Resenhas IME-USP 2000, Vol. 4, No. 3, 249 - 281.

Regeneration in One-Dimensional Gibbs States and Chains with Complete Connections

S. P. Lalley

Abstract: A *chain with complete connections* is a stationary sequence of random variables, valued in a finite state space $\mathcal{Y}$, whose joint distribution is a *DLR state* (or *equilibrium state*, in the terminology of Ruelle [19]). It is shown that every chain with complete connections whose *interaction function* γ_m decays exponentially or polynomially admits a representation as a *block process*, that is, a stationary process obtained by stringing together i.i.d. random blocks (words) from the alphabet $\mathcal{Y}$. The length distribution of the random blocks in this representation has exponentially decaying tail if γ_m decays exponentially, and has finite $\beta - \varepsilon$ moment if $\gamma_m = O(m^{-\beta})$.

Key words: Gibbs state, chain with complete connections, regeneration point.

0. Foreword

This paper was written in 1984, and a much-abbreviated version was published in the 1986 *Annals of Probability* (vol. 14, no. 4, pp. 1262–1271). In the published version, only exponentially decaying interaction functions were considered. At the time, there was little interest in the case of polynomially decaying interactions (or so the editors of the *Annals of Probability* decreed), and so I abandoned the problem and moved on to other things. Recently, there has been renewed interest in regenerative representations. Therefore, at the suggestion of **Roberto Fernandez**, I have decided to publish the original manuscript in full.

1. Introduction

Perhaps the most important technical device in the study of recurrent Markov chains on denumerable state spaces is the decomposition of the sample path into i.i.d. "blocks" (in the terminology of Freedman [10]) by means of the process of successive returns to a distinguished state. This decomposition, apparently first used by Doeblin [6], [7], affords an easy approach to the ergodic theory of such chains, and also to the standard limit theorems for numerical functionals on the chains, such as the Central Limit Theorem and Law of the Iterated Logarithm (cf. Chung [4]) and the Renewal Theorem (cf. Smith [20]).

Recently Athreya and Ney [2] and Nummelin [17] discovered that a decomposition into i.i.d. blocks may be achieved for certain Markov chain without recurrent points, the so-called $(A, \lambda, \varphi, 1)$ - - recurrent chains. In this decomposition

---

Columbia University and now University of Chicago

the recurrent set A serves as a surrogate for the distinguished recurrent point of Doeblin's decomposition; an auxiliary randomization is used to determine which visits to A are labeled "regeneration" times. Athreya, Mac Donald, and Ney [1] have also shown that the block decomposition may be used as the basis for an easy proof of a general renewal theorem for Markov chains first established by Kesten [13].

The purpose of this paper is to exhibit a block decomposition for a class of stationary processes, the so-called "chains with complete connections". These processes were introduced by Onicescu and Mihoc [18]; the basic ergodic theory was developed by Doeblin and Fortet [8] and Harris [11]. Iosifescu and Theodorescu [12] and others have used chains with complete connections and various related processes as models for learning behavior. Recently Ledrappier [15] noted that the class of chains with complete connections includes the one-dimensional Gibbs states (also known as DLR states), which have been studied extensively in the literature of statistical mechanics and topological dynamics (cf. Ruelle [19] and Bowen [3] for surveys).

The decomposition described here is similar to that of Athreya and Ney in one respect, to wit, it relies on an auxiliary randomization to select the regeneration times. In others respects it is quite different: without the Markov property it becomes necessary to deal with conditional probabilities involving the entire past. The construction is explicit enough to give a relation between the "strength" of the dependence (i.e., the rate of decay of the interaction) and the number of finite moments allowed the regeneration time.

2. Statement of Principal Results

A *chain with complete connections* is a stationary process $\{Y_n\}_{n\in\mathbb{Z}}$ taking values in a finite state space $\mathcal{Y}$ such that

$$P(Y_1 = \xi_1, \ldots, Y_n = \xi_n) > 0, \quad \forall \xi_1, \ldots, \xi_n \in \mathcal{Y}; \tag{2.1}$$

$$\begin{aligned} \lim_{m\to\infty} & P(Y_0 = \xi_0 \mid Y_n = \xi_n, -m \le n \le -1) \\ & = P(Y_0 = \xi_0 \mid Y_n = \xi_n,\ n \le -1) \text{ exists for all} \\ & \xi_0, \xi_{-1}, \xi_{-2}, \ldots \in \mathcal{Y}; \end{aligned} \tag{2.2}$$

and

$$\gamma_m \downarrow 0, \tag{2.3}$$

where γ_m is the *interaction function*, defined by

$$\gamma_m \triangleq \sup\Big\{ \Big| \frac{P(Y_n = \xi_n,\ 0 \le n \le r \mid Y_n = \xi_n,\ n \le -1)}{P(Y_n = \xi_n, 0 \le n \le r \mid Y_n = \xi_n^*,\ n \le -1)} - -1 \Big| : \\ r < \infty;\ \xi_j,\ \xi_j^* \in \mathcal{Y};\quad \text{and}\quad \xi_n = \xi_n^*,\ \ \forall n,\ -m \le n \le -1 \Big\}. \tag{2.4}$$

This definition is somewhat different than that of Doeblin and Fortet [8], but more suitable for our purposes. Observe that k-step Markov dependence is equivalent to $\gamma_m = 0,\ \forall m \ge k$. Notice also that the conditional probabilities in (2.4) are defined for all sequences ξ_n, ξ_n^* from $\mathcal{Y}$, by (2.2) and the stationarity of $\{Y_n\}_{n\in\mathbb{Z}}$.

To state our main results concerning chains with complete connections we must clarify the notion of a regenerative representation. Roughly, a *block process* is a stationary process obtained by piecing together i.i.d. blocks of symbols from $\mathcal{Y}$ (not necessarily of the same length). A *block* b is an element of the set $\mathcal{Y}^n$ of finite sequences (words) with entries in the set $\mathcal{Y}$; its length is denoted by $\lambda(b)$. Suppose that on some probability space are defined independent random blocks $\{B_n\}_{n\in\mathbb{Z}}$ and an integer-valued random variable M such that

$$\begin{aligned} &\{B_n\}_{n\le -1, n\ge 1} \quad \text{are i.i.d. };\\ &E\lambda(B_1) < \infty;\\ &P(B_0 = b) = \lambda(b)\ P(B_1 = b)/E\lambda(B_1);\quad \text{and}\\ &P(M = k \mid \{B_n\}_{n\in\mathbb{Z}}) = 1/\lambda(B_0),\ k = 1, 2, \ldots, \lambda(B_0). \end{aligned} \tag{2.5}$$

Let

$$\begin{aligned} N(n) &\triangleq \min\{m \ge 0 : \sum_{i=0}^{m} \lambda(B_i) - M \ge n\},\ n \ge 0;\\ &\triangleq \min\{m \ge 0 : M - 1 + \sum_{i=1}^{m} \lambda(B_{-i}) \ge -n\},\ n < 0; \end{aligned} \tag{2.6}$$

define

$$\begin{aligned} Y_n^* &\triangleq B_{N(n)}(\lambda(B_{N(n)}) - (\sum_{i=0}^{N(n)} \lambda(B_i) - M - n)),\ n \ge 0;\\ &\triangleq B_{-N(n)}(M + \sum_{i=1}^{N(n)} \lambda(B_{-i}) + n),\ n < 0. \end{aligned} \tag{2.7}$$

Then the process $\{Y_n^*\}_{n\in\mathbb{Z}}$ is called a block process. Every block process is stationary: the proof, which is an easy exercise in elementary renewal theory, is left to the reader. It should be noted that block processes are derived from a special kind of semi-Markov process (cf. Smith [20]), in which the blocks are the successive states, and the block lengths $\lambda(B_n)$ are the sojourn times. In section 3 we shall prove the following two theorems.

THEOREM 1: *Suppose* $\{Y_n\}_{n\in\mathbb{Z}}$ *is a chain with complete connections for which the sequence* $\{\gamma_m\}_{m\geq 0}$ *is exponentially decaying. Then there is a version* $\{Y_n^*\}_{n\in\mathbb{Z}}$ *of* $\{Y_n\}_{n\in\mathbb{Z}}$ *which is a block process, and such that the block length variable* $\lambda(B_1)$ *has an exponentially decaying tail, i.e.,*

$$Ee^{\theta\lambda(B_1)} < \infty, \quad \text{for some } \theta > 0. \tag{2.8}$$

NOTE: $\{\gamma_m\}$ exponentially decaying means that there are constants $C < \infty$, $0 < \beta < 1$ such that $\gamma_m \leq C\beta^m$, for all m.

THEOREM 2: *Suppose* $\{Y_n\}_{n\in\mathbb{Z}}$ *is a chain with complete connections for which*

$$\gamma_m = O(m^{-\beta}) \tag{2.9}$$

for some constant $\beta > 1$. *Then for every* $p < \beta$ *there exists a version* $\{Y_n^*\}_{n\in\mathbb{Z}}$ *of* $\{Y_n\}_{n\in\mathbb{Z}}$ *which is a block process, and for which the block length variable* $\lambda(B_1)$ *satisfies*

$$E\lambda(B_1)^p < \infty. \tag{2.10}$$

It is not at all obvious that the hypotheses concerning the decay of γ_m are the "right" ones. To demonstrate that polynomial decay of γ_m is the appropriate hypothesis for (2.10), we shall prove (in section 4) the following.

PROPOSITION 1: *For each* $p > 1$ *there is a chain with complete connections* $\{Y_n\}_{n\in\mathbb{Z}}$ *with* $\gamma_m = O(m^{-p+1})$ *such that if* $\{Y_n^*\}_{n\in\mathbb{Z}}$ *is a block process version of* $\{Y_n\}_{n\in\mathbb{Z}}$, *then the block length variable* $\lambda(B_1)$ *must satisfy*

$$E\lambda(B_1)^{p+1} = \infty. \tag{2.11}$$

What is the minimal rate of decay of γ_m sufficient to guarantee the existence of a regenerative representation with (2.10)? We have not succeeded in determining this; however, Theorem 2 and Proposition 1 show that the answer is somewhere between $O(m^{-p})$ and $O(m^{-(p-2)})$.

The hypotheses concerning the rate of decay of γ_m are by no means *necessary* for the conclusions (2.8) and (2.10). It is quite possible for (2.8) to hold in a regenerative representation of a chain with complete connections for which $\gamma_m \downarrow 0$ very slowly. The easiest examples are stationary renewal processes. Let $0 < \beta < 1$

and $\{a_m\}_{m\geq 1}$ be constants such that $a_m \to 1$ as $m \to \infty$, and $a_m > \beta a_{m+1}$ for all m, and $1 > \beta a_1$. Define a distribution on blocks of zeros and ones as follows:

$$P(B_1 = (1)) = 1 - \beta a_1;$$
$$P(B_1 = (0,1)) = \beta a_1 - \beta^2 a_2;$$
$$P(B_1 = (0,0,1)) = \beta^2 a_2 - \beta^3 a_3;$$
$$\dots .$$

Then $\lambda(B_1)$ clearly satisfies $Ee^{\theta\lambda(B_1)} < \infty$ for $e^\theta < 1/\beta$; moreover for the block process constructed from this distribution

$$\gamma_m \geq \sup_{n>m} \left| \frac{1-\beta(a_{n+1}/a_n)}{1-\beta(a_{m+1}/a_m)} - 1 \right|$$

which may converge to zero quite slowly, depending on the rate at which $a_m \to 1$.

An important reason for studying chains with complete connections is the fact that they arise from the so-called *Gibbs states* of ergodic theory (cf. Dobrushin [5] and Lanford and Ruelle [14]), which we shall now describe. Let H (the "Hamiltonian", or "energy-per-site" function) be a continuous function on the "configuration space" $\mathcal{Y}^{\mathbb{Z}}$, i.e., suppose that

$$\delta_m \downarrow 0, \tag{2.12}$$

where

$$\delta_m \triangleq \sup\{\, | H(\xi) - H(\xi^*) | \, : \xi, \xi^* \in \mathcal{Y}^{\mathbb{Z}} \text{ and } \xi_n = \xi_n^*, \;\; 0 \leq |n| \leq m\}. \tag{2.13}$$

For each finite interval $\Lambda \subset \mathbb{Z}$ and each configuration $\zeta \in \mathcal{Y}^{\mathbb{Z}\setminus\Lambda}$ on $\mathbb{Z}\setminus\Lambda$, define a probability measure $\mu_{\Lambda|\zeta}$ on $\mathcal{Y}^\Lambda$ (the "Gibbs ensemble" for the boundary condition ζ), by

$$\mu_{\Lambda|\zeta}(\{\xi\}) = \exp\{-\sum_{n\in\Lambda} H(\sigma^n(\xi \vee \zeta))\}/Z(\Lambda;\zeta), \;\; \xi \in \mathcal{Y}^\Lambda, \tag{2.14}$$

where

$$Z(\Lambda;\zeta) = \sum_{\xi^*\in\mathcal{Y}^\Lambda} \exp\{-\sum_{n\in\Lambda} H(\sigma^n(\xi^* \vee \zeta))\}. \tag{2.15}$$

Here $\xi \vee \zeta \in \mathcal{Y}^{\mathbb{Z}}$ denotes the configuration on $\mathbb{Z}$ obtained by amalgamating ξ and ζ, and σ denotes the shift operator on $\mathcal{Y}^{\mathbb{Z}}$, i.e., $\forall n, j \in \mathbb{Z}, \xi \in \mathcal{Y}^{\mathbb{Z}}$

$$(\sigma^n \xi)_j = \xi_{n+j}. \tag{2.16}$$

A probability measure P on $\mathcal{Y}^{\mathbb{Z}}$ is called a *Gibbs state* (or a *DLR state*) if for every finite subset $\Lambda \subset \mathbb{Z}$, every $\xi \in \mathcal{Y}^\Lambda$, and every $\zeta \in \mathcal{Y}^{\mathbb{Z}\setminus\Lambda}$

$$P(Y_\Lambda = \xi \mid Y_{\mathbb{Z}\setminus\Lambda} = \zeta) = \mu_{\Lambda|\zeta}(\{\xi\}), \tag{2.17}$$

where Y_Λ and $Y_{\mathbb{Z}\setminus\Lambda}$ denote the coordinate projections $\eta \to \eta \mid \Lambda$ and $\eta \to \eta \mid \mathbb{Z}\setminus\Lambda$ for $\eta \in \mathcal{Y}^{\mathbb{Z}}$. In other words, a probability measure on $\mathcal{Y}^{\mathbb{Z}}$ is a Gibbs state if the Gibbs ensembles $\mu_{\Lambda|\zeta}$ form a system of regular conditional distributions for the configurations on Λ. If P is a Gibbs state, then the stochastic process $\{Y_n\}_{n\in\mathbb{Z}}$ consisting of the coordinate random variables will be called a Gibbs process.

It is known that for a Hamiltonian function H satisfying $\sum_{m=1}^{\infty} \delta_m < \infty$, there is a unique Gibbs state $P = P_H$. This measure is translation-invariant, so the resultant Gibbs process is stationary. (This result is, in essence, proved in Ruelle [19], Ch. 5: cf. Corollary 5.6. It is not difficult to give a direct proof: P is just the weak limit on $\mathcal{Y}^{\mathbb{Z}}$ of the Gibbs ensembles $\mu_{\Lambda_n \mid \zeta_n}$, where $\Lambda_n = \{-n, \ldots, n-1, n\}$ and ζ_n is *any* configuration on $\mathbb{Z}\setminus\Lambda_n$).

PROPOSITION 2: *Suppose the Hamiltonian function H satisfies $\sum_{m=1}^{\infty} \delta_m < \infty$. Then the Gibbs process $\{Y_n\}_{n\in\mathbb{Z}}$ with distribution P_H is a chain with complete connections. Moreover, for γ_m defined by (2.4),*

$$\gamma_m = O\Big(\sum_{j=m}^{\infty} \delta_j\Big). \tag{2.18}$$

The proof will be given in section 5.

It is relation (2.18) that makes Theorems 1 and 2 worthwhile. For (2.18) allows one to relate the number of finite moments permitted the regeneration time in a block process version directly to the continuity properties of the Hamiltonian function H (or, equivalently, to the decay of the corresponding "interaction function" : cf. Ruelle [19]). In particular, if H is Hölder continuous (i.e., δ_m decays exponentially) then there is a regenerative representation for which the block-length variable has finite exponential moments. We remark that all of the Gibbs states considered by Bowen in [3] have Hölder continuous Hamiltonian functions.

3. Regeneration in Chains with Complete Connections

The problem of giving a block decomposition of a stationary process is essentially the same as that of constructing a single regeneration point. This is formalized in

LEMMA 3.1: *Suppose that on some probability space are defined a $\mathcal{Y}$-valued stationary process $\{Y_n\}_{n\geq 0}$ and a random variable $T \in \mathbb{Z}^+$ with $ET < \infty$, such that for all $\xi_j, \zeta_j \in \mathcal{Y}$ and $k, m \in \mathbb{Z}^+$,*

$$\begin{aligned} P(Y_{m+n} = \xi_n,\ 0 \le n \le k \mid T = m; Y_j = \zeta_j,\ j = 0, 1, \ldots, m-1) \\ = P(Y_n = \xi_n,\ 0 \le n \le k). \end{aligned} \tag{3.1}$$

Then there is a block process $\{Y_n^*\}_{n \in \mathbb{Z}}$ such that $\{Y_n^*\}_{n \ge 0} \overset{\mathcal{D}}{=} \{Y_n\}_{n \ge 0}$. Moreover, the block process may be chosen so that the block length variable $\lambda(B_1)$ is identical in law to T.

PROOF: Let $\{B_n\}_{n \in \mathbb{Z}}$ be a sequence of i.i.d. random blocks, and M a random variable satisfying (2.10), such that the distribution of the block variables B_n satisfies

$$P\{B_1 = (\xi_1, \ldots, \xi_m)\} = P\{T = m \text{and} Y_n = \xi_{n+1}, \forall\ 0 \le n < m\}. \tag{3.2}$$

We must show that the block process $\{Y_n^*\}$ obtained from $\{B_n\}$ and M according to (2.12) has the same law as $\{Y_n\}$. We may assume WLOG that the random variables $\{Y_n\}_{n \ge 0}$, $\{B_n\}_{n \in \mathbb{Z}}$, T, and M are all defined on the same probability space, with $(\{B_n\}_{n \in \mathbb{Z}}, M)$ independent of $(\{Y_n\}_{n \ge 0}, T)$.

For each $k = 1, 2, \ldots$, let $\{\overline{Y}_n^k\}_{n \ge 0}$ be the sequence of $\mathcal{Y}$- valued random variables obtained by prefacing the sequence $\{Y_n\}_{n \ge 0}$ by the first k blocks B_1, B_2, $\ldots$, B_k : specifically, let

$$(\overline{Y}_0^k, \overline{Y}_1^k, \ldots, \overline{Y}_{\lambda(B_1)-1}^k) = (B_1(1), \ldots, B_1(\lambda(B_1)));$$

$$(\overline{Y}_{\lambda(B_1)}^k, \ldots, \overline{Y}_{\lambda(B_1)+\lambda(B_2)-1}^k) = (B_2(1), \ldots, B_2(\lambda(B_2)));$$

$$\ldots$$

$$(\overline{Y}^k_{\sum_{j=1}^k \lambda(B_j)}, \ldots, \overline{Y}^k_{\sum_{j=1}^k \lambda(B_j)-1}) = (B_k(1), \ldots, B_k(\lambda(B_k)));$$

$$(\overline{Y}^k_{\sum_{j=1}^k \lambda(B_j)}, \overline{Y}^k_{\sum_{j=1}^k \lambda(B_j)+1}, \ldots) = (Y_0, Y_1, \ldots).$$

Then for each $k \ge 1$, the sequence $\{\overline{Y}_n^k\}_{n \ge 0}$ has the same law as $\{Y_n\}_{n \ge 0}$. This follows by an easy induction argument (on k). That it is true for $k = 1$ follows immediately from (3.1), (3.2), and the postulated independence of B_1 and $\{Y_n\}_{n \ge 0}$; that it is true for $k = K + 1$ follows by combining the results for $k = K$ and $k = 1$, since $(B_2, \ldots, B_{K+1})$ has the same law as $(B_1, \ldots, B_K)$.

Next notice that for each $n \ge 0, \overline{Y}_n^k \to \overline{Y}_n^\infty$ as $k \to \infty$; $\{\overline{Y}_n^\infty\}_{n \ge 0}$ is the process obtained by stringing together the entire sequence $\{B_k\}_{k \ge 1}$ of blocks.

Since each of the sequences $\{\overline{Y}^k_n\}_{n\geq 0}$, $k \geq 1$, has the same distribution as $\{Y_n\}_{n\geq 0}$, it follows that $\{\overline{Y}^{\infty}_n\}_{n\geq 0}$ is also equal to $\{Y_n\}_{n\geq 0}$ in law.

Now let $\{Y^*_n\}_{n\in\mathbb{Z}}$ be the block process obtained from $\{B_k\}_{k\in\mathbb{Z}}$ and M according to the specification (2.12). Since $\{\overline{Y}^{\infty}_n\}_{n\geq 0} \stackrel{\mathcal{D}}{=} \{\overline{Y}^{\infty}_{n+m}\}_{n\geq 0} \stackrel{\mathcal{D}}{=} \{Y_n\}_{n\geq 0}$ for all $m \geq 0$, it is now enough to show that $\{\overline{Y}^{\infty}_{n+m}\}_{n\geq 0} \stackrel{\mathcal{D}}{\to} \{Y^*_n\}_{n\geq 0}$ as $m \to \infty$ (here convergence in distribution refers to the product topology on $\mathcal{Y}^{\mathbb{N}}$). But this is an easy consequence of the Feller-Erdös-Pollard renewal theorem (cf. [19], Ch. 13): in particular, since $\{B_k\}_{k\geq 1}$ are i.i.d., if $\tau(m) \stackrel{\Delta}{=} \min\{k : \sum_{j=1}^{k} \lambda(B_j) \geq m+1\}$, then

$$\lim_{m\to\infty} P\{\sum_{j=1}^{\tau(m)} \lambda(B_j) - m - 1 = n; B_{\tau(m)} = b_0, \ldots; B_{\tau(m)+r} = b_r\}$$

$$= P\{M = \lambda(b_0) - n; B_0 = b_0; B_1 = b_1; \ldots; B_r = b_r\}.$$

Hence $\{\overline{Y}^{\infty}_{m+n}\}_{n\geq 0} \stackrel{\mathcal{D}}{\to} \{Y^*_n\}_{n\geq 0}$ as $m \to \infty$.

\\\

Notational Convention. If $(\Omega, \mathcal{F}, Q)$ is a probability space on which are defined random variables $\{Y_n\}_{n\in\mathbb{Z}}$ valued in $\mathcal{Y}$, then for all subsets $A, B \subset \mathbb{Z}$ and each sequence of values $\xi_n \in \mathcal{Y}$, $Q(\xi(A))$ will denote $Q(Y_n = \xi_n, \forall\, n \in A)$ and $Q(\xi(A) \mid \xi(B))$ will denote $Q(Y_n = \xi_n\ \forall n \in A \mid Y_n = \xi_n\ \forall n \in B)$. In addition, the interval notations [,], (,], etc., will be used to denote intervals of *integers*, e.g. $[m, n] = \{m, m+1, \ldots, n\}, [m, n) = \{m, m+1, \ldots, n-1\}$.

LEMMA 3.2: *Suppose $\{Y_n\}_{n\in\mathbb{Z}}$ is a chain with complete connections. Then there exists $\delta_0 > 0$ such that for every $A \subset (-\infty, -1]$ and finite $B \subset [0, \infty)$, and all values $\xi_n \in \mathcal{Y}$,*

$$P(\xi(B) \mid \xi(A)) \geq \delta_0 P(\xi(B)). \tag{3.3}$$

PROOF. It clearly suffices to consider only cases $A = (-\infty, -1]$ and $B = [0, m]$. According to (2.3) there exist $\delta_1 > 0$ and an integer $1 \leq r < \infty$ such that for all choices of $\xi_n \in \mathcal{Y}$ and all $m < \infty$

$$P(\xi[0, m] \mid \xi(-\infty, -1]) \geq \delta_1 P(\xi[0, m] \mid \xi[-r, --1]), \tag{3.4}$$

and consequently, for all $k \geq r$,

$$P(\xi[0, m] \mid \xi[-k, -1]) \geq \delta_1 P(\xi[0, m] \mid \xi[-r, --1]). \tag{3.5}$$

Now since there are only finitely many configurations (choices of ξ_n) on $[-r, r-1]$ it follows from (2.1) that there exists $\delta_2 > 0$ such that for all $\xi_n \in \mathcal{Y}$

$$P(\xi[0, r-1] \mid \xi[-r, -1]) \geq \delta_2 P(\xi[0, r-1]). \tag{3.6}$$

Therefore for all $\xi_n \in \mathcal{Y}$ and $m > r-1$,

$$\begin{aligned} P(\xi[0, m] \mid \xi[-r, -1]) &= P(\xi[r, m] \mid \xi[-r, r-1]) \cdot P(\xi[0, r-1] \mid \xi[-r, -1]) \\ &\geq \delta_1 P(\xi[r, m] \mid \xi[0, r-1]) \delta_2 P(\xi[0, r-1]) \\ &= \delta_1 \delta_2 P(\xi[0, m]). \end{aligned} \tag{3.7}$$

(Here we have used (3.5) and (3.6), together with the stationarity of $\{Y_n\}$,which guarantees that the conditional probabilities are translation invariant). Combining (3.4) and (3.7), we conclude that for all $\xi_n \in \mathcal{Y}$ and all $m < \infty$

$$P(\xi[0, m] \mid \xi(-\infty, -1]) \geq \delta_1^2 \delta_2 P(\xi[0, m]).$$

\\\

PROOF OF THEOREM 1.

First we show that there is no loss of generality in assuming that

$$\gamma_k = \gamma_k^Y \leq 16^{-k} \tag{3.8}$$

for all $k \geq 1$. For suppose $\{Y_n\}_{n \in \mathbb{Z}}$ is a chain with complete connections on the state space $\mathcal{Y}$, such that $\gamma_k^Y \leq C\theta^k$ for some $C < \infty$ and $0 < \theta < 1$. Then the process

$$Z_n \triangleq (Y_{nr+1}, Y_{nr+2}, \ldots, Y_{(n+1)r}), \; n \in \mathbb{Z},$$

is a chain with complete connections on the state space $\mathcal{Y}^r$ such that $\gamma_k^Z = \gamma_{kr}^Y \leq C\theta^{rk}$. By choosing r sufficiently large, one obtains $\gamma_k^Z \leq (16)^{-k}$ for all $k \geq 1$. Now clearly if there is a block process with the same law as $\{Z_n\}$, then there is a block process with the same law as $\{Y_n\}_{n \in \mathbb{Z}}$, for which the block length variable is multiplied by a factor of r.

Next we define a (countable) family of probability distributions Q_n on the sequence space $\mathcal{Y}^{[0,\infty)}$. Let Q_0 be the probability measure defined by

$$Q_0(\xi[0, m]) = P(\xi[0, m]) \tag{3.9}$$

for all $m \geq 0, \xi_0, \xi_1, \ldots, \xi_m \in \mathcal{Y}$. Fix $\delta > 0$ small. For each $k \geq 0$ and each choice of $\xi_0, \xi_1, \ldots, \xi_k \in \mathcal{Y}$, let $Q_{k+1}^{\xi[0,k]}$ be the probability distribution on $\mathcal{Y}^{[k+1,\infty)}$ specified by

$$\begin{aligned} Q_{k+1}^{\xi[0,k]}&(\xi[k+1,k+m]) \\ &= (1-\delta)^{-1}\left[Q_k^{\xi[0,k-1]}(\xi[k+1,k+m] \mid \xi_k) - \delta P(\xi[k+1,k+m])\right] \end{aligned} \tag{3.10}$$

for all $\xi_{k+1},\ldots,\xi_{k+m} \in \mathcal{Y}$, $m \geq 1$. In order that this be a valid recursive definition it must be shown (inductively) that for $\delta > 0$ sufficiently small,

$$Q_k^{\xi[0,k-1]}(\cdot \mid \xi_k) \geq \delta P(\cdot). \tag{3.11}$$

Before proving (3.11) we will show how to use the probability distributions $Q_k^{\xi[0,k-1]}$ to build a regeneration point for $\{Y_n\}$. Let $(\Omega, \mathcal{F}, Q)$ be a probability space on which are defined random variables $\{Y_n^A\}_{n\geq 0}$, $\{Y_n^B\}_{n\geq 0}$ (all valued in $\mathcal{Y}$) and T (valued in $\{1,2,\ldots\}$) such that

$$\begin{aligned} Q(Y_n^A =&\xi_n,\ n \in \Lambda_1; Y_n^B = \zeta_n,\ n \in \Lambda_2;\ T = k) \\ &= Q(Y_n^A = \xi_n,\ n \in \Lambda_1) Q(Y_n^B = \zeta_n,\ n \in \Lambda_2)\ Q(T = k) \end{aligned} \tag{3.12}$$

for all $\xi_n, \zeta_n \in \mathcal{Y}$ and $k \in \mathbb{Z}^+$, and all finite subsets $\Lambda_1, \Lambda_2 \subset \mathbb{Z}$; and

$$Q(T = k) = \delta(1-\delta)^{k-1},\ k = 1,2,\ldots;$$

$$\begin{aligned} &Q(Y_n^B = \xi_n,\ n \in \Lambda) = P(\xi(\Lambda)); \\ &Q(Y_n^A = \xi_n,\ 0 \leq n \leq k) = Q_0(\xi_0) \prod_{j=1}^{k} Q_j^{\xi[0,j-1]}(\xi_j) \end{aligned} \tag{3.13}$$

for all $\xi_n \in \mathcal{Y}$, $\Lambda \subset \mathbb{Z}$, and $k \geq 1$. Define new random variables $\{Y_n^*\}_{n\geq 0}$ on $(\Omega, \mathcal{F}, P)$ by

$$Y_n^* = \begin{cases} Y_n^A, & n < T, \\ Y_n^B, & n \geq T. \end{cases} \tag{3.14}$$

We will argue that $\{Y_n^*\}_{n\geq 0}$ has the same distribution as the original process $\{Y_n\}_{n\geq 0}$. It is clear from the construction that for all $\xi_j,\ \zeta_j \in \mathcal{Y}, k, m \in \mathbb{Z}^+$

$$\begin{aligned} Q(Y_{m+n}^* = \xi_n, 0 \leq n \leq k \mid T = m; Y_j = \zeta_j, 0 \leq j \leq m) \\ = Q(Y_{m+n}^B = \xi_n, 0 \leq n \leq k) \\ = P(Y_n = \xi_n, 0 \leq n \leq k) \end{aligned}$$

so by Lemma 3.1, proving that $\{Y_n\}_{n\geq 0} \stackrel{\mathcal{D}}{=} \{Y_n^*\}_{n\geq 0}$ will suffice to show that $\{Y_n\}_{n\geq 0}$ has a representation as a block process.

To see that $\{Y_n\}_{n\geq 0} \stackrel{\mathcal{D}}{=} \{Y_n^*\}_{n\geq 0}$, use (3.12), (3.13), and (3.14) to write the finite-dimensional distributions of $\{Y_n^*\}_{n\geq 0}$ as

$$
\begin{aligned}
&Q(Y_n^* = \xi_n,\ 0 \le n \le k) \\
&= \sum_{m=1}^{k} Q(T = m) Q(Y_n^A = \xi_n, 0 \le n < m) Q(Y_n^B = \xi_n, m \le n \le k) \\
&+ Q(T > k) Q(Y_n^A = \xi_n, 0 \le n \le k) \\
&= \sum_{m=1}^{k} \delta(1-\delta)^{m-1} [\prod_{j=0}^{m-1} Q_j^{\xi[0,j-1]}(\xi_j)] P(\xi[m,k]) \\
&+ (1-\delta)^k \prod_{j=0}^{k} Q_j^{\xi[0,j-1]}(\xi_j).
\end{aligned} \tag{3.15}
$$

Now use the relation (3.10) successively for $Q_k^{\xi[0,k-1]}$, then $Q_{k-1}^{\xi[0,k-2]}$, etc. , to get

$$
\begin{aligned}
\text{RHS } (3.15) &= \sum_{m=1}^{k-1} \delta(1-\delta)^{m-1} [\prod_{j=0}^{m-1} Q_j^{\xi[0,j-1]}(\xi_j)] P(\xi[m,k]) \\
&+ (1-\delta)^{k-1} [\prod_{j=0}^{k-1} Q_j^{\xi[0,j-1]}(\xi_j)] Q_{k-1}^{\xi[0,k-2]}(\xi_k \mid \xi_{k-1}) \\
&= \sum_{m=1}^{k-2} \delta(1-\delta)^{m-1} [\prod_{j=0}^{m-1} Q_j^{\xi[0,j-1]}(\xi_j)] P(\xi[m,k]) \\
&+ (1-\delta)^{k-2} [\prod_{j=0}^{k-2} Q_j^{\xi[0,j-1]}(\xi_j)]\ Q_{k-2}^{\xi[0,k-3]}(\xi[k-1,k] \mid \xi_{k-2}) \\
&= \dots \\
&= Q_0(\xi_0) Q_0(\xi[1,k] \mid \xi_0) \\
&= Q_0(\xi[0,k]) \\
&= P(\xi[0,k]).
\end{aligned}
$$

It remains to be shown that $\delta > 0$ can be chosen sufficiently small that (3.11) holds for all k, so that the recursive definition of the probability measures $Q_k^{\xi[0,k-1]}$ is valid. Recall from Lemma 3.2 that there exists $\delta_0 > 0$ so small that

$$
P(\xi(B) \mid \xi(A)) \ge \delta_0 P(\xi(B)) \tag{3.16}
$$

for all $A \subset (-\infty, k-1]$ and finite $B \subset [k, \infty)$, and all choices of $\xi_n \in \mathcal{Y}$. Choose

$$
\delta < \delta_0/4 \le 1/4. \tag{3.17}
$$

We will show by induction on k that for all $k \geq 0, m \geq 1$, and $r \geq 0$, and all choices of $\xi_n \in \mathcal{Y}$,

$$\left| \frac{Q_k^{\xi[0,k-1]}(\xi[k+m,k+m+r] \mid \xi[k,k+m-1])}{P(\xi[k+m,k+m+r] \mid \xi[k,k+m-1])} - 1 \right| \leq (4)(16)^{-m}. \tag{3.18}$$

Since RHS (3.18) $\leq 1/2$, it follows immediately from (3.16), (3.17), and (3.18) that

$$Q_k^{\xi[0,k-1]}(\xi[k+1,k+n] \mid \xi_k) \geq \frac{1}{2}\delta_0 P(\xi[k+1,k+n]) \tag{3.19}$$

for all $k \geq 0, n \geq 1$, and all choices of $\xi_j \in \mathcal{Y}$. This clearly will prove (3.11).

Notice that (3.18), and (3.19), are trivial for $k = 0$, by (3.9). We now assume that (3.18) and (3.19) are true for some indeterminate value of k, and proceed to show that (3.18), and hence (3.19), must also hold for $k+1$.

