The Annals of Applied Probability
1995, Vol. 5, No. 1, 261–293

TAIL EVENTS OF SOME NONHOMOGENEOUS MARKOV CHAINS[1]

BY WOJCIECH NIEMIRO[2] AND PIOTR POKAROWSKI

University of Warsaw and Polish Academy of Sciences

We consider finite state nonhomogeneous Markov chains with one-step transition probabilities roughly proportional to powers of a small parameter, converging to zero. We examine asymptotic properties of trajectories. The analysis is based on the so-called orders of recurrence. Transient states, recurrent classes and periodic subclasses can be identified in terms of the matrix of powers. This leads to a complete description of the tail sigma field. Our theorems generalize the classical results for homogeneous chains and can also be applied to chains generated by stochastic algorithms of the "simulated annealing" type.

0. Introduction. We consider finite state, discrete time nonhomogeneous Markov chains (X_n) with transition probabilities satisfying

$$c^{-1}\varepsilon_{n+1}^{v_{ij}} \le \mathbf{P}(X_{n+1} = j \mid X_n = i) \le c\varepsilon_n^{v_{ij}},$$

where $\varepsilon_n \searrow 0$. Similar classes of chains were considered by Tsitsiklis (1989), Connors and Kumar (1989), Chiang and Chow (1989) and Borkar (1992). The primary motivation comes from the field of stochastic algorithms for global optimization. The "simulated annealing" algorithm generates Markov chains of this type [general references on simulated annealing are van Laarhoven and Aarts (1987) and the special issue of *Algorithmica* 6 (1991)]. The class of chains under consideration is fairly large. All chains of Doeblin's type (A), in particular all homogeneous ones, satisfy our assumption. On the other hand, the assumption is stronger than Doeblin's condition (B) and, consequently, leads to more conclusive results [see Cohn (1981), for example]. Classification of states according to asymptotic properties of trajectories can be described explicitly, in an algorithmic way, in terms of the matrix (v_{ij}). The state space can be decomposed into recurrent classes of asymptotically communicating states and the set of transient states. Recurrent classes can be further divided into periodic subclasses. This decomposition is equivalent to the complete description of the tail σ-field of the chain, if $\mathbf{P}(X_0 = i) > 0$ for all i.

Our approach is based on the so-called pathwise orders of recurrence, introduced by Borkar (1992). For each state i, we consider random variable α_i

Received August 1993; revised June 1994.
[1]Research supported by KBN Grant 21168/91/01.
[2]Author to whom correspondence should be addressed.
AMS 1991 *subject classifications*. Primary 60J10; secondary 90C15.
Key words and phrases. Nonhomogeneous Markov chain, tail sigma field, orders of recurrence, simulated annealing.

defined by

$$\alpha_i = \sup\left\{h \geq 0 : \sum_n \varepsilon_n^h \mathbf{1}(X_n = i) = \infty\right\}.$$

Given $\varrho = \sup\{h \geq 0 : \sum_n \varepsilon_n^h = \infty\}$, all values of the α_i's can be expressed in terms of the matrix (v_{ij}). Connors and Kumar (1989) defined orders of recurrence in a slightly different way, setting

$$\beta_i = \sup\left\{h \geq 0 : \sum_n \varepsilon_n^h \mathbf{P}(X_n = i) = \infty\right\}.$$

Following Borkar, we call β_i's mean orders of recurrence. They may depend on transition probabilities in a more complicated way, not only through the matrix (v_{ij}). Moreover, even if a β_i is known, we cannot tell whether state i is recurrent (visited infinitely often with positive probability) or transient. The analysis based on the pathwise orders is simpler and leads to results with clear probabilistic interpretation.

The outline of the paper is the following. In Section 1 we introduce the class of chains under consideration. The definitions and basic properties of the pathwise orders of recurrence are recalled in Section 2. The key results here are Theorem 2.1 [the propagation rule; due essentially to Borkar (1992)] and Theorem 2.2 [the balance equations; due to Connors (1988)]. To make the paper self-contained, we give proofs of these theorems under our slightly relaxed assumptions. Section 3 has more to do with algebra and graph theory than with probability. In this section we focus on adjusting the results of Connors and Kumar (1989) to fit our purposes. Our main results appear in Sections 4 and 5. The description of the recurrent classes of the chain, given in Section 4, is based on the preceding graph considerations. In Section 5, we decompose recurrent classes into periodic subclasses and thus identify all the atoms of the tail σ-field. In Section 6, we apply the general results to chains, produced by stochastic optimization algorithms of the "simulated annealing" type.

1. Chains with regularly diminishing transitions. Suppose $(X_n, n \geq 0)$ is a nonhomogeneous Markov chain on a finite state space $\mathscr{S}$. Assume that

$$c^{-1}\varepsilon_{n+1}^{v_{ij}} \leq \mathbf{P}(X_{n+1} = j \mid X_n = i) \leq c\varepsilon_n^{v_{ij}}, \qquad 0 < c < \infty, 0 \leq v_{ij} \leq \infty, \tag{1.1}$$

where (ε_n) is a real sequence such that

$$\begin{gathered} 0 < \varepsