] for chains on countable state spaces. The general state space case has received attention in more recent times, with the development of regeneration structures for MC's in the setting of Harris and Orey (see Orey's book [O], and [A-Ne], [Nu1]). One of the limitations of this theory is that it requires the existence of so-called C-sets (see [O]) or minorizations, which may not always exist, or whose existence may be hard to establish. Since the existence of a regeneration automatically provides us with a tool for proving limit theorems, it is of interest to develop such constructions for processes not known to have reference measures with associated C-sets and minorizations.

One such process is the infinite memory chain (IMC), namely a sequence of random variables $\{X_n; n=0, 1, \ldots\}$, taking values in a countable set $\mathscr{S}$, for which there is *no fixed m* such that

$$P\{X_{n+1} \in \cdot \mid X_0, \ldots, X_n\} = P\{X_{n+1} \in \cdot \mid X_{n-m+1}, \ldots, X_n\}$$

for all $n \geqq m$. Such processes have been extensively studied in earlier works going back to [D-F] under the name "chains with complete connections". (See e.g. [I-G, N, B, L] for further references.)

* Research supported by NSF Grant DMS 89-01464

The sequences $\{X_n\}$ will in general be non-Markovian, but will, under some hypothesis, admit a regeneration structure. The process $\{(\ldots, X_{n-1}, X_n); n=0, 1, \ldots\}$ taking values in the sequence space $\mathscr{S}^{Z^-}$, ($Z^-=\{\ldots, -2, -1, 0\}$) *will* be Markovian, but will *not*, in general, have a regeneration structure. Both processes will be uniquely determined by an initial state in $\mathscr{S}^{Z^-}$ and the transition mechanism described below. We are interested in both these processes and in their relationship.

Our attention was drawn to these chains through an interesting paper by Lalley [L], who proved the existence of a regeneration for a class of IMC's. Shortly thereafter there was another interesting paper by Berbee [B], which did not refer explicitly to regeneration, but used random times and coupling in a similar spirit. These authors also pointed out the relevance of IMC's to the construction of Gibbs' measures.

In the present work we only require memory to decay along a single recurrent path of observed history of the process. Under various conditions such as conservativity there will be many recurrent paths, and our theory then gives a decomposition of these (and their domains of attraction) into ergodic classes, each with an associated regeneration structure and an (extremal) invariant measure on $\mathscr{S}^Z$ (Z = the integers).

The hypothesis in Lalley's, Berbee's and our papers differ in several respects. Lalley works with a finite alphabet, uniform memory loss, and stationary sequences. Berbee allows countably infinite alphabet, but requires the transition probabilities to be bounded from below by a positive constant, as well as uniform memory decay. As indicated above, non-uniform memory decay is allowed in our model. Our construction draws on ideas from our earlier work [A-N, Nu1], and is similar in spirit to Lalley's but there are differences in the constructions.

As with any sequence of r.v.'s $\{X_n; n=1, 2, \ldots\}$, the IMC induces a measure $P(\cdot)$ on $\mathscr{S}^{\mathbf{Z}^+}$ ($\mathbf{Z}^+=\{1, 2, \ldots\}$) (actually a family of such measures is generated depending on initial measures) and the process can then be identified with the deterministic shift map on $\mathscr{S}^{\mathbf{Z}^+}$.

Section 2 contains notation and definitions. Results are summarized in Sect. 3, the proofs are in Sect. 4.

There are many examples of infinite memory chains and their applications in [I-G] and [N]. For an application to number theory, see e.g. [K]. An example illustrating our particular set-up is in Sect. 5.

2 Notation and definitions

Let $\mathscr{S}$ be a countable state space,

$$x_m^n = (x_m, \ldots, x_n),\ x_i \in \mathscr{S},\ -\infty \leqq m \leqq n \leqq \infty,$$

$$\mathscr{S}_m^n = \{x_m^n : x_i \in \mathscr{S},\ i=m, \ldots, n\}.$$

Thus, for example

$$\mathscr{S}_{-\infty}^0 = \{(\ldots, x_{-2}, x_{-1}, x_0)\}$$

$$\mathscr{S}_1^\infty = \{(x_1, x_2, \ldots)\}.$$

We will sometimes call $\mathscr{S}$ the "alphabet" and x_m^n a "word".

If $x=x_k^\ell$ and $y=y_m^n$ with $k\leqq\ell$ and $m\leqq n$, write $xy=(x_k,\ldots,x_\ell,y_m,\ldots,y_n)$. If $y=x_{\ell+1}^n$ with $\ell<n$, then $xy=x_k^n$.

The datum of our model is a stochastic transition kernel $g:\mathscr{S}_{-\infty}^0\times\mathscr{S}\to\mathbb{R}$, $\sum_{x_1\in\mathscr{S}} g(x_{-\infty}^0,x_1)=1$. We consider a process $\{X_n;\,-\infty<n<\infty\}$ taking values in $\mathscr{S}$ with time homogeneous transition probabilities

$$P\{X_{n+1}=x_1\,|\,X_{-\infty}^n=x_{-\infty}^0\}=g(x_{-\infty}^0,x_1)\quad\text{for all } n. \tag{2.1}$$

Intuitively, think of $\mathscr{S}_{-\infty}^0$ as the "past". By iteration g induces

$$P\{X_{n+1}^{n+N}=x_1^N\,|\,X_{-\infty}^n=x_{-\infty}^0\}=g^N(x_{-\infty}^0,x_1^N). \tag{2.2}$$

Let

$$P_x(X_{n+1}^\infty\in\cdot)=G(x_{-\infty}^0,\cdot)=P\{X_{n+1}^\infty\in\cdot\,|\,X_{-\infty}^n=x_{-\infty}^0\},\quad\text{where}\quad x=x_{-\infty}^0. \tag{2.3}$$

Note that the process $\{X_n;n=0,1,\ldots\}$ on $\mathscr{S}$, is *not* Markovian, but $\{X_{-\infty}^n\}$ is a Markov chain (MC).

Definition 1 Call $\alpha=\alpha_{-\infty}^0\in\mathscr{S}_{-\infty}^0$ a *recurrent point for* g, if there exists an $x=x_{-\infty}^0\in\mathscr{S}_{-\infty}^0$ such that for every $m=1,2,\ldots$

$$P_x\{X_{n-m+1}^n=\alpha_{-m+1}^0\text{ for infinitely many } n\geqq 0\}=1. \tag{2.4}$$

Let $R=$ the set of recurrent points. If (2.4) holds we will say that x *is attracted to* α, and define the *attraction set of* α by

$$B_\alpha=\{x=x_{-\infty}^0: x\text{ is attracted to }\alpha\}. \tag{2.5}$$

The situation is quite analogous to the standard setup for Markov chains. The kernel $g(\cdot,\cdot)$ plays the role of the transition function, the point $x=x_{-\infty}^0$ is an "initial point", and our definition of recurrence is the standard notion of topological recurrence to the neighborhoods of α for the MC $\{X_{-\infty}^n\}$, with open sets being the cylinder sets in $\mathscr{S}_{-\infty}^0$. Under various standard conditions there will be lots of recurrent points. For example if $\{X_{-\infty}^n\}$ is (topologically) μ-*conservative* for some measure μ on $\mathscr{S}_{-\infty}^0$, (in the sense that for open sets $\mathcal{O}\subset\mathscr{S}_{-\infty}^0$, one has $P_x\{X_{-\infty}^n\in\mathcal{O}\text{ i.o.}\}=1$ for μ-a.e. $x\in\mathcal{O}$), then since the base for the topology is countable, almost all points in Supp (μ) are recurrent. The B_α's will be the analogs of the usual ergodic classes.