Write

$$\begin{aligned} &Q_{k+1}^{\xi[0,k]}(\xi[k+1+m,k+1+m+r] \mid \xi[k+1,k+m]) \\ &= \frac{Q_{k+1}^{\xi[0,k]}(\xi[k+1,k+1+m+r])}{Q_{k+1}^{\xi[0,k]}(\xi[k+1,k+m])}. \end{aligned} \tag{3.20}$$

Now apply (3.10) to both numerator and denominator of RHS (3.20), then divide by $P(\xi[k+1+m,k+1+m+r] \mid \xi[k+1,k+m])$ to obtain

$$\begin{aligned} &\left| \frac{Q_{k+1}^{\xi[0,k]}(\xi[k+1+m,k+1+m+r] \mid \xi[k+1.k+m])}{P(\xi[k+1+m,k+1+m+r] \mid \xi[k+1,k+m])} - 1 \right| \\ &= \left| \frac{Q_k^{\xi[0,k-1]}(\xi[k+1+m,k+1+m+r] \mid \xi[k,k+m])}{P(\xi[k+1+m,k+1+m+r] \mid \xi[k+1.k+m])} - 1 \right| \\ &\quad \cdot \left| 1 - \delta \frac{P(\xi[k+1,k+m])}{Q_k^{\xi[0,k-1]}(\xi[k+1,k+m] \mid \xi_k)} \right|^{-1} \\ &\leq (1 - 2\delta_0^{-1}\delta)^{-1} \left| \frac{Q_k^{\xi[0,k-1]}(\xi[k+1+m,k+1+m+r] \mid \xi[k,k+m])}{P(\xi[k+1+m,k+1+m+r] \mid \xi[k+1,k+m])} - 1 \right|. \end{aligned} \tag{3.21}$$

The last inequality follows from (3.19), which holds by virtue of the induction hypothesis. Now by (3.17),

$$(1 - 2\delta_0^{-1}\delta)^{-1} \leq 2,$$

so to complete the proof it suffices to show that the $|\cdot|$ factor on RHS (3.21) is no larger than $2(16)^{-m}$. But

$$\frac{Q_k^{\xi[0,k-1]}(\xi[k+1+m,k+1+m+r] \mid \xi[k,k+m])}{P(\xi[k+1+m,k+1+m+r] \mid \xi[k+1,k+m])}$$

$$= \Big[\frac{Q_k^{\xi[0,k-1]}(\xi[k+1+m,k+1+m+r] \mid \xi[k,k+m])}{P(\xi[k+m,k+m+r] \mid \xi[k,k+m])}\Big] \tag{3.22}$$

$$\cdot \Big[\frac{P(\xi[k+m,k+m+r] \mid \xi[k,k+m])}{P(\xi[k+m,k+m+r] \mid \xi[k+1,k+m])}\Big]$$

$$= \big[1 \pm (4)(16)^{-m-1}\big]\big[1 \pm \gamma_m\big]$$

by the induction hyphothesis (3.18) and the definition (2.4) of γ_m. Consequently the $|\cdot|$ factor on RHS (3.21) is no larger than

$$(1/4)(16)^{-m} + \gamma_m + (1/4)(16)^{-}m\gamma_m \le (2)\cdot(16)^{-m},$$

by (3.8).

\\\

The reader should note that the hypothesis (3.8) of exponential decay of γ_k is only used in this very last inequality. Notice also that we are forced to prove (3.18) for all values of m, not just $m = 1$, because the case $(m+1,k)$ is used in the induction step to prove the case $(m,k+1)$. This explains the need for exponential decay of γ_m.

For polynomially decreasing γ_k the construction just completed fails: it is impossible to choose $\delta > 0$ so that (3.11) holds for all k. However, it is still possible to represent chains with complete connections for which $\gamma_k = O(k^{-\beta}), \beta > 1$, as block processes. The construction will be somewhat more complicated. First, regeneration will only be allowed following an ocurrence of a certain "codeword" in the process $\{Y_n\}$(by redefining the state space $\mathcal{Y}$ we will arrange matters so that regeneration is allowed only after $Y_n = y_*$, where y_* is a distinguished "letter" (state) of $\mathcal{Y}$). Second, "regeneration opportunities " will be separated by increasingly lengthy intervals of time. Consequently the block length variable $\lambda(B_1)$ will end up having a somewhat longer tail.

LEMMA 3.3: *Let $\{Y_n\}_{n\in\mathbb{Z}}$ be a $\mathcal{Y}$-valued chain with complete connections. Suppose that on some probability space are defined a $\mathcal{Y}$-valued process $\{\tilde{Y}_n\}_{n\geq 0}$ and a random variable $T \in \mathbb{Z}^+$ with $ET < \infty$, such that for all $\xi_j, \zeta_j \in \mathcal{Y}$ and $k.m \in \mathbb{Z}^+$*

$$P(\tilde{Y}_{T-1} = y_*) = 1, \tag{3.23}$$

$$\begin{aligned} P(\tilde{Y}_{m+n} = \xi_n,\ 0 \le n \le k \mid T = m; \tilde{Y}_n = \zeta_n,\ 0 \le n \le m-2) \\ = P(\tilde{Y}_n = \xi_n,\ 0 \le n \le k), \end{aligned} \tag{3.24}$$

and

$$P(\tilde{Y}_n = \xi_n,\ 0 \le n \le k) = P(Y_n = \xi_n,\ 0 \le n \le k \mid Y_{-1} = y_*). \tag{3.25}$$

Then there is a block process $\{Y_n^*\}_{n\in\mathbb{Z}}$ identical in law to $\{Y_n\}_{n\in\mathbb{Z}}$, for which the block length variable $\lambda(B_1)$ satisfies

$$\lambda(B_1) \overset{\mathcal{D}}{=} T. \tag{3.26}$$

The difference between this lemma and Lemma 3.1 is that the "post-regeneration" process $\{\tilde{Y}_{T+n}\}_{n\ge 0}$ is not required to have the same law as $\{Y_n\}_{n\ge 0}$.

PROOF: Let $\{B_n\}_{n\in\mathbb{Z}}$ be a sequence of random blocks and M an integer-valued random variable satisfying (2.10), and with block distribution

$$P(B_1 = (\xi_1, \ldots, \xi_n)) = P(\tilde{Y}_n = \xi_{n+1},\ 0 \le n < m; T = m).$$

We will show that the block process $\{Y_n^*\}_{n\in\mathbb{Z}}$ obtained from $\{B_n\}$ and M via (2.12) is identical in law to $\{Y_n\}_{n\in\mathbb{Z}}$.

Let $\{\overline{Y}_n\}_{n\ge 0}$ be the process obtained by stringing together the random blocks $B_1, B_2, \ldots$, i.e.,

$$\begin{aligned} &(\overline{Y}_0, \overline{Y}_1, \ldots, \overline{Y}_{\lambda(B_1)-1}) = (B_1(1), B_1(2), \ldots, B_1(\lambda(B_1))), \\ &(\overline{Y}_{\lambda(B_1)}, \overline{Y}_{\lambda(B_1)+1}, \ldots, \overline{Y}_{\lambda(B_1)+\lambda(B_2)-1}) = \\ &(B_2(1), B_2(2), \ldots, B_2(\lambda(B_2))), \\ &\ldots . \end{aligned}$$

Using (3.23), (3.24), and an induction argument almost identical to that in the proof of Lemma 3.1, it may be shown that $\{\overline{Y}_n\}_{n\ge 0}$ is identical in law to $\{\tilde{Y}_n\}_{n\ge 0}$. Furthermore, an argument based on the Feller-Erdös-Pollard renewal theorem, much like that in the proof of Lemma 3.1, shows that as $m \to \infty$

$$\{\overline{Y}_{m+n}\}_{n\ge 0} \overset{\mathcal{D}}{\to} \{Y_n^*\}_{n\ge 0}.$$

Consequently, to show that $\{Y_n^*\}_{n\in\mathbb{Z}} \overset{\mathcal{D}}{=} \{Y_n\}_{n\in\mathbb{Z}}$ is suffices to show that as $m \to \infty$,

$$\{\tilde{Y}_{m+n}\}_{n\geq 0} \xrightarrow{\mathcal{D}} \{Y_n\}_{n\geq 0},$$

i.e., that

$$\begin{aligned}
&\lim_{m\to\infty} P(\tilde{Y}_{m+n} = \xi_n,\ 0 \leq n \leq k) \\
&= \lim_{m\to\infty} P(Y_{m+n} = \xi_n,\ 0 \leq n \leq k \mid Y_{-1} = y_*) \qquad (3.27)\\
&= P(Y_n = \xi_n,\ 0 \leq n \leq k).
\end{aligned}$$

Assume that $\{Y_n\}_{n\in\mathbb{Z}}$ and $\{\tilde{Y}_n\}_{n\geq 0}$ are defined on the same probability space, and that the two processes are independent. Define random times τ_k by

$$\tau_k = \inf\{m \geq k : \tilde{Y}_{m-n} = Y_{m-n} \ \text{ for } 0 \leq n \leq k\}.$$

It is an easy consequence of Lemma 3.2 that $\tau_k < \infty$ with probability one, for every $k = 1, 2, \ldots$ (Lemma 3.2 guarantees that regardless of the comportment of $\{Y_n\}_{0\leq n<m}$ and $\{\tilde{Y}_n\}_{0\leq n<m}$, the probability of a match at time $n = m$ is at least δ_0^2). Now for any $k, m, r \in \mathbb{Z}^+$ with $k < m$, and any choice of $\xi_n \in \mathcal{Y}$,

$$\begin{aligned}
&\mid P(\tilde{Y}_{m+n} = \xi_n,\ 0 \leq n \leq r) - P(Y_{m+n} = \xi_n,\ 0 \leq n \leq r) \mid \\
&\leq \sum_{j=k}^{m-1} \mid P(\tilde{Y}_{m+n} = \xi_n,\ 0 \leq n \leq r \mid \tau_k = j) \\
&- P(Y_{m+n} = \xi_n,\ 0 \leq n \leq r \mid \tau_k = j) \mid \cdot P(\tau_k = j) + P(\tau_k \geq m) \\
&\leq \gamma_k + P(\tau_k \geq m),
\end{aligned}$$

by the definition (2.4) of γ_k. Since $\gamma_k \downarrow 0$ (cf. (2.3)), (3.27) follows.

\\\

PROOF OF THEOREM 2.

Assume that $\{Y_n\}_{n\in\mathbb{Z}}$ is a chain with complete connections satisfying $\gamma_n = O(n^{-\beta})$ for some real $\beta > 1$. Fix $\rho < \beta$. We will show that on some probability space are defined a $\mathcal{Y}-$valued process $\{\tilde{Y}_n\}_{n\geq 0}$ and a random time $T \in \mathbb{Z}^+$ such that $ET^\rho < \infty$ and (3.23) - (3.25) are satisfied. We may assume without loss of generality that

$$\gamma_1 \leq \gamma_*, \qquad (3.28)$$

where $\gamma_* > 0$ is some small constant which will be specified shortly (see the argument in the first paragraph of the proof of Theorem 1).

Fix $\delta, 0 < \delta < 1$, and choose constants $0 < \alpha, C, C_*, \gamma_* < 1$ such

$$(1-\alpha)^\beta < (1-\delta) < (1-\alpha)^\rho, \qquad (3.29)$$

$$\delta(1 + C + \gamma_* + C\gamma_*) < 1, \tag{3.30}$$

$$(1 + C_*)(1 - \alpha)^{\beta} + (C_*/C) < 1 - \delta(1 + C + \gamma_* + C\gamma_*), \tag{3.31}$$

$$0 < (1 - C(1 - \alpha)^{\beta})^{-1} \leq 1 + C, \tag{3.32}$$

and

$$(1 - C)(1 - \gamma_*) > \delta. \tag{3.33}$$

Since $\gamma_n = O(n^{-\beta})$, there exists an increasing sequence $\{r_k\}_{k \geq 1}$ of positive integers satisfying

$$\sum_{j=1}^{k} r_j = O((1 - \alpha)^{-k}), \ k \geq 1, \tag{3.34}$$

and

$$\gamma_{\sum_{j=1}^{k} r_j} \leq C_*(1 - \alpha)^{\beta k}, \ k \geq 1. \tag{3.35}$$

Now we shall define measures $Q_k^{\cdot}$ on various sequence spaces; these will be used to build $\{\tilde{Y}_n\}_{n \geq 0}$ and T in much the same way as in the proof of Theorem 1. Both the indexing and definition of the family $\{Q_k^{\cdot}\}$ are somewhat more complicated, since "regeneration " is to take place only after occurrences of y_*, and "regeneration opportunities" are to be widely separated.

Given a sequence $\{\xi_n\}_{n \geq 0}$, define

$$\begin{aligned} t_0 &= t_0(\{\xi\}) = -1, \\ t_1 &= t_1(\{\xi\}) = \inf\{n \geq r_1 : \xi_n = y_*\}, \\ t_{k+1} &= t_{k+1}(\{\xi\}) = \inf\{n \geq t_k + r_{k+1} : \xi_n = y_*\}, \end{aligned} \tag{3.36}$$

where r_k are the integers satisfying (3.34)-(3.35). The dependence of t_k on $\{\xi\}$ will be suppressed. The $Q_k^{\cdot}$ measures are indexed by finite sequences $\xi[0, t_k]$ having precisely k occurrences of y_* at successive gaps of (at least) $r_1, r_2, \ldots, r_k$. The measure $Q_k^{\xi[0,t_k]}$ is defined on the sequence space $\mathcal{Y}^{[t_k+1,\infty)}$ as follows:

$$Q_0(\xi[0, m]) = P(\xi[0, m] \mid \xi_{-1} = y_*),$$

$$\begin{aligned} &Q_{k+1}^{\xi[0,t_{k+1}]}(\xi[1 + t_{k+1}, m + t_{k+1}]) = \\ &\quad = (1 - \delta)^{-1}\big\{Q_k^{\xi[0,t_k]}(\xi[1 + t_{k+1}, m + t_{k+1}] \mid \xi[1 + t_k, t_{k+1}]) \\ &\quad - \delta P(\xi[1 + t_{k+1}, m + t_{k+1}] \mid \xi(t_{k+1}))\big\}. \end{aligned} \tag{3.37}$$

Notice that $\xi(t_{k+1}) = y_*$, so $P(\cdot \mid \xi(t_{k+1}))$ is just the t_{k+1} translate of Q_0; moreover if $\{\tilde{Y}_n\}_{n \geq 0}$ is a process with distribution Q_0 then $\{\tilde{Y}_n\}_{n \geq 0}$ satisfies (3.25).