Continuity conditions on $g(\cdot,\cdot)$ are expressed in terms of the "memory" of the process. Namely, a function $f:\mathscr{S}_{-\infty}^0\to\mathbb{R}$ is continuous (at $y_{-\infty}^0$) if

$$\sup\{|f(x_{-\infty}^0)-f(y_{-\infty}^0)|:x_{-n}^0=y_{-n}^0\}\to 0\quad\text{as}\quad n\to\infty. \tag{2.6}$$

The simplest "loss of memory" is just the condition that

$$g(\cdot,x_1):\mathscr{S}_{-\infty}^0\to\mathbb{R}\text{ is continuous in the sense of (2.6) for each } x_1\in\mathscr{S}. \tag{2.7}$$

A stronger condition was imposed by Berbee [B]. He lets

$$e^{-r_n}=\inf\left\{\frac{g(x_{-\infty}^0,x_1)}{g(y_{-\infty}^0,y_1)}:x_{-n}^1=y_{-n}^1\right\}, \tag{2.8}$$

and requires $\sum r_n < \infty$ or $\sum_{n=1}^{\infty} \exp\left(-\sum_{i=1}^{n} r_i\right) = \infty$ for various of his results.

The continuity condition we will need in the present work is defined in terms of the *memory* of the process $\{X_n\}$ *along the sequence* $\alpha = \alpha_{-\infty}^{0}$, which we quantify by

$$\text{(2.9)} \qquad \gamma_m(\alpha) = \inf\left\{\frac{g^N(x_{-\infty}^0, z_1^N)}{g^N(y_{-\infty}^0, z_1^N)} : \text{all } x_{-\infty}^0, y_{-\infty}^0 \text{ such that } x_{-m+1}^0 = y_{-m+1}^0 = \alpha_{-m+1}^0, \text{ all } z_1^N, \text{ and all } N = 1, 2, \ldots\right\}.$$

Definition 2 We will say that g *loses memory along* α, or for short, satisfies condition $M(\alpha)$, if

$$(M(\alpha)) \qquad \gamma_m(\alpha) \nearrow 1 \quad \text{as} \quad m \to \infty.$$

Remark. This is just *equicontinuity* of $\log g^N(\cdot, z_1^N)$ at α, over $z_1^N \in \mathcal{S}_1^N$, $N = 1, 2, \ldots$

3 Results

The purpose of this paper is to study regeneration properties of the ∞-memory chain.

Definition 3 A *regeneration structure for* $\{X_n; n = 1, 2, \ldots\}$ is a sequence of randomized stopping times $\{T_i; i = 0, 1, \ldots\}$ and a measure ν on $\mathcal{S}_1^\infty$, such that
(i) $\{T_{i+1} - T_i; i = 0, 1, \ldots\}$ are i.i.d. random variables,
(ii) $X_0^{T_0}, X_{T_0+1}^{T_1}, X_{T_1+1}^{T_2}, \ldots$ are independent blocks,
(iii) $P\{X_{n+1}^{\infty} \in \cdot \mid T_i = n, \mathcal{F}_0^n\} = \nu(\cdot)$ for all $i \geq 0$, $n \geq 0$, where $\mathcal{F}_0^n$ = the σ-field generated by $(X_0, \ldots, X_n)$.
Call $\{T_i\}$ the *regeneration times* and $\nu(\cdot)$ the *regeneration measure*.

Theorem 1 *Let* $\alpha = \alpha_{-\infty}^0$ *be a recurrent point for* g, *and* B_α *be its attraction set as defined in* (2.5). *Let* $u = u_{-\infty}^0 \in B_\alpha$. *Let* $\{X_n\}$ *be the process generated by* g *and the initial point* u. *Assume that* g *satisfies condition* $M(\alpha)$. *Then there exists a regeneration structure for* $\{X_n\}$.

Remarks about Theorem 1 (i) The regeneration measure $\nu(\cdot)$ is specified in (4.3).
(ii) The simplest examples of regeneration structures are the return times to a fixed reference state for a Markov chain on a finite or countable state space. Clearly regenerations are not unique. Even for this simple MC there will be many different regenerations associated with return times to the various states; and there will in general be other regenerations not related to particular recurrent points. If the chain is not irreducible, there may be several ergodic classes, and these can be associated with recurrent points.
(iii) A similar situation prevails here. Theorem 1 is an assertion about a process $\{X_n\}$. As in the MC case, there is really a family of processes associated with various initial points in $H = \bigcup_{\alpha \in R} B_\alpha$, or with measures on this space. The theorem says that there exist regeneration structures for all of these processes. There

is no claim of uniqueness. Even a single process $\{X_n\}$, stemming from a single point u, may have many associated regeneration structures.
(iv) The regenerations we construct are associated with recurrent points, and the recurrent point plays an intimate role in the construction. This will become apparent in the proof. We will refer below to *regeneration structures* and *measures associated with a recurrent point* α. There is no claim of uniqueness, and there may also exist regenerations not associated with any particular recurrence point. However, we show in Theorem 2, that regeneration measures associated with recurrent points are either singular with respect to each other or are equivalent in the sense below.

Theorem 2 *Let v and v' be regeneration measures associated with recurrent points α and α' along which there is loss of memory. (It may be that $\alpha=\alpha'$ or $\alpha\neq\alpha'$). Let $\theta(x_1^\infty)=x_2^\infty$ denote the shift map on $\mathcal{S}_1^\infty$, and define the maximal measures*

$$\psi=\sum_{n=0}^{\infty} 2^{-n} v\circ\theta^{-n} \quad \text{and} \quad \psi'=\sum_{n=0}^{\infty} 2^{-n} v'\circ\theta^{-n}.$$

Assume that g satisfies $M(\alpha)$ and $M(\alpha')$. Then
(i) *ψ and ψ' are either equivalent or singular.*
(ii) *Either $B_\alpha\cap B_{\alpha'}=\phi$ or $B_\alpha=B_{\alpha'}$.*
(iii) *The number of attraction classes B_α is countable.*

Remarks about Theorem 2 (i) The relation between a recurrent point and a regeneration structure is established in the definition (4.3) of regeneration measure.
(ii) The decomposition asserted in the theorem depends purely on g. The recurrence points α and the sets B_α are defined in terms of g alone. The measures ψ and ψ' of course are constructed from processes $\{X_n\}$, but nothing is assumed about ψ or ψ' other than that they stem from recurrent points α and α'.
(iii) The assertion that ψ and ψ' are either singular or equivalent is made only for measures ψ, ψ' associated with recurrent points. There may be other regeneration structures with associated maximal measures, and we do not assert the "singular/equivalent" dichotomy for these. In the case of φ-irreducible Markov chains, all such maximal measures satisfy this dichotomy, but *we leave it as an open problem* whether this holds for ∞-memory chains.
(iv) It may be that $\alpha\notin B_\alpha$ and hence that $R-H\neq\emptyset$. Theorem 2 gives a partition of H into attracting sets. (Recall $R=$ the set of recurrent points and $H=\bigcup_{\alpha\in R} B_\alpha$.)