As in the proof of Theorem 1, it must be verifed inductively (on k) that (3.37) is a proper recursive definition, i.e., that

$$\begin{aligned} Q_k^{\xi[0,t_k]}(\cdot \mid \xi[1+t_k, t_{k+1}]) \\ \geq \delta P(\cdot \mid \xi(t_{k+1})). \end{aligned} \tag{3.38}$$

To accomplish this, we will prove by induction on k that

$$Q_k^{\xi[0,t_k]}(t_{k+n} < \infty) = 1 \tag{3.39}$$

and

$$\begin{aligned} &\left| \frac{Q_k^{\xi[0,t_k]}(\xi[1+t_{k+n}, m+t_{k+n}] \mid \xi[1+t_k, t_{k+n}])}{P(\xi[1+t_{k+n}, m+t_{k+n}] \mid \xi[t_k, t_{k+n}])} - 1 \right| \\ &\leq C(1-\alpha)^{\beta n} \end{aligned} \tag{3.40}$$

for all $k \geq 0, n \geq 0, m \geq 1$ (with the convention $\xi_{-1} = y_*$ for the case $k = 0$), and for all choices of $\xi_n \in \mathcal{Y}$, subject to the restrictions imposed by (3.36). Since $(1-C)(1-\gamma_*) > \delta$ (cf. (3.33)) and $\gamma_1 < \gamma_*$ (cf. (3.28)), the inequality (3.38) follows easily from (3.40) with $n = 1$.

To start the induction notice that by the definition (3.37) of Q_0, LHS (3.40) $=$ 0 whenever $k = 0$. Furthermore (3.39) must hold for $k = 0$ because by the ergodic theorem and (2.1), $P(Y_n = y_*$ for infinitely many $n > 0) = 1$, and $Q_0 \ll P$.

Assume now that (3.39) has been established for all k, $0 \leq k < K$. Then $Q_{k-1}^{\xi[0,t_{k-1}]}$, and hence $Q_{k-1}^{\xi[0,t_{k-1}]}(\cdot \mid \xi[1+t_{k-1}, t_k])$, attach all their mass to sequences $(\xi_n)_{n \geq 1+t_k}$ for which the entry y_* appears, i.o. . It therefore follows from (3.37) that $Q_k^{\xi[0,t_k]}$ attaches all its mass to sequences with infinitely many occurrences of y_*. Thus, conditional on the validity of (3.38), we have established (3.39).

To prove (3.40), assume that (3.40) has been established for all values of k with $0 \leq k \leq K$, and that (3.39) has been established for $0 \leq k \leq K+1$. Using the definition (3.37) of Q_{k+1}, the induction hypothesis (3.40) for $k = K$, the assumption (3.35) and the fact that $t_{k+1+n} - t_{k+1} \leq \sum_{j=1}^{n} r_j$, (3.28), (3.32) and (3.31), we have

$$\begin{aligned} &\left| \frac{Q_{K+1}^{\xi[0,t_{K+1}]}(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[1+t_{K+1}, t_{K+1+n}])}{P(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[t_{K+1}, t_{K+1+n}])} - 1 \right| \\ &= \left| \frac{Q_K^{\xi[0,t_K]}(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[1+t, t_{K+1+n}])}{P(\xi[1+t_{K+1+n}, m+t_{K+1+m}] \mid \xi[t_{K+1}, t_{K+1+n}])} - 1 \right| \\ &\cdot \left| 1 - \delta \frac{P(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi[1+t_K, t_{K+1+n}])}{Q_k^{\xi[0,t_K]}(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi[1+t_K, t_{K+1}]} \right|^{-1} \end{aligned}$$

$$= \left| \frac{Q_K^{\xi[0,t_K]}(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[1+t_K, t_{K+1+n}])}{P(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[t, t_{K+1+n}])} \cdot \right.$$

$$\left. \cdot \frac{P(\xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[t_K, t_{K+1+n}])}{P(xi[1+t_{K+1+n}, m+t_{K+1+n}] \mid \xi[t_{K+1}, t_{K+1+n}])} - -1 \right|$$

$$\cdot \left| 1 - \delta \left\{ \frac{P(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi[t_K, t_{K+1}])}{Q_K^{\xi[0,t_K]}(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi[1+t_K, t_{K+1}])} \right. \right.$$

$$\left. \left. \cdot \frac{P(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi(t_{K+1}))}{P(\xi[1+t_{K+1}, m+t_{K+1+n}] \mid \xi[t_K, t_{K+1}])} \right\} \right|^{-1}$$

$$= \left| (1 \pm C(1-\alpha)^{\beta(n+1)})(1 \pm \gamma_{t_{K+1+n} - -t_{K+1}}) - 1 \right|$$

$$\cdot \left| 1 - \delta(1 \pm C(1-\alpha)^\beta)^{-1}(1 \pm \gamma_1) \right|^{-1}$$

$$\leq \left\{ C(1-\alpha)^{\beta(n+1)} + C_*(1-\alpha)^{\beta n} + CC_*(1-\alpha)^{\beta(2n+1)} \right\}$$

$$\cdot \{1 - \delta(1+C)(1+\gamma_*)\}^{-1}$$

$$\leq C(1-\alpha)^{\beta n} \{(1-\alpha)^\beta + (C_*/C) + C_*(1-\alpha)^\beta\}$$

$$\cdot \{1 - \delta(1 + C + \gamma_* + C\gamma_*)\}^{-1}$$

$$\leq C(1-\alpha)^{\beta n},$$

as desired. This proves (3.40) and hence (3.38), legitimizing the recursive definition of $Q_K^{\cdot}$.

We will now use the measures $Q_K^{\cdot}$ to build a regeneration point, i.e., to carry out the construction called for by Lemma 3.3. Let $(\Omega, \mathcal{F}, Q)$ be a probability space on which are defined *independent* random processes $\{\tilde{Y}_n^A\}_{n \geq 0}$ and $\{\tilde{Y}_n^B\}_{n \geq 0}$ valued in $\mathcal{Y}$, and a random variable $N \in \{1, 2, \ldots\}$ independent of $\{\tilde{Y}_n^A\}_{n \geq 0}$ and $\{\tilde{Y}_n^B\}_{n \geq 0}$, such that

$$Q(N = k) = \delta(1-\delta)^{k-1}, \ k = 1, 2, \ldots; \tag{3.41}$$

$$Q(\tilde{Y}_n^B = \xi_n, \ n \in \Lambda) = P(\xi(\Lambda) \mid \xi_{-1} = y_*); \tag{3.42}$$

$$Q(\tilde{Y}_n^A = \xi_n;\ 0 \le n \le t_k + m \le t_{k+1}) = \prod_{j=0}^{k-1} Q_j^{\xi[0,t_j]}(\xi[1+t_j, t_{j+1}]) \cdot Q_k^{\xi[0,t_k]}(\xi[1+t_k, m+t_k]) \tag{3.43}$$

for all $\Lambda \subset \mathbb{Z}^+, k \ge 0, m \ge 1$, and all choices of $\xi_n \in \mathcal{Y}$. Define a new random process $\{\tilde{Y}_n\}_{n\ge 0}$ and a random variable $T \in \{1, 2, \ldots\}$ as follows:

$$T = t_N(\{\tilde{Y}_n^A\}_{n\ge 0}) + 1; \tag{3.44}$$

and

$$\tilde{Y}_n = \begin{cases} \tilde{Y}_n^A, & n < T, \\ \tilde{Y}_n^B, & n \ge T. \end{cases} \tag{3.45}$$

Note that (3.39) implies $Q(T < \infty) = 1$. Notice also that since $\{\tilde{Y}_n^A\}$, $\{\tilde{Y}_n^B\}$, and N are independent,

$$Q(\tilde{Y}_{m+n} = \xi_n,\ 0 \le n \le k \mid T = m; \tilde{Y}_n = \zeta_n,\ 0 \le n \le m-2) = P(Y_n = \xi_n,\ 0 \le n \le k \mid Y_{-1} = \xi_*)$$

by (3.42), as required by (3.24) and (3.25). It is clear from the construction that $\tilde{Y}_{T-1} \equiv y_*$. Therefore, by the result of Lemma 3.3, to complete the proof of Theorem 2 it suffices to show that

$$Q(\tilde{Y}_n = \xi_n,\ 0 \le n \le k) = P(Y_n = \xi_n,\ 0 \le n \le k \mid Y_{-1} = y_*) \tag{3.46}$$

and

$$ET^\rho < \infty. \tag{3.47}$$

The proof of (3.46) is virtually identical to the argument in the proof of Theorem 1 showing that $\{Y_n^*\}_{n\ge 0} = \{Y_n\}_{n\ge 0}$ (see (3.15) and the ensuing discussion); only the indexing needs changing. Consequently we shall omit it.

To prove (3.47) notice first that there is a constant $\epsilon > 0$ such that for all $k = 0, 1, \ldots$ and all choices of $\xi_n \in \mathcal{Y}$

$$Q_k^{\xi[0,t_k]}(\xi(t_k + m + 1) = y_* \mid \xi[1 + t_k, m + t_k]) \ge \epsilon. \tag{3.48}$$

This is an easy consequence of (3.40) (with $n = 0$), Lemma 3.2, and assumption (2.1). Hence

$$Q_k^{\xi[0,t_k]}(t_{k+1} > t_k + r_{k+1} + n) \le (1 - \epsilon)^n. \tag{3.49}$$

Using (3.41), (3.43), and (3.44). we therefore may stochastically bound T by

$$1 + \sum_{j=1}^{N} M_j + \sum_{j=1}^{N} r_j$$

where $M_1, M_2, \ldots$ are i.i.d., independent of N, and satisfy

$$P(M_j = n) = \epsilon(1-\epsilon)^{n-1}, \ n = 1, 2, \ldots.$$

The fact that $ET^\rho < \infty$ now follows from Minkowski' s inequality, $E(\sum_{j=1}^{N} M_j)^\rho < \infty$, (3.29) and (3.34):

$$E(\sum_{j=1}^{N} r_j)^\rho \leq K \cdot \sum_n \delta(1-\delta)^{n-1}(1-\alpha)^{-n\rho} < \infty. \qquad \backslash\backslash\backslash$$

4. The Attainable Rate of Regeneration

The purpose of this section is to show that the number of finite moments permitted the block length variable $\lambda(B_1)$ in a regenerative representation of a chain with complete connections may be limited by the rate of decay of γ_n. Specifically, we will show that for each $\rho > 1$ there is a chain with complete connections $\{Y_n\}_{n\in\mathbb{Z}}$ with the following properties:

$$\gamma_n = O(n^{-\rho+1}), \tag{4.1}$$

and if $(\{Y_n^*\}_{n\in\mathbb{Z}}, \{B_k\}_{k\in\mathbb{Z}}, M)$ is a regenerative representation of $\{Y_n\}_{n\in\mathbb{Z}}$, then

$$E\lambda(B_1)^{\rho+1} = \infty. \tag{4.2}$$

The important features of the processes in question are summarized in

PROPOSITION 4.1: *For each $\rho > 1$ there is a chain with complete connections $\{Y_n\}_{n\in\mathbb{Z}}$ valued in $\mathcal{Y} = \{0,1\}$, such that for any sequence $\xi_n \in \mathcal{Y}, \ n \geq 1$,*

$$P(Y_0 = 1 \mid Y_{-n} = \xi_n, \ n \geq 1) = 1/3 + \sum_{k=0}^{\infty} 2^{-k_\rho}\xi_{2^k}/3C_\rho,$$

where $C_\rho = \sum_{k=0}^{\infty} 2^{-k_\rho} = (1-2^{-\rho})^{-1}$. This process is unique, in the sense that any stationary process satisfying (4.3) must have the same distribution on $\{0,1\}^{\mathbb{Z}}$ as $\{Y_n\}_{n\in\mathbb{Z}}$.