Some remarks on invariant measures. The regeneration structure immediately yields an invariant measure on $\mathcal{S}_1^\infty$, under the shift map. Namely, define a measure π^+ on $\mathcal{S}_1^\infty$ by

$$\pi^+(A)=\int\sum_{n=1}^{T_0} 1_A(x_n^\infty)\,dv(x_1^\infty),$$

where T_0 is the first regeneration time, $v(\cdot)$ is the regeneration measure, and $1_A(\cdot)$ is the indicator function. Then π^+ is a σ-finite measure and $\pi^+=\pi^+\circ\theta^{-1}$. Via the shift map, extend $\pi^+(\cdot)$ to a measure π on $\mathcal{S}_{-\infty}^\infty$, which will be shift invariant on $\mathcal{S}_{-\infty}^\infty$, and let π^- be the projection of π on $\mathcal{S}_{-\infty}^0$. Then π^- is a σ-finite invariant measure for the Markov chain $\{X^n_{-\infty}\}$. These facts are true for *any* sequence $\{X_n\}$ with a regeneration structure, and are not particular to our set up.

The question of uniqueness of invariant measures is more complicated. Berbee [B] has proved uniqueness under his condition $\sum_{n=1}^{\infty} \exp\left(-\sum_{1}^{\infty} r_i\right)=\infty$, with r_i defined in (2.8). Bramson and Kalikow [Br-K] have shown (by a counterexample) that mere continuity of g in the sense of (2.7) is not sufficient, even when accompanied by a strong mixing condition.

In our setting, since there may be many ergodic classes, there is clearly no uniqueness. Any mixture of (mutually singular) invariant measure supported by different B'_αs will be invariant.

4 Proofs

Proof of Theorem 1 There will be regeneration structures associated with each class B_α and each initial point $u=u^0_{-\infty}\in B_\alpha$. *Fix α and u in the following argument.*

For each i, let $\{\gamma_{m,i}; m=1,2,\dots\}$ be a sequence satisfying

$$0\leqq \gamma_{m,i}\nearrow 1 \quad \text{as} \quad m\to\infty, \tag{4.1}$$

with

$$\gamma_{m,0}=\gamma_m(\alpha) \quad \text{as in (2.9)},$$

and the remaining $\gamma_{m,i}$'s to be specified later. Since α is fixed throughout this proof we may not always display it.

Fix $0<p<1/2$, and let

$$m_i=\inf\{m: \gamma_{m,i}(\alpha)>2p\}. \tag{4.2}$$

Since g loses memory along α, $m_i<\infty$ for each i.

Recalling the notation in (2.3), let

$$\nu(\cdot)=P(\cdot\,|X^0_{-\infty}=\alpha^0_{-\infty})=P_\alpha(\cdot)\doteq G(\alpha,\cdot). \tag{4.3}$$

We will prove that $\nu(\cdot)$ *is a regeneration measure for* $\{X_n\}$.

Recall $u\in B_\alpha=$ the attraction set of α. The proof is based on the fact that a suitable sequence of measures on $\mathcal{S}_1^\infty$ can be split away from $G(u^0_{-\infty}x^n_1,\cdot)$. Starting with the (fixed) initial state $u=u^0_{-\infty}$, let

$$\tau_0=\inf\{n: X^n_{n-m_0+1}=\alpha^0_{-m_0+1}\}. \tag{4.4}$$

Note that $\tau_0<\infty$ a.s. since α is a recurrent point. Thus

$$X^{\tau_0}_{\tau_0-m_0+1}=\alpha^0_{-m_0+1}. \tag{4.5}$$

We abbreviate

$$X_1^{\tau_0-m_0}=X^{(0)}. \tag{4.6}$$

Thus τ_0 is the first time that the sequence $\{X_1^n\}$ "sees" the word $\alpha_{-m_0+1}^0$. In general, let

$$(4.7)\qquad \tau_i = \inf\{n: X_{n-m_i+1}^n = \alpha_{-m_i+1}^0\} = \text{"the first time that } \{X_1^n\} \text{ sees } \alpha_{-m_i+1}^0\text{"}.$$

(Again $\tau_i < \infty$ a.s. since α is a recurrent point.) Let $X^{(i)} = X_1^{\tau_i - m_i}$ = the sample path of $\{X_n\}$ until $\tau_i - m_i$. Thus

$$(4.8)\qquad X_1^{\tau_i} = X_1^{\tau_i - m_i} X_{\tau_i - m_i + 1}^{\tau_i} = X^{(i)} \alpha_{-m_i+1}^0.$$

Let

$$(4.9)\qquad P^{(0)}(\cdot) = P_u(\cdot) = G(u, \cdot)$$

and

$$(4.10)\qquad Q^{(0)}(X_1^n, \cdot) = P^{(0)}\{X_{n+1}^\infty \in \cdot \mid X_1^n\} = G(u_{-\infty}^0 X_1^n, \cdot).$$

Recall (4.5), (4.6), and let

$$(4.11)\qquad P^{(1)}(\cdot) = (1-p)^{-1}[Q^{(0)}(X^{(0)} \alpha_{-m_0+1}^0, \cdot) - p\,\nu(\cdot)].$$

By condition $M(\alpha)$, for all $x_{-\infty}^0$, y_1^{n-m}, α_{-m+1}^0

$$(4.12)\qquad \gamma_m G(u_{-\infty}^0 y_1^{n-m} \alpha_{-m+1}^0, \cdot) \leqq G(x_{-\infty}^0 \alpha_{-m+1}^0, \cdot) \leqq \gamma_m^{-1} G(u_{-\infty}^0 y_1^{n-m} \alpha_{-m+1}^0, \cdot).$$

In particular, taking $m = m_0$, $n = \tau_0$, and $y_1^{n-m_0} = X^{(0)}$,

$$\gamma_{m_0} G(u_{-\infty}^0 X^{(0)} \alpha_{-m_0+1}^0, \cdot) \leqq G(x_{-\infty}^0 \alpha_{-m_0+1}^0, \cdot) \leqq \gamma_{m_0}^{-1} G(u_{-\infty}^0 X^{(0)} \alpha_{-m_0+1}^0, \cdot),$$

and hence by (4.10)

$$(4.13)\qquad \gamma_{m_0} Q^{(0)}(X^{(0)} \alpha_{-m_0+1}^0, \cdot) \leqq G(\alpha_{-\infty}^0, \cdot) \leqq \gamma_{m_0}^{-1} Q^{(0)}(X^{(0)} \alpha_{-m_0+1}^0, \cdot),$$

namely

$$(4.14)\qquad \gamma_{m_0} \leqq \frac{Q^{(0)}(X^{(0)} \alpha_{-m_0+1}^0, \cdot)}{\nu(\cdot)} \leqq \gamma_{m_0}^{-1}.$$

(Note that here we have non-random bounds for the random quantity in the middle.) Therefore $P^{(1)}$ is a probability measure on $\mathcal{S}_1^\infty$. Similarly, replacing m_0 by arbitrary m and $X^{(0)}$ by y_1^{n-m} in (4.13), recalling $\gamma_{m,0} = \gamma_m$,

$$(4.15)\qquad \gamma_{m,0} \leqq \frac{Q^{(0)}(y_1^{n-m} \alpha_{-m+1}^0, \cdot)}{\nu(\cdot)} \leqq \gamma_{m,0}^{-1}.$$

These inequalities will be needed later. In terms of the probability measure $P^{(1)}$, define

$$Q^{(1)}(X_1^n,\cdot)=P^{(1)}\{X_{n+1}^{\infty}\in\cdot\,|\,X_1^n\},\ n=1,2,\dots,$$

and

$$P^{(2)}(\cdot)=(1-p)^{-1}[Q^{(1)}(X^{(1)}\,\alpha^0_{-m_1+1},\cdot)-p\nu(\cdot)].$$

(We cannot *yet* assert that $P^{(2)}$ is a probability measure. This will follow from the lemma below.)

We now proceed to define $(P^{(i)}, Q^{(i)})$, $i=1,2,\dots$ inductively. Having defined $(P^{(i)}, Q^{(i)})$ let

$$P^{(i+1)}(\cdot)=(1-p)^{-1}[Q^{(i)}(X^{(i)}\,\alpha^0_{-m_i+1},\cdot)-p\nu(\cdot)], \tag{4.17}$$

and

$$Q^{(i+1)}(X_1^n,\cdot)=P^{(i+1)}\{X_{n+1}^{\infty}\in\cdot\,|\,X_1^n\}. \tag{4.18}$$

The fact that $P^{(i)}(\cdot)$ are probability measures and the consistency of the above definitions follows from Lemma 1 below.

Definition 4 Call a word y_1^n *$P^{(i)}$-observable* if $P^{(i)}\{X_1^n=y_1^n\}>0$.

Lemma 1 *Assume that g satisfies condition $M(\alpha)$. Then for each $i=0,1,2,\dots$ there exists a (non-random) sequence $\{0<\gamma_{m,i};\ m=1,2,\dots\}\nearrow 1$ as $m\to\infty$ such that for all $n\geqq m\geqq 1$*

$$\gamma_{m,i}\leqq\frac{Q^{(i)}(y_1^{n-m}\,\alpha^0_{-m+1},\cdot)}{\nu(\cdot)}\leqq\gamma_{m,i}^{-1} \tag{4.19}$$

for all $P^{(i)}$-observable y_1^{n-m}.

The proof of the lemma is at the end of the proof of Theorem 1.

We have observed that $P^{(1)}$ is a probability measure. Proceeding by induction, suppose $P^{(i)}$ is a probability measure. Then $Q^{(i)}$ is a stochastic kernel, and by the lemma $P^{(i+1)}$ is a probability measure. Hence we have

Corollary 1 *For all $i=0,1,\dots,P^{(i)}$ are probability measures and $Q^{(i)}$ are stochastic kernels.*

The construction of the regeneration times and regeneration measure proceeds via a series of "cycles" consisting of "stages". Regeneration occurs at the end of a cycle, after a random number of these stages, stopped when a certain outcome occurs. A typical cycle is illustrated in the flow diagram below.

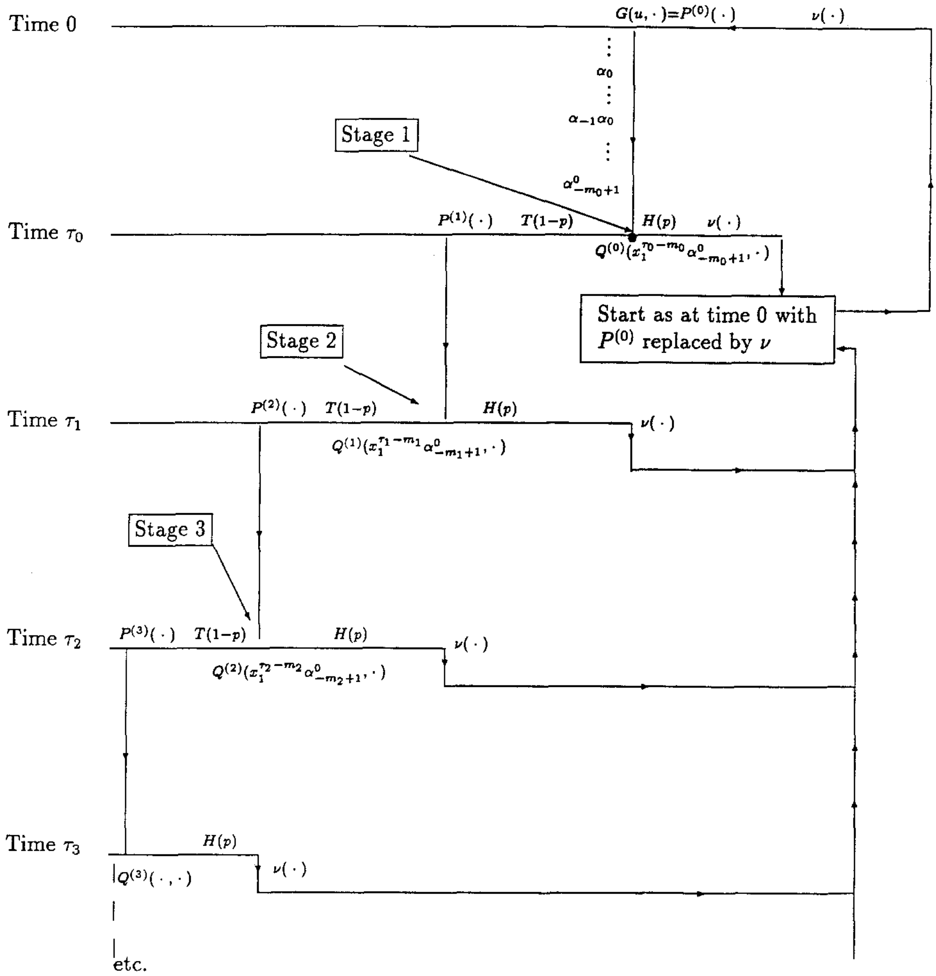
Stage 0 Starting with initial state $u^0_{-\infty}=u\in\mathscr{S}^0_{-\infty}$, generate $(X_1,X_2,\dots)\in\mathscr{S}_1^{\infty}$ via the kernel $g(\cdot,\cdot)$. Let $\gamma_{m,0}=\gamma_m(\alpha)$ as in the theorem, $\alpha=\alpha^0_{-\infty}$ be the given recurrent point, $m_0=\inf\{m:\gamma_m(\alpha)>0\}$, and

$$\tau_0=\inf\{n>m_0: X^n_{n-m_0+1}=\alpha^0_{-m_0+1}\}. \tag{4.20}$$

Recall $\gamma_{m_0}(\alpha^0_{-m_0+1})\geqq 2p$ (by definition). Note that since α is a recurrent point

$$P\{\tau_0<\infty\}=1.$$

One cycle in the regeneration construction



Colloquially, $\{X^n_{-\infty}\}$ first "sees the word $\alpha^0_{-m_0+1}$ at time τ_0". The conditional distribution of $X^\infty_{\tau_0+1}$, given the history up to τ_0, is

(4.20) $$P\{X^\infty_{n+1}\in\cdot\mid\tau_0=n, X^0_{-\infty}=u^0_{-\infty}, X_1^{n-m_0}=x_1^{n-m_0}, X^n_{n-m_0+1}=\alpha^0_{-m_0+1}\}.$$

Referring to the diagram, we are now at the point marked by a heavy dot ●.