The process $\{Y_n\}_{n\in\mathbb{Z}}$ is symmetric with respect to interchange of 0 and 1, i.e.,

$$\{(1-Y_n)\}_{n\in\mathbb{Z}} \stackrel{\mathcal{D}}{=} \{Y_n\}_{n\in\mathbb{Z}}; \tag{4.4}$$

moreover, $\{Y_n\}_{n\in\mathbb{Z}}$ has the property

$$\gamma_n = O(n^{-(\rho-1)}); \tag{4.5}$$

and there is a constant $C > 0$ such that for each $k = 1, 2, \ldots,$

$$P(Y_{2^k} = 1 \mid Y_0 = 1) - P(Y_{2^k} = 1) \geq C2^{-k\rho}. \tag{4.6}$$

PROOF: Let $(\Omega, \mathcal{F}, P)$ be a probability space on which are defined i.i.d. uniform $(0,1)$ random variables $\{U_n\}_{n\in\mathbb{Z}}$ (i.e., $P\{U_n \leq t) = t,\ \forall t \in [0,1]$). Define $\{Y_n\}_{n\in\mathbb{Z}}$ as follows:

$$Y_n = \begin{cases} 1, & \text{if } U_n < 1/3; \\ 0, & \text{if } U_n > 2/3; \\ Y_{n-2^k}, & \text{if } \sum_{j=0}^{k-1} 2^{-j\rho} \leq 3C_\rho(U_n - -1/3) < \sum_{j=0}^{k} 2^{-j\rho}. \end{cases} \tag{4.7}$$

for $n \in \mathbb{Z}$, $k \in \mathbb{Z}$, $k \geq 0$. Notice that this definition "back references" at those n for which $1/3 \leq U_n \leq 2/3$; however, since the $\{U_m\}_{m\in\mathbb{Z}}$ are i.i.d. and $P(1/3 \leq U_m \leq 2/3) = 1/3$, the number of back-references needed to ascertatin the value of Y_n is geometrically distributed, hence finite with probability one. Thus, with probability one. all of the variables Y_n are unambiguously determined by (4.7).

It is clear that the process $\{Y_n\}_{n\in\mathbb{Z}}$ is stationary, since $\{U_n\}_{n\in\mathbb{Z}}$ is . It is also clear that (4.4) holds, because $\{1 - Y_n\}_{n\in\mathbb{Z}}$ is the process which would be produced by the mechanism (4.7), if, instead of using the i.i.d. uniform $-(0,1)$ sequence $\{U_n\}_{n\in\mathbb{Z}}$ one used the i.i.d. uniform $-(0,1)$ sequence $\{U_n^*\}_{n\in\mathbb{Z}}$, where

$$U_n^* = \begin{cases} U_n, & \text{if } 1/3 \leq U_n \leq 2/3; \\ U_n - 2/3 & \text{if } 2/3 \leq U_n \leq 1; \\ U_n + 2/3 & \text{if } 0 \leq U_n < 1/3. \end{cases}$$

Furthermore, all finite configurations of 0's and 1's have positive probability, as demanded by (2.1), because (4.7) clearly implies that for all $\xi_n \in \{0,1\}$,

$$P\{Y_n = \xi_n,\ 1 \leq n \leq r\} \geq 3^{-r}.$$

To prove that $\{Y_n\}_{n\in\mathbb{Z}}$ is a chain with complete connections it will suffice to show that for any sequence $\xi_n \in \{0,1\}$, $n \geq 1$,

$$\begin{aligned} \lim_{m\to\infty} & P(Y_0 = 1 \mid Y_{-n} = \xi_n,\ 1 \leq n \leq m) \\ &= 1/3 + \sum_{k=0}^{\infty} 2^{-k_\rho} \xi_{2^k}/3C_\rho, \end{aligned} \tag{4.8}$$

and that (4.5) holds. (cf. (2.2) and (2.3)). To establish (4.8) it suffices to look only at $m = 2^k,\ k = 1, 2, \ldots$. Now

$$\begin{aligned}
P(Y_0 = 1 \mid Y_{-n} = \xi_n,\ 1 \le n \le 2^k) \\
&= P(U_0 < 1/3) \\
&\quad + \sum_{j=0}^{k} \xi_{2^j} \cdot P\Big(\sum_{i=0}^{j-1} 2^{-i\rho} \le 3C_\rho(U_0 - 1) < \sum_{i=0}^{j} 2^{-i\rho}\Big) \\
&\quad + P\Big(Y_0 = 1,\ 1/3 + \sum_{j=0}^{k} 2^{-j\rho}/3C_\rho \le U_0 \le 2/3 \mid Y_{-n} = \xi_n,\ 1 \le n \le 2^k\Big) \\
&= 1/3 + \sum_{j=0}^{k} 2^{-j\rho} \xi_{2^j}/3C_\rho \\
&\quad + P\Big(Y_0 = 1, \sum_{j=0}^{k} 2^{-j\rho}/3C_\rho \le U_0 - 1/3 \le 1/3 \mid Y_{-n} = \xi_n,\ 1 \le n \le 2^k\Big).
\end{aligned} \tag{4.9}$$

Since U_0 is independent of $\{U_n\}_{n \le -1}$, it is independent of $\{Y_n\}_{n \le -1}$; consequently the last term on RHS (4.9) is no larger than

$$\begin{aligned}
P\Big(\sum_{j=0}^{k} 2^{-j\rho}/3C_\rho \le U_0 - 1/3 \le 1/3\Big) \\
&= (1/3)\Big(1 - \sum_{j=0}^{k} 2^{-j\rho}/C_\rho\Big) \\
&\to 0 \text{ as } k \to \infty.
\end{aligned}$$

This proves (4.8).

To show that $\gamma_n = O(n^{-(\rho-1)})$ it suffices to show that $\log(1 + \gamma_n) = O(n^{-(\rho-1)})$. Recall that

$$\gamma_m = \sup\Big\{\Big|\frac{P(Y_n = \xi_n,\ 0 \le n \le k \mid Y_n = \xi_n,\ n < 0)}{P(Y_n = \xi_n^*,\ 0 \le n \le k \mid Y_n = \xi_n^*,\ n < 0)} - 1\Big| :$$

$$\xi_n,\ \xi_n^* \in \mathcal{Y} \text{ and } \xi_n = \xi_n^* \text{ for } -m \le n \le k\Big\}.$$

Since

$$P(Y_n = \xi_n,\ 0 \le n \le k \mid Y_n = \xi,\ n < 0),$$

$$= \prod_{r=0}^{k} P(Y_r = \xi_r \mid Y_n = \xi_n,\ n < r),$$

to show that $\log(1 + \gamma_m) = O(m^{-(\rho-1)})$ it suffices to show that

$$\sup\left\{\left|\log \frac{P(Y_0=\xi_0 \mid Y_n=\xi_n,\ n<0)}{P(Y_0=\xi_0^* \mid Y_n=\xi_n^*,\ n<0)}\right| : \xi_n=\xi_n^* \ \text{ for } -m\le n\le 0\right\}$$

$$= O(m^{-\rho}).$$

But since $P(Y_0=\xi_0^* \mid Y_n=\xi_n^*,\ n<0)\ge 1/3$,

$$\left|\log \frac{P(Y_0=\xi_0 \mid Y_n=\xi_n,\ n<0)}{P(Y_0=\xi_0^* \mid Y_n=\xi_n^*,\ n<0)}\right|$$

$$\le \bigl|\log(1+3|P(Y_0=\xi_0 \mid Y_n=\xi_n,\ n<0) \\ - P(Y_0=\xi_0^* \mid Y_n=\xi_n^*,\ n<0)|)\bigr|$$

$$\le |\log(1+3\sum_{k:2^k\ge m} 2^{-k\rho}/3C_\rho)| \\ = O(m^{-\rho}).$$

for all choices of $\xi_n,\xi_n^*\in\{0,1\}$ such that $\xi_n=\xi_n^*$ for $-m\le n\le 0$. This proves (4.5).

To prove that $\{Y_n\}_{n\in\mathbb{Z}}$ is the essentially *unique* stationary process for which (4.3) holds. we assume that $\{\overline{Y}_n\}_{n\in\mathbb{Z}}$ is another. Without loss of generality we may assume that $\{\overline{Y}_n\}_{n\in\mathbb{Z}}$ is defined on $(\Omega,\mathcal{F},P)$. For each sequence $\xi=(\xi_n)_{n\le -1}$, define a process $\{Y_n(\xi)\}_{n\in\mathbb{Z}}$ as follows:

$$Y_n(\xi)=\xi_n,\ n\le -1;$$

$$Y_n(\xi)=\begin{cases} 1, & \text{if } U_n<1/3,\ n\ge 0;\\ 0, & \text{if } U_n>2/3,\ n\ge 0;\\ Y_{n-2^m}(\xi), & \text{if } \sum_{j=0}^{m-1} 2^{-j\rho}\le 3C_\rho(U_n--1/3)\le \sum_{j=0}^{m} 2^{-j\rho},\\ & \quad n\ge 0, m\ge 0.\end{cases}$$

Let $Y=(Y_n)_{n\le -1}$ and $\overline{Y}=(\overline{Y}_n)_{n\le -1}$. By (4.7),

$$Y_n=Y_n(Y),\ \forall n\ge 0.$$

Define a new process $\{Z_n\}_{n\in\mathbb{Z}}$ as follows:

$$Z_n=\overline{Y}_n,\ n\le -1;$$

$$Z_n=Y_n(\overline{Y}),\ n\ge 0.$$

Since we assumed that $\{\overline{Y}_n\}_{n\in\mathbb{Z}}$ satisfies (4.3), i.e., that

$$P(\overline{Y}_0 = 1 \mid \overline{Y}_n = \xi_{-n},\ n < 0) = 1/3 + \sum_{k=0}^{\infty} 2^{-k\rho}\xi_{2^k}/3C_\rho,$$

it follows that the processes $\{Z_n\}_{n\in\mathbb{Z}}$ and $\{\overline{Y}_n\}_{n\in\mathbb{Z}}$ are identical in law. We will show that $\{Z_n\}_{n\in\mathbb{Z}}$ and $\{Y_n\}_{n\in\mathbb{Z}}$ are equal in law by showing that there exists a random time $T \in \mathbb{Z}^+$ such that

$$P(T < \infty) = 1, \tag{4.10}$$

and

$$Y_n = Z_n \quad \text{on} \quad n \geq T. \tag{4.11}$$

(By the ergodic theorem (4.10) and (4.11) prove that $\{Y_n\} \overset{\mathcal{D}}{=} \{Z_n\}$).