By (4.14)

(4.22) $$Q^{(0)}(X_1^{\tau_0-m_0}\alpha^0_{-m_0+1},\cdot)\geqq 2p\nu(\cdot).$$

Let $P^{(1)}$ be as specified in the construction of $\{P^{(i)}; i=0,1,\ldots\}$. Namely

(4.23) $$P^{(1)}(\cdot)=\frac{Q^{(0)}(X_1^{\tau_0-m_0}\alpha^0_{-m_0+1},\cdot)-p\nu(\cdot)}{1-p}.$$

By the corollary, $P^{(1)}$ is a probability measure. We may thus "split" $Q^{(0)}(X_1^{\tau_0-m_0}\alpha_{-m_0+1}^0,\cdot)$ into a mixture of two measures:

$$\begin{cases} \nu(\cdot) & \text{with probability } p, \\ P^{(1)}(\cdot) & \text{with probability } 1-p. \end{cases} \tag{4.24}$$

Let us express this in terms of coin tossing and refer to it as

Stage 1 Toss a coin with Prob $(H)=p$.
(a) If (H), distribute $X_{\tau_0+1}^{\infty}$ according to $\nu(\cdot)$. In this case $T_0=\tau_0$ is the first regeneration time. Now start again as at time 0, but with the initial measure $G(u,\cdot)$ replaced by ν.
(b) If (T) at stage 1 of cycle 1, distribute $X_{\tau_0+1}^{\infty}$ according to $P^{(1)}$. The conditional distribution of $X_{\tau_1+1}^{\infty}$ is $Q^{(1)}(X^{(1)}\alpha_{-m_1+1}^0,\cdot)$, where $X^{(1)}$ is the observed sequence up to time τ_1-m_1. Now (conditioned on (T) at stage 1 of cycle 1) proceed at time τ_1 to

Stage 2 Toss a coin with Prob $(H)=p$.
(a) If (H) distribute $X_{\tau_1+1}^{\infty}$ according to $\nu(\cdot)$, and we have regeneration at time τ_1.
(b) If (T), distribute $X_{\tau_1+1}^{\infty}$ according to $P^{(2)}$. The fact that $P^{(2)}$ is a probability measure, and that $Q^{(1)}(X^{(1)}\alpha_{-m_1+1}^0,\cdot)$ can be split as indicated, follows from Lemma 1 and Corollary 1. Define m_2 and τ_2 as indicated at the start of this section.

The conditional distribution of $X_{\tau_2+1}^{\infty}$ is $Q^{(2)}(X^{(2)}\alpha_{-m_2+1}^0,\cdot)$. Now (conditioned on (T) at stages 1 and 2), proceed at time τ_2 to stage 3.

Proceeding inductively, suppose that the initial coin tosses at stages 1, ..., i all come out (T). Then, defining m_i and τ_i as before, namely

$$\tau_i=\inf\{n: X_{n-m_i+1}^{n}=\alpha_{-m_i+1}^0\},$$

the conditional distribution of $X_{\tau_i+1}^{\infty}$ is

$$Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0,\cdot).$$

One then proceeds to

Stage i+1 Toss the p-coin. Distribute $X_{\tau_i+1}^{\infty}$ according to

$$\begin{cases} \nu(\cdot) & \text{if } (H), \\ P^{(i+1)} & \text{if } (T). \end{cases} \tag{4.25}$$

(All coin tosses are independent.) If (H), we have regeneration at time τ_i; otherwise start stage $i+2$.

$$T_0=\inf\{\tau_i:(H) \text{ at time } \tau_i \text{ by } p\text{-coin}\}.$$

The first cycle ends at T_0.

The law of $X_{T_0+1}^{\infty}$ is ν. Now start again (the next cycle) at time T_0+1, but with the initial measure $G(u,\cdot)$ replaced by ν, and generate the next regeneration time T_1, and inductively $T_0, T_1, T_2, \ldots$ Due to the geometric probabilities produced by the independent coint tosses, it is easy to see that $P\{T_i<\infty\}=1$, and the other properties required of a regenerative structure are satisfied by construction.

This proves Theorem 1. □

Proof of Lemma 1 For $i=0$, (4.19) is just (4.15). Proceeding by induction assume (4.19) for arbitrary fixed i. By definition of $Q^{(i)}$

$$Q^{(i+1)}(y_1^{n-m}\alpha_{-m+1}^0,\cdot)=\frac{P^{(i+1)}\{y_1^{n-m}\alpha_{-m+1}^0\cap\theta^n(\cdot)\}}{P^{(i+1)}\{y_1^{n-m}\alpha_{-m+1}^0\}},\tag{4.26}$$

and by definition of $P^{(i+1)}$ this

$$=\frac{Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0,y_1^{n-m}\alpha_{-m+1}^0\cap\theta^n(\cdot))-p\nu(y_1^{n-m}\alpha_{-m+1}^0\cap\theta^n(\cdot))}{Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0,y_1^{n-m}\alpha_{-m+1}^0)-p\nu(y_1^{n-m}\alpha_{-m+1}^0)}$$

$$\doteq\frac{c-d}{a-b}\text{ (say).}$$

By the induction hypothesis (4.19) and the definition of m_i

$$2p\leqq\frac{Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0,\cdot)}{\nu(\cdot)}\leqq\frac{1}{2p},\tag{4.27}$$

and hence (noting the factor p in b)

$$2\leqq\frac{a}{b}\leqq\frac{1}{2p^2}.\tag{4.28}$$

Remark. $Q^{(i)}$ (and $P^{(i)}$) depend on the sample path of the process and are thus random quantities. Hence so is a/b. However the $\gamma_{m,i}$ are *not* random and hence the bounds in (4.27) and (4.28) are *not* random.

Note that a and b are the probabilities of particular sets, c and d are measures. Let $Q(x_1^n,\cdot)=\nu(X_{n+1}^\infty\in\cdot\,|\,X_1^n=x_1^n)$. Now by definition of conditional probability

$$\frac{c}{a}=Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0\,y_1^{n-m}\alpha_{-m+1}^0,\cdot)\doteq u\text{ (say),}$$

and

$$\frac{d}{b}=Q(y_1^{n-m}\alpha_{-m+1}^0,\cdot)=v\text{ (say).}$$

Thus (4.26) can be rewritten as

$$Q^{(i+1)}(y_1^{n-m}\alpha_{-m+1}^0,\cdot)=\frac{c-d}{a-b}=\frac{au-bv}{a-b}.\tag{4.29}$$

Now

$$\begin{aligned}\frac{u}{v}&=\frac{Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0\,y_1^{n-m}\alpha_{-m+1}^0,\cdot)}{Q(y_1^{n-m}\alpha_{-m+1}^0,\cdot)}\\&=\left(\frac{Q^{(i)}(X^{(i)}\alpha_{-m_i+1}^0\,y_1^{n-m}\alpha_{-m+1}^0,\cdot)}{\nu(\cdot)}\right)\left(\frac{Q(y_1^{n-m}\alpha_{-m+1}^0,\cdot)}{\nu(\cdot)}\right)^{-1}.\end{aligned}\tag{4.30}$$

Applying the induction hypothesis to the first factor in (4.30) and applying (4.16) to the second we get

$$(4.31) \qquad \gamma_{m,0}\,\gamma_{m,i} \leqq \frac{u}{v} \leqq \gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1},$$

and hence

$$(4.31) \qquad \frac{\left(\frac{a}{b}\gamma_{m,0}\,\gamma_{m,i}-1\right)v}{\frac{a}{b}-1} \leqq \frac{au-bv}{a-b} \leqq \frac{\left(\frac{a}{b}\gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1}-1\right)v}{\frac{a}{b}-1}.$$