To prove (4.10) and (4.11) we will show that there is a random time $T \in \mathbb{Z}^+$ satisfying (4.10), and such that for any two sequences $\xi = (\xi_n)_{n\leq -1}$ and $\zeta = (\zeta_n)_{n\leq -1}$,

$$Y_n(\xi) = Y_n(\zeta) \quad \text{on} \quad n \geq T. \tag{4.12}$$

Notice first that $Y_n(\xi)$ is monotone in ξ, i.e., if $\xi_n \leq \zeta_n$ for all $n \leq -1$, then $Y_n(\xi) \leq Y_n(\zeta)$ for all $n \in \mathbb{Z}$. Thus if $\xi^{(1)}$ and $\xi^{(0)}$ are the sequences given by $\xi_n^{(1)} = 1$ for all $n \leq -1$, $\xi_n^{(0)} = 0$ for all $n \leq -1$, then for all sequences ξ,

$$Y_n(\xi^{(0)}) \leq Y_n(\xi) \leq Y_n(\xi^{(1)}).$$

It therefore will suffice to show that there is a random time $T \in \mathbb{Z}^+$ satisfying (4.10), such that

$$Y_n(\xi^{(0)}) = Y_n(\xi^{(1)}) \quad \text{if} \quad n \geq T. \tag{4.13}$$

We introduce some terminology borrowed from the realm of branching processes. Call each 1 appearing in the sequence $\{Y_n(\xi)\}_{n\in\mathbb{Z}}$ an *individual*. If $n \geq 0$, $\sum_{j=0}^{m-1} 2^{-j\rho} \leq 3C_\rho(U_n - 1/3) \leq \sum_{j=0}^{m} 2^{-j\rho}$, and $Y_{n-2^m}(\xi) = 1$, then call the individual born at time n an *offspring* of the individual born at time $n - 2^m$. For a given individual, the offspring, the offspring's offspring, etc., will be referred to as *descendants* of the individual. Now for any individual born at a time $n \geq 0$ the expected number of offspring is

$$\begin{aligned}
\sum_{m=0}^{\infty} E\mathbb{1}\{\sum_{j=0}^{m-1} 2^{-j\rho} \leq 3C_\rho(U_{n+2^m} - 1/3) \leq \sum_{j=0}^{m} 2^{-j\rho}\} \\
= \sum_{m=0}^{\infty} 2^{-m\rho}/3C_\rho \\
= 1/3;
\end{aligned}$$

consequently the expected number of descendants of such an individual is $\sum_{n=1}^{\infty} 3^{-n} = 1/2$. Similarly, for an individual born at time $-n$, where $2^m < n \leq 2^{m+1}$, the expected number of offspring is $2^{-(m+1)\rho}/3$ and the expected number of descendants is $2^{-(m+1)\rho}/2$.

The upstart of all this is that the expected *total* number of descendants of all individuals born at times $n \leq -1$ in $\{Y_n(\xi^{(1)})\}_{n\in\mathbb{Z}}$ is finite. because $\sum_{m\geq 0} 2^m 2^{-m\rho} < \infty$ for $\rho > 1$. Since such descendants are born precisely at those times $n \geq 0$ when $Y_n(\xi^{(1)}) \neq Y_n(\xi^{(0)})$, the processes $\{Y_n(\xi^{(1)})\}_{n\geq 0}$ and $\{Y_n(\xi^{(0)})\}_{n\geq 0}$ must coalesce after a finite (but random) amount of time. Thus there exists $T \in \mathbb{Z}^+$ such that (4.10) and (4.12) hold: this completes the proof of uniqueness.

To prove (4.6), it suffices to show that there is a constant $C^* > 0$ such that for all $k = 1, 2, \ldots,$

$$P(Y_{2^k} = 1 \mid Y_0 = 1) - P(Y_{2^k} = 1 \mid Y_0 = 0) \geq C^* 2^{-k\rho}, \tag{4.14}$$

because by (4.4) $P(Y_{2^k} = 1) = 1/2(P(Y_{2^k} = 1 \mid Y_0 = 1) + P(Y_{2^k} = 1 \mid Y_0 = 0))$.

We will rely on the formula

$$P(Y_m = 1 \mid Y_0 = \zeta_0) = \int_{(\xi_n)_{n<m}} P(Y_m = 1 \mid Y_n = \xi_n,\ n < m) \cdot P((Y_n)_{n<m} \in d(\xi_n)_{n<m} \mid Y_0 = \zeta_0) \tag{4.15}$$

for $\zeta_0 = 0$ and 1. The range of integration is the set of all possible configurations $(\xi_n)_{n<m}$ of zeros and ones on the interval $(-\infty, m)$.

Note that for all pairs of configurations $(\xi_n)_{n<2^k}$ and $(\xi_n^*)_{n<2^k}$ such that $\xi_0 = 1,\ \xi_0^* = 0$, and $\xi_n^* = \xi_n$ for all $n \neq 0,\ n < 2^k$,

$$P(Y_{2^k} = 1 \mid Y_n = \xi_n,\ n < 2^k) - P(Y_{2^k} = 1 \mid Y_n = \xi_n^*,\ n < 2^k) = 2^{-k\rho}/3C_\rho,$$

by (4.3). Notice also that $P(Y_{2^k} = 1 \mid Y_0 = 0; Y_n = \xi_n,\ n < 2^k \text{ and } n \neq 0)$ is a nondecreasing function of $(\xi_n)_{n<2^k;\ n\neq 0}$, again by (4.3) (here nondecreasing means with respect to the partial order $(\xi_n) \leq (\xi_n^*)$ iff $\xi_n \leq \xi_n^*$ for all n). Consequently to prove (4.14) with $C^* = 1/3C_\rho$, it suffices to show that for every measurable nondecreasing function $f((\xi_n)_{n<2^k;n\neq 0})$ of configurations on $(-\infty, 0) \cup (0, 2^k)$

$$\int_{(\xi_n)_{n<2^k;n\neq 0}} f((\xi_n))P(Y_n=\xi_n,\ n<2^k,\ n\neq 0\mid Y_0=1)$$
$$\geq \int_{(\xi_n)_{n<2^k;n\neq 0}} f((\xi_n))P(Y_n=\xi_n,\ n<2^k,\ n\neq 0\mid Y_0=0) \tag{4.16}$$

To establish this, we will use the sequence $\{U_n\}_{n\in\mathbb{Z}}$ of uniform -(0, 1) random variables which were employed in the definition (4.7) of $\{Y_n\}_{n\in\mathbb{Z}}$ to create two new sequences $\{Y^{(0)}\}_{n\in\mathbb{Z}}$ and $\{Y_n^{(1)}\}_{n\in\mathbb{Z}}$ such that

$$Y_n^{(0)} \leq Y_n^{(1)} \text{ for all } n\in\mathbb{Z}\,; \tag{4.17}$$

and

$$P(Y_n^{(\zeta)}=\xi_n,\ n\in\Lambda)=P(Y_n=\xi_n,\ n\in\Lambda\mid Y_0=\zeta) \tag{4.18}$$

for $\zeta=0$ and 1 and any choice of $\xi_n\in\{0,1\}$, any finite subset $\Lambda\subset\mathbb{Z}$.

Given a realization of the uniform sequence $\{U_n\}$, let $(K_1,K_2,\ldots,K_R)$ be the longest sequence of nonnegative integers such that

$$\sum_{j=0}^{K_1-1}2^{-j\rho}\leq 3C_\rho(U_0-1/3)\leq\sum_{j=0}^{K_1}2^{-j\rho};$$

and

$$\sum_{j=0}^{K_{r+1}-1}2^{-j\rho}\leq 3C_\rho(U_{-\sum_{n=1}^{r}2^{K_n}}-1/3)\leq\sum_{j=0}^{K_{r+1}}2^{-j\rho}$$

for $r=1,2,\ldots,R-1$. Notice that this is the chain of back-references one would make to determine the value of Y_0 by (4.7); it is finite with probability one. Define

$$Y_n^{(\zeta)}=\begin{cases}\zeta, & \text{if } n=0;\\ \zeta, & \text{if } n=-\sum_{m=1}^{r}2^{K_m},\ 1\leq r\leq R;\\ Y_n, & \text{otherwise.}\end{cases}$$

for $\zeta=0$ and 1. It is clear that $Y_n^{(1)}\geq Y_n^{(0)}$ for all $n\in\mathbb{Z}$. That (4.18) holds follows from (4.7), because conditioning on $Y_0=\zeta$ is the same as conditioning on the event that the chain of back-references starting at $n=0$ in (4.7) ends in a ζ; i.e., that

$$U_{-(\sum_{n=1}^{R}2^{K_n})}\begin{cases}<1/3 & (\text{for } \zeta=1),\\ >2/3 & (\text{for } \zeta=0).\end{cases}$$

This completes the proof of (4.14), and of Proposition 4.1.

\\\

That the processes $\{Y_n\}_{n\in\mathbb{Z}}$ constructed in Proposition 4.1 enjoy property (4.2) will follow immediately from (4.6) and the next result.

PROPOSITION 4.2: *Suppose $\{Y_n^*\}_{n\in\mathbb{Z}}$ is a block process valued in $\mathcal{Y}=\{0,1\}$, constructed from the sequence $\{B_n\}_{n\in\mathbb{Z}}$ of random blocks, and suppose that $\{Y_n^*\}$ is a chain with complete connections. Suppose also that for all choices of $\xi_1,\xi_2,\ldots,\xi_m\in\{0,1\}$*

$$P(Y_n^*=\xi_n,\ 1\le n\le m)>0. \tag{4.19}$$

If $E\lambda(B_1)^{\rho+1}<\infty$ for some $\rho>1$, then

$$\left|P(Y_m^*=1\mid Y_0^*=1)-P(Y_m^*=1)\right|=o(m^{-\rho})\ \text{ as }\ m\to\infty. \tag{4.20}$$

PROOF: This is a relatively straight forward application of a coupling result due to Ney [16]. We paraphrase Ney's result as follows:

> Given a probability distribution $\{p_n\}_{n\ge0}$ such that $\sum np_n<\infty$ and g.c.d. $\{n\ge1:p_n>0\}=1$, given independent "initial" random variables $X_0,\overline{X}_0\ge0$, both integer-valued, and given independent i.i.d. sequences $\{X_n\}_{n\ge1}$ and $\{\tilde{X}_n\}_{n\ge1}$ all with distribution $\{p_m\}_{m\ge0}$, and such that $\{X_n\}_{n\ge0}$ and $\{\tilde{X}_n\}_{n\ge0}$ are all independent, there exist random (integer) times $L,\tilde{L}\ge0$ such that
>
> for each $n,\tilde{n}\in\mathbb{Z}^+$ the event $(L\ge n,\tilde{L}\ge\tilde{n})$ is a function of $(X_0,X_1,\ldots,X_n;\tilde{X}_0,\tilde{X}_1,\ldots,\tilde{X}_{\tilde{n}})$; (4.21)
>
> $$S_L=\tilde{S}_{\tilde{L}},\ \text{ where } S_n=\sum_{i=0}^n X_i\ \text{ and }\ \tilde{S}_n=\sum_{i=0}^n\tilde{X}_i; \tag{4.22}$$
>
> $$P(S_L>x)=O(P(X_0+\tilde{X}_0+\sum_{i=1}^N Z_i>x)), \tag{4.23}$$
>
> where $N,X_0,\tilde{X}_0$, and $Z_n,\ n\ge1$, are all independent, $\{Z_n\}_{n\ge1}$ are i.i.d., $Ee^{\theta N}<\infty$ for some $\theta>0$, and
>
> $$P(Z_n>z)=O(\sum_{n=z}^{\infty}(n-z)p_n). \tag{4.24}$$

This is slightly different than the statement of Lemma 1 in [16]: there it is only asserted that $S_L \overset{\mathcal{D}}{=} \tilde{S}_{\tilde{L}}$. However, the proof of Lemma 1 in [16] actually establishes the stronger statement $S_L = \tilde{S}_{\tilde{L}}$.

Recall from (2.5)-(2.7) that a block process $\{Y_n^*\}$ is constructed from a sequence of random variable M satisfying (2.1). Let $p_n = P\{\lambda(B_1) = n\}$, $n \geq 0$, and assume that

$$\text{g.c.d. } \{n \geq 1 : p_n > 0\} = 1, \tag{4.25}$$

so that the hypothesis of Ney's result is satisfied. We will show later how this assumption may be removed. Finally, suppose that $E\lambda(B_1)^{\rho+1} < \infty$.