Abbreviate

$$\frac{Q^{(i)}(y_1^{n-m}\,\alpha_{-m+1}^0,\cdot)}{v(\cdot)} = R_m^{(i)}.$$

Thus by (4.29) and the definition of v

$$(4.32) \qquad \frac{\frac{a}{b}\gamma_{m,0}\,\gamma_{m,i}-1}{\frac{a}{b}-1}\cdot\frac{Q(y_1^{n-m}\,\alpha_{-m+1}^0,\cdot)}{v(\cdot)}$$

$$\leqq R_m^{(i+1)} \leqq \frac{\frac{a}{b}\gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1}-1}{\frac{a}{b}-1}\cdot\frac{Q(y_1^{n-m}\,\alpha_{-m+1}^0,\cdot)}{v(\cdot)}$$

and by (4.16)

$$(4.33) \qquad \frac{\frac{a}{b}\gamma_{m,0}\,\gamma_{m,i}-1}{\frac{a}{b}-1}\gamma_{m,0} \leqq R_m^{(i+1)} \leqq \frac{\frac{a}{b}\gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1}-1}{\frac{a}{b}-1}\gamma_{m,0}^{-1}.$$

Now a/b is a random quantity, but by (4.28) we have (uniformly) that $2 \leqq \frac{a}{b} < \infty$. Hence a little calculating shows that

$$(2\gamma_{m,0}\,\gamma_{m,i}-1)\,\gamma_{m,0} \leqq R_m^{(i+1)} \leqq [(2\gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1}-1)^{-1}\,\gamma_{m,0}]^{-1}.$$

Thus taking $\gamma_{m,i+1} = \min\{2\gamma_{m,0}\,\gamma_{m,i}-1)\,\gamma_{m,0},\ (2\gamma_{m,0}^{-1}\,\gamma_{m,i}^{-1}-1)^{-1}\,\gamma_{m,0}\}$, and recalling that $\gamma_{m,0} \nearrow 1$, $\gamma_{m,i} \nearrow 1$ as $m \to \infty$, we can choose m so that (4.19) holds for $i+1$. □

Proof of Theorem 2 Part (i) Let α and α' be recurrent points for g; let v, v' and ψ, ψ' be regeneration and maximal measures associated with α, α' respectively; and let $u=u_{-\infty}^{0}$, $u'=(u')_{-\infty}^{0}$ be points in $\mathscr{S}_{-\infty}^{0}$ attracted to α, α', and m_0, m_0' be as defined in (4.1) for α, α'. Let

$$W=\{x_1^\infty : x_{n+1}^{n+m_0}=\alpha_{-m_0+1}^{0} \text{ i.o.}(n)\},$$

with W' defined similarly for α'. By definition of recurrence

$$G(u, W)=1. \tag{4.34}$$

Suppose $\psi \not\perp \psi'$. We claim that this implies

$$G(u', W)=1, \tag{4.35}$$

and similarly that

$$G(u, W')=1. \tag{4.36}$$

Suppose we have (4.35) and (4.36). Let $D\subset\mathscr{S}_1^\infty$ be such that $\psi(D)>0$. Then by regeneration

$$\nu\{X_n^\infty\in D \text{ i.o.}\}=1 \tag{4.37}$$

and hence

$$\psi\{X_n^\infty\in D \text{ i.o.}\}=1. \tag{4.38}$$

Let

$$D(w')=\{x_1^\infty : x_{k-m_0'+1}^{k}=\alpha'^{0}_{-m_0'+1},\, x_{k+n}^\infty\in D \text{ for some } k, \text{and infinitely many } n\},$$

i.e. this is the set of sequences in $\mathscr{S}_1^\infty$ that "enter D i.o. after they have seen $\alpha'_{-m_0'+1}$." Then by (4.36) and (4.38)

$$G(u, D(w'))=1. \tag{4.39}$$

Let $\tau=\inf\{n:\ X_{n-m_0'+1}^{n}=\alpha'^{0}_{-m_0'+1}\}$. Then

$$G(uX_1^{\tau_1-m_0'}\alpha'^{0}_{-m_0'+1}, \{X_n^\infty\in D \text{ i.o.}\})>0. \tag{4.40}$$

But by $M(\alpha')$, for any set $B\subset\mathscr{S}_1^\infty$

$$G(x_{-\infty}^{-m_0'}\alpha'^{0}_{-m_0'+1}, B)>0 \Rightarrow \nu'(B)>0, \tag{4.41}$$

and applied to (4.40) this

$$\begin{aligned}&\Rightarrow \nu'\{X_n^\infty\in D \text{ i.o.}\}>0\\ &\Rightarrow \nu'\{X_n^\infty\in D \text{ i.o.}\}=1 \Rightarrow \psi'(D)>0.\end{aligned} \tag{4.42}$$

Similarly $\psi'(D)>0$ implies $\psi(D)>0$ and we conclude that $\psi\sim\psi'$.

It remains to prove that if $\psi\not\perp\psi'$ then (4.35) also holds.

Certainly $G(u, W)=1\Rightarrow \nu(W)>0$

$$\begin{aligned}&\Rightarrow \nu(W)=1 \Rightarrow \nu\circ\theta^n(W)=1\\ &\Rightarrow \psi(W)=1.\end{aligned} \tag{4.43}$$

Now if $\psi \not\perp \psi'$ then this

$$(4.45) \quad \begin{aligned} &\Rightarrow \psi'(W) > 0 && \text{(trivially)} \\ &\Rightarrow \nu' \circ \theta^{-n}(W) > 0 && \text{(by definition of } \psi') \\ &\Rightarrow \nu'(W) = 1 && \text{(by regeneration)} \\ &\Rightarrow G(u', W) = 1. \end{aligned}$$

This proves part (i) of the theorem.

Part (ii). Next suppose $u \in B_\alpha \cap B_{\alpha'}$ with $\alpha^0_{-m_0+1} \neq \alpha'^0_{-m_0+1}$. Then for a process initiated at $u^0_{-\infty} = u$, there will be regeneration measures ν and ν' associated with α and α'. Hence for any $A \subset \mathcal{S}_1^\infty$

$$\nu\{X_n^\infty \in A \text{ i.o.}\} = 1 \Leftrightarrow \nu'\{X_n^\infty \in A \text{ i.o.}\} = 1.$$

We claim that

$$\psi(A) > 0 \Leftrightarrow \psi'(A) > 0.$$

To see this note that

$$\begin{aligned} \psi(A) > 0 &\Rightarrow \nu\{X_n^\infty \in A\} > 0 \text{ for some } n \\ &\Rightarrow \nu\{X_n^\infty \in A \text{ i.o.}\} = 1 \text{ by regeneration} \\ &\Rightarrow P_u\{X_n^\infty \in A \text{ i.o.}\} = 1 \Rightarrow \nu'\{X_n^\infty \in A \text{ i.o.}\} = 1 \\ &\Rightarrow \nu'(X_n^\infty \in A\} > 0 \text{ for some } n \Rightarrow \psi'(A) > 0. \end{aligned}$$

Thus $\psi \perp \psi'$, and by part (i) $\psi \sim \psi'$. Now suppose $B_\alpha \cap B_{\alpha'} \neq \phi$ and let $v \in B_\alpha$. Then $G(v, \{X^n_{n-m+1} = \alpha^0_{-m+1}$ i.o. (n), for all $m\}) = 1$ and since $\psi \sim \psi'$

$$G(v, \{X^n_{n-m+1} = \alpha'^0_{-m+1} \text{ i.o. } (n) \text{ for all } m\}) = 1,$$

i.e. $v \in B_{\alpha'}$, and $B_\alpha = B_{\alpha'}$.

Part (iii) Note that each class B_α is determined by a recurrent point α and an m_0, leading to a regeneration measure $\nu(\cdot)$. Namely we claim that if α and α' are recurrent points along which g loses memory and if

$$\alpha^0_{-m_0+1} = \alpha'^0_{-m_0+1} \quad \text{then } B_\alpha = B_{\alpha'}.$$

To see this, let E, $F_{\alpha'}$ be the events

$$\begin{aligned} E &= \{X_1^\infty \text{ sees the word } \alpha^0_{-m_0+1}\} \\ F_{\alpha'} &= \{X_1^\infty \text{ sees all the words of } \alpha' \text{ i.o.}\}. \end{aligned}$$

If $u \in B_\alpha$ then $G(u, E) = 1$. Choose $v \in B_{\alpha'}$ and note that $G(v, E) = 1$ and $G(v y \alpha^0_{-m_0+1}, F_{\alpha'}) = 1$ for every y such that $G(v, y\alpha^0_{-m_0+1} \mathcal{S}_1^\infty) > 0$. By (4.1) and (2.9) the measures $G(x^0_{-\infty} \alpha^0_{-m_0+1}, \cdot)$ and $G(y^0_{-\infty} \alpha^0_{-m_0+1}, \cdot)$ are mutually absolutely continuous for any $x^0_{-\infty}, y^0_{-\infty}$, and hence the measures $G(v y \alpha^0_{-m_0+1}, \cdot)$ and $G(w \alpha^0_{-m_0+1}, \cdot)$ are mutually absolutely continuous for any $w = w^0_{-\infty}$. Therefore $G(w \alpha^0_{-m_0+1}, F_{\alpha'}) = 1$ for any w. In particular $G(u z \alpha^0_{-m_0+1}, F_{\alpha'}) = 1$ for any z. This combined with $G(u, E) = 1$ implies $G(u, F_{\alpha'}) = 1$, i.e. $u \in B_{\alpha'}$. Since the number of finite words $\alpha^0_{-m_0+1}$ is countable, so are the B_α's. □

5 Examples

(i) Here is a class of ∞-memory chains where all points are recurrent and in the same attraction set, and the loss of memory rate depends on the point $\alpha=\alpha_{-\infty}^{0}$. The idea is to make $g(\cdot,\cdot)$ so that the memory depends on a part of the history going back a random time, depending on α.

Let $\mathscr{S}=(0,1,\ldots,d)$, $d<\infty$,

$$f_n:\mathscr{S}_{-n}^{0}\times\mathscr{S}\to\mathbb{R},\ \sum_{x_1\in\mathscr{S}} f_n(x_{-n}^{0},x_1)=1,$$

with

$$0<\varepsilon\leqq f_n(x_{-n}^{0},x_1)\qquad\text{for all } x_{-n}^{1}. \tag{5.1}$$

For each α let $L=L(\alpha)$ and $W=W(\alpha)$ be independent r.v.'s, with $P\{L=k\}=p_k(\alpha)=p_k$, $P\{W=k\}=w_k(\alpha)=w_k$.

Different α's will have different associated $L(\alpha)$'s and $W(\alpha)$'s, which will lead to processes with different memory decay. All these processes can be constructed on the same sample space, but there is no other particular relation between them. For definiteness consider for example $\alpha=\alpha_{-\infty}^{0}=(\ldots,0,0)$, i.e. the sequence of all 0's. Let $t_\ell(x_{-\infty}^{0})=\sup\{n: x_{-n}=\ldots=x_{-n+\ell-1}=0\}\leqq\infty$. ="the most recent time that a run of ℓ zeros was seen". Let $t_{k,\ell}(x_{-\infty}^{0})=k\wedge t_\ell(x_{-\infty}^{0})$. Define

$$\begin{aligned} g(x_{-\infty}^{0},x_1)&=\sum_{k,\ell} w_k\,p_\ell\, f_{t_{k,\ell}(x_{-\infty}^{0})}(x_{-t_{k,\ell}(x_{-\infty}^{0})}^{0},x_1)\\ &=Ef_{t_{W,L}(x_{-\infty}^{0})}(x_{-t_{W,L}(x_{-\infty}^{0})}^{0},x_1).\end{aligned}$$

Note that (5.1) implies

$$0<\varepsilon\leqq g(x_{-\infty}^{0},x_1)\quad\text{for all } x_{-\infty}^{1}. \tag{5.2}$$

Hence all points in $\mathscr{S}_{-\infty}^{0}$ are recurrence points and there is one ergodic class. Abbreviate $t_{k,\ell}(x_{-\infty}^{0})=t(k,\ell)$. Let $A_m=\{(k,\ell): t(k,\ell)\leqq m\}$ and write $g(x_{-\infty}^{0},x_1)=I_m(x_{-\infty}^{1})+II_m(x_{-\infty}^{1})$, where

$$I_m(x_{-\infty}^{1})=I_m=\sum_{A_m} w_k\,p_\ell\, f_{t(k,\ell)}(x_{-t(k,\ell)}^{0},x_1) \tag{5.3}$$

and

$$II_m(x_{-\infty}^{1})=II_m=\sum_{A_m^c} w_k\,p_\ell\, f_{t(k,\ell)}(x_{-t(k,\ell)}^{0},x_1). \tag{5.4}$$

Since $0<\varepsilon\leqq f_n(x_{-n}^{0},x_1)\leqq 1$ for all x_{-n}^{1}

$$\varepsilon\,\delta_m\leqq II_m(x_{-\infty}^{1})\leqq\delta_m\quad\text{for all } x_{-\infty}^{0}, \tag{5.5}$$

where

$$\delta_m=P\{t(W,L)>m\}.$$

Notice (by the definition (5.3)) that

$$I_m(x_{-\infty}^{1})=I_m(x_{-m}^{1}) \tag{5.6}$$

i.e. it depends only on x^1_{-m}. Hence

(5.7) $$g(x^0_{-\infty}, x_1) = I_m(x^1_{-m}) + II_m(x^1_{-\infty})$$

where $II_m(x^0_{-\infty})$ satisfies (5.5) uniformly in $x^0_{-\infty}$. Now if $x^0_{-m} = y^0_{-m}$, then

(5.8) $$\begin{aligned} g(y^0_{-\infty}, x_1) &= I_m(y^0_{-m}\, x_1) + II_m(y^0_{-\infty}\, x_1) \\ &= I_m(x^1_{-m}) + II_m(x^1_{-\infty}) + II_m(y^0_{-\infty}\, x_1) - II_m(x^1_{-\infty}) \\ &= g(x^0_{-\infty}, x_1) + r_m(x^1_{-\infty}, y^0_{-\infty}), \end{aligned}$$

where $r_m(x^1_{-\infty}, y^0_{-\infty}) = II_m(y^0_{-\infty}\, x_1) - II_m(x^1_{-\infty})$, and

(5.9) $$|r_m(x^1_{-\infty}, y^0_{-\infty})| \leqq \delta_m \quad \text{for all } x^1_{-\infty}, y^0_{-\infty}.$$

Thus

(5.10) $$\frac{g(y^0_{-\infty}, x^N_1)}{g(x^0_{-\infty}, x^N_1)} = \prod_{i=0}^{N-1} \frac{g(y^0_{-\infty}\, x^i_1, x_{i+1})}{g(x^i_{-\infty}, x_{i+1})},$$

and if $y^0_{-m} = x^0_{-m}$ then $y^0_{-m}\, x^i_1 = x^i_{-m}$ and hence by (5.8)

$$(5.10) = \prod_{i=0}^{N-1} \left(1 + \frac{r_{m+1}(x^{i+1}_{-\infty}, y^0_{-\infty}\, x^{i+1}_1)}{g(x^i_{-\infty}, x_{i+1})}\right).$$