Let $\{\tilde{B}_n\}_{n\in\mathbb{Z}}$ be a sequence of random blocks, and $\tilde{M} \geq 1$ an integer - valued random variable, satisfying

$$\{\tilde{B}_n\}_{n\in\mathbb{Z}} \text{ and } \tilde{M} \text{ are independent of } \{B_n\}_{n\in\mathbb{Z}} \text{ and } M; \tag{4.26}$$

$$\begin{aligned}&\text{the variables } \tilde{B}_n,\ n \in \mathbb{Z},\ \text{are independent, and for}\\ &\text{all } n \neq 0, \tilde{B}_n \overset{\mathcal{D}}{=} B_1; \text{ and}\end{aligned} \tag{4.27}$$

$$\begin{aligned}P(\tilde{B}_0 = b, \tilde{M} = k \mid \{\tilde{B}_n\}_{n\neq 0}) \\ = \frac{P(B_1 = b)\mathbb{1}\{1 \leq k \leq \lambda(b); b(k) = 1\}}{E\sum_{j=0}^{\lambda(B_1)} \mathbb{1}\{B_1(j) = 1\}}.\end{aligned} \tag{4.28}$$

Notice that (4.28) is just the conditional distribution of (B_0, M) given that $Y_0^* = 1$: cf. (2.5). Define random variables $\{X_n\}_{n\geq 0}$, $\{\tilde{X}_n\}_{n\geq 0}$, $\{S_n\}$, $\{\tilde{S}_n\}$, $\{N(n)\}$, and $\{\tilde{N}(n)\}$ by

$$\begin{aligned}
X_n &= \lambda(B_n),\ n \geq 1; \\
\tilde{X}_n &= \lambda(\tilde{B}_n),\ n \geq 1; \\
X_0 &= \lambda(B_0) - M; \\
\tilde{X}_0 &= \lambda(\tilde{B}_0) - \tilde{M}; \\
S_n &= \sum_{j=0}^{n} X_j,\ n \geq 0; \\
\tilde{S}_n &= \sum_{j=0}^{n} \tilde{X}_j,\ n \geq 0; \\
N(n) &= \min\{m \geq 0 : S_m \geq n\},\ n \geq 0; \text{ and} \\
\tilde{N}(n) &= \min\{m \geq 0 : \tilde{S}_m \geq n\},\ n \geq 0.
\end{aligned} \tag{4.29}$$

By Ney's result there exist random times $L, \tilde{L}$ such that (4.21)-(4.23) hold; observe that by (4.22),

$$N(S_L) = L,$$
$$\tilde{N}(\tilde{S}_{\tilde{L}}) = \tilde{L}.$$

Define a random process $\{\tilde{Y}_n\}_{n\geq 0}$ as follows:

$$\tilde{Y}_n = \begin{cases} \tilde{B}_{\tilde{N}(n)}(\lambda(\tilde{B}_{\tilde{N}(n)}) - \tilde{S}_{\tilde{N}(n)} + n) & \text{if } \tilde{N}(n) \leq \tilde{L}; \\ B_{N(n)}(\lambda(B_{N(n)}) - S_{N(n)} + n) & \text{if } N(n) > L. \end{cases} \tag{4.30}$$

Observe that by construction (cf. (4.28)), $\tilde{Y}_n = 1$; in fact, for all $k \geq 1$, $\xi_1, \ldots, \xi_k \in \{0, 1\}$,

$$P(\tilde{Y}_n = \xi_n,\ 1 \leq n \leq k) = P(Y_n^* = \xi_n,\ 1 \leq n \leq k \mid Y_0^* = 1). \tag{4.31}$$

Moreover, since after $S_L = \tilde{S}_{\tilde{L}}$ the construction of $\{\tilde{Y}_n\}$ proceeds by stringing together the *same* blocks $\{B_n\}$ that are used in the construction of $\{Y_n^*\}$,

$$\tilde{Y}_n = Y_n^* \text{ for } n > S_L. \tag{4.32}$$

Combining (4.31) and (4.32), we have

$$\begin{aligned} &\left|P(Y_m^* = 1 \mid Y_0^* = 1) - P(Y_m^* = 1)\right| \\ &\qquad \leq P(S_L \geq m) \end{aligned} \tag{4.33}$$

Consequently, to prove (4.20), it suffices to show that

$$P(S_L \geq m) = O(m^{-\rho}). \tag{4.34}$$

But this follows almost immediately from (4.23) and the independence of $N, X_0, \tilde{X}_0$, and Z_n, $n \geq 1$. For $E\lambda(B_1)^{\rho+1} < \infty$ implies that $EX_0^\rho < \infty$ (cf. (2. ?)), $E\tilde{X}_0^\rho < \infty$ (cf. (4.28)), and $EZ_n^\rho < \infty$ (cf. (4.24)). Thus $E(X_0 + \tilde{X}_0 + \sum_{i=1}^{N} Z_i)^\rho < \infty$, and (4.34) follows.

In proving (4.20) we made an extraneous assumption about the distribution of the block length variable $\lambda(B_1)$, to wit, (4.25), according to which the distribution of $\lambda(B_1)$ is supported by no proper subgroup of $\mathbb{Z}$. Assume now instead of (4.25) that

$$\text{g.d.c.}\{n \geq 1 : \rho_n > 0\} = d > 1. \tag{4.35}$$

Using the fact that $\{Y_n^*\}_{n\in\mathbb{Z}}$ is ergodic (a chain with complete connections is always ergodic), we will argue that

$$\mathcal{D}(\{Y_n^*\}_{n\in\mathbb{Z}} \mid M \equiv 0 \bmod d) \\ = \mathcal{D}(\{Y_n^*\}) \tag{4.36}$$

This will allow us to replace the variables M and $\tilde{M}$ in the preceding argument by variables which are congruent to 0 mod d; the rest of the argument may then be repeated almost verbatim.

To prove (4.36), define a new process $\{W_n\}_{n\in\mathbb{Z}}$ by

$$W_n = (Y_{nd+1}^*, Y_{nd+2}^*, \dots Y_{(n+1)d}^*);$$

these random variables assume values in $\{0,1\}^d$. Clearly the process $\{W_n\}_{n\in\mathbb{Z}}$ is stationary and ergodic, since $\{Y_n^*\}_{n\in\mathbb{Z}}$ is. Moreover, *conditional on* $M \equiv 0(\bmod d)$, $\{W_n\}_{n\in\mathbb{Z}}$ is still a stationary process, because it is a block process!

This implies

$$\mathcal{D}(\{W_n\}_{n\in\mathbb{Z}} \mid M \equiv 0 \bmod d) \\ = \mathcal{D}(\{W_n\}_{n\in\mathbb{Z}}),$$

because otherwise the probability distribution $\mathcal{D}(\{W_n\}_{n\in\mathbb{Z}})$ could be written as a nontrivial convex combination of shift-invariant probability distribution on $(\{0,1\}^d)^{\mathbb{Z}}$, contraditing the fact that $\{W_n\}_{n\in\mathbb{Z}}$ is ergodic.

\\\

5. Gibbs Processes are Chains with Complete Connections

The sole purpose of this section is to prove Proposition 2 of section 2. To prove that a Gibbs process $\{Y_n\}_{n\in\mathbb{Z}}$ is a chain with complete connections, we must verify (2.1), (2.2) and (2.3). Now the definition (2.17) of a Gibbs state already guarantees that there is a system of regular conditional distributions for the configurations on each finite set, given the "paste" and "future": hence (2.2) is trivial. Moreover, each of the Gibbs ensembles $\mu_{\Lambda|\xi}$ attributes positive probability to every configuration on Λ (cf. (2.14)): consequently (2.1) is trivial. To prove Proposition 2, therefore, it suffices to show (2.18).

Define new constants γ_m^* by

$$\gamma_m^* = \sup_{k\geq 1}\Big\{\Big|\frac{P(Y_n=\xi_n,\ 0\leq n\leq k \mid Y_n=\zeta_n,\ n<0, n>k)}{P(Y_n=\xi_n,\ 0\leq n\leq k \mid Y_n=\zeta_n^*,\ n<0, n>k)} - 1\Big| : \\ \xi_n,\ \zeta_n,\ \zeta_n^* \in \mathcal{Y} \quad \text{and} \\ \zeta_n = \zeta_n^*,\ n>k \ \text{ and } \ -m\leq n\leq -1\Big\}.$$

Integrating over the set $\mathcal{Y}^{(k,\infty)}$ of all possible "futures" one easily obtains

$$\gamma_m \leq \gamma_m^*.$$

Thus it suffices to show that $\gamma_m^* = O(\sum_{j=m}^{\infty} \delta_j)$.

Fix $k \geq 1$, and $\xi_n, \zeta_n, \zeta_n^* \in \mathcal{Y}$ such that $\zeta_n = \zeta_n^*$ for $n > k$ and $-m \leq n \leq -1$. Let $\Lambda_k = \{0, 1, \ldots, k\}$, and let ζ, ζ^* be the configurations on $\mathbb{Z} \setminus \Lambda_k$ with coordinates ζ_n, ζ_n^*, respectively. Then

$$\begin{aligned} &\frac{P(Y_n = \xi_n,\ 0 \leq n \leq k \mid Y_n = \zeta_n,\ n \in \mathbb{Z} \setminus \Lambda_k)}{P(Y_n = \xi_n,\ 0 \leq n \leq k \mid Y_n = \zeta_n^*,\ n \in \mathbb{Z} \setminus \Lambda_k)} \\ &= \frac{\mu_{\Lambda_k|\zeta}(\xi)}{\mu_{\Lambda_k|\zeta^*}(\xi)} \\ &= \frac{Z(\Lambda_k;\zeta^*)}{Z(\Lambda_k;\zeta)} \cdot \frac{\exp\{-\sum_{n=0}^{k} H(\alpha^n(\xi \vee \zeta))\}}{\exp\{-\sum_{n=0}^{k} H(\alpha^n(\xi \vee \zeta^*))\}} \end{aligned} \tag{5.1}$$

by (2.17) and (2.14). Now the condition $\zeta_n = \zeta_n^*$ for $n > k$ and $-m \leq n \leq -1$ implies

$$\begin{aligned} \sum_{n=0}^{k} &\left| H(\alpha^n(\xi \vee \zeta)) - H(\alpha^n(\xi \vee \zeta^*)) \right| \\ &\leq \sum_{n=0}^{k} \delta_{n+m} \\ &\leq \sum_{j=m}^{k} \delta_j, \end{aligned}$$

by (2.13). Now this holds for all configuration ξ on Λ_k. Hence, using (2.15), we have

$$\text{RHS } (5.1) = (1 \pm O(\sum_{j=m}^{\infty} \delta_j))\ (1 \pm O(\sum_{j=m}^{\infty} \delta_j))$$

which shows that $\gamma_m^* = O(\sum_{j=m}^{\infty} \delta_j)$.

\\\

REFERENCES

[1] Athreya, K., Mac Donald, D., and Ney, P. (1978) Limit theorems for semi-Markov processes and renewal theory for Markov chains. *Ann. Prob.* 6, 788-797.

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

[3] Bowen, R. (1975) *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Lecture Notes in Math. No. 470. Springer, Berlin.

[4] Chung, K.L. (1961) *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin.

[5] Dobrushin, R. (1968) The description of a random field by means of conditional probabilities and conditions of its regularity. *Th. Prob. and its Appl.* 13, 197-224.

[6] Doeblin, W. (1938) Sur deux problèmes de M. Kolmogorov concernant les chaines denombrables. *Bull. Soc. Math. France* 66, 210-220.

[7] Doeblin, W. (1939) Sur certains mouvements aleatoires discontinus. *Skand. Akt.* 22, 211-222.

[8] Doeblin, W. et Fortet, R. (1937) Sur les chaines à liaisons complètes. *Bull. Soc. Math. France* 65, 132-148.

[9] Feller, W. (1968) *An Introduction to Probability Theory and its Applications* Vol. I, 3rd ed. Wiley, New York.

[10] Freedman, D. (1971) *Markov Chains.* Holden-Day, San Francisco.

[11] Harris, T. E. (1955) On chains of infinite order. *Pac. J. Math.* 5, 707-724.

[12] Iosifescu, M. and Theodorescu, R. (1969) *Random Processes and Learning.* Springer, Berlin.

[13] Kesten, H. (1974) Renewal theory for Markov chains. *Ann. Prob.* 3, 355-387.

[14] Lanford, O. E., and Ruelle, D. (1969) Observables at infinity and states with short range correlations in statistical mechanics. *Comm. Math. Phys.* 13, 194-215.

[15] Ledrappier, F. (1974) Sur la condition de Bernoulli faible et ses applications. *Theorie Ergodique,* Lecture Notes in Math. No. 532. Springer, Berlin.

[16] Ney, P. (1981) A refinement of the coupling method in renewal theory. *Stoch. Proc. and their Appl.* 11, 11-26.

[17] Nummelin, E. (1978) A splitting technique for Harris recurrent Markov chains. *Z. Warsch.* 43, 309-318.

[18] Onicescu, O. and Mihoc, G. (1935) Sur les chaines statistiques. *C. R. Acad. Sci. Paris* 200, 511-512.

[19] Ruelle, D. (1978) *Thermodynamic Formalism.* Addison-Wesley, Reading, Mass.

[20] Smith, W. (1955) Regenerative stochastic processes. *Proc. Roy. Soc. London,* Ser. A. 232, 6-31

Steven P. Lalley
Department of Statistics
University of Chicago
5734 S. University Avenue
Chicago, IL 60637
United States of America