Thus $f_n(x^0_{-\infty}, x_1) \geqq \varepsilon$ and (5.9) imply

(5.11) $$\begin{aligned} \prod_{i=0}^{\infty} \left(1 - \frac{\delta_{m+i}}{\varepsilon}\right) \leqq \prod_{i=0}^{N-1} \left(1 - \frac{\delta_{m+i}}{\varepsilon}\right) &\leqq \frac{g(y^0_{-\infty}, x^N_1)}{g(x^0_{-\infty}, x^N_1)} \\ &\leqq \prod_{i=0}^{N-1} \left(1 + \frac{\delta_{m+i}}{\varepsilon}\right) \leqq \prod_{i=0}^{\infty} \left(1 + \frac{\delta_{m+i}}{\varepsilon}\right). \end{aligned}$$

Now *assume* that

(5.12) $$\sum_k \delta_k < \infty.$$

Then for sufficiently large m, and some c

(5.13) $$1 - c \sum_{i=m}^{\infty} \delta_i \leqq \frac{g(y^0_{-\infty}, x^N_1)}{g(x^0_{-\infty}, x^N_1)} \leqq 1 + c \sum_{i=m}^{\infty} \delta_i$$

uniformly in $x^0_{-\infty}, y^0_{-\infty}, N$ subject to $x^0_{-m} = y^0_{-m}$. Hence

(5.14) $$\gamma_m(\alpha) \geqq 1 - c \sum_{i=m}^{\infty} \delta_i \nearrow 1.$$

Of course $\{\delta_i\}$ and hence γ_m depend on α, and hence so does the memory loss.

At the cost of some further restriction on the f_n's one can also claim an upper bound for γ_m, and hence a true rate. Suppose for example that there exist an $x^0_{-\infty}$, $y^0_{-\infty}$, $a \in \mathscr{S}$ and a $\beta > 0$ such that

$$f_n(x^0_{-n}, a) > f_n(y^0_{-n}, a) + \beta \quad \text{for } n = 0, 1, \ldots \tag{5.15}$$

Then by (5.4)

$$r_m(x^0_{-\infty}\, a, y^0_{-\infty}\, a) \leqq -\beta \delta_m$$

and

$$\begin{aligned} \frac{g(y^0_{-\infty}, a)}{g(x^0_{-\infty}, a)} &= 1 + \frac{r_m(x^0_{-\infty}\, a, y^0_{-\infty}\, a)}{g(x^0_{-\infty}\, a, y^0_{-\infty}\, a)} \\ &\leqq 1 - \beta \delta_m. \end{aligned} \tag{5.16}$$

Taking $a_i^N = (a, \ldots, a)$ we have

$$\frac{g(y^0_{-\infty}, a_1^N)}{g(x^0_{-\infty}, a_1^N)} \leqq 1 - \varepsilon' \sum_{i=m}^{\infty} \delta_i \quad \text{for some } \varepsilon' > 0. \tag{5.17}$$

Hence

$$\gamma_m(\alpha) \leqq 1 - \varepsilon' \sum_{i=1}^{\infty} \delta_i(\alpha) \tag{5.18}$$

and thus

$$c' \sum_{i=m}^{\infty} \delta_i(\alpha) \leqq 1 - \gamma_m(\alpha) \leqq c'' \sum_{i=m}^{\infty} \delta_i(\alpha) \tag{5.19}$$

for some $0 < c' \leqq c'' < \infty$. This identifies the rate of $\gamma_m(\alpha) \nearrow 1$.

(ii) Here is a (trivial) example where there exist two attraction sets, as defined in (2.5), and one transient class. Take $\mathscr{S} = \{0, 1, \ldots, 9\}$. For $x_0 = 0$ or 1, $x_{-1} = 0$ or 1 define g so that

$$g(x^0_{-\infty}, 0) + g(x^0_{-\infty}, 1) = 1.$$

For $x_0 = 8$ or 9, $x_{-1} = 8$ or 9, take $g(x^0_{-\infty}, 8) + g(x^0_{-\infty}, 9) = 1$. For all other cases take $g(x^0_{-\infty}, x_1) \geqq \varepsilon > 0$. Then

$$\begin{aligned} C_1 &= \{x^0_{-\infty} : x_0 \text{ and } x_{-1} = 0 \text{ or } 1\} \subset \mathscr{S}^0_{-\infty}, \text{ and} \\ C_2 &= \{x^0_{-\infty} : x_0 \text{ and } x_{-1} = 8 \text{ or } 9\} \subset \mathscr{S}^0_{-\infty} \end{aligned}$$

are closed classes, $\{x^0_{-\infty} : x_i = 0 \text{ or } 1\} = R_1$ are recurrent points in C_1, and $\{x^0_{-\infty} : x_i = 8 \text{ or } 9\} = R_2$ are recurrent points in C_2.

$C_3 = \mathscr{S}^0_{-\infty} \cap C_1^c \cap C_2^c$ is a transient class. All points in C_i are attracted to all points in R_i for $i = 1, 2$.

Acknowledgement. We thank the referee for many helpful comments and suggestions, particularly in the proof of Theorem 2.

References

[A-Ne] Athreya, K., Ney, P.: A new approach to the limit theory of recurrent Markov chains. Trans. Am. Math. Soc. **245**, 493–501 (1978)

[B] Berbee, H.: Chains with infinite connections: uniqueness and Markov representations. Probab. Theory Relat. Fields **76**, 243–253 (1987)

[B-K] Bramson, M., Kalikow, S.: Nonuniqueness in *g*-functions. University of Wisconsin Technical Report (1992)

[D] Doeblin, W.: Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables. Bull. Soc. Math. Fr. **66**, 210–220 (1938)

[D-F] Doeblin, W., Fortet, E.: Sur les chaines à liaisons complétes. Bull. Soc. Math. Fr. **65**, 132–148 (1937)

[I-G] Iosifescu, M., Grigorescu, S.: Dependence with complete connectins and applications. Cambridge: Cambridge University Press 1990

[K] Kalpazidou, S.: A Gaussian measure for certain continued fractions. Proc. Am. Math. Soc. **96**, 629–635 (1986)

[L] Lalley, S.P.: Regenerative representation for one dimensional Gibbs states. Ann. Probab. **14**, 1262–1271 (1986)

[N] Norman, M.F.: Markov processes and learning models. New York: Academic Press 1972

[N-Nu] Ney, P., Nummelin, E.: Regeneration for chains of infinite order and random maps. In: Proceedings of conference in honor of Doeblin. Blaubeuern 1991

[Nu1] Nummelin, E.: A splitting technique for Harris recurrent Markov chains. Z. Wahrscheinlichkeitstheor. Verw. Geb. **43**, 309–318 (1978)

[Nu2] Nummelin, E.: Renewal representations for Markov operators. Adv. Math. **90**, 15–46 (1991)

[O] Orey, S.: Lecture notes on limit theorems for Markov chain transition probabilities. London: Van Nostrand, Reinhold 1971

[R] Rosenblatt, M.: Markov processes. Structures and asymptotic behaviour. Berlin Heidelberg New York: Springer 1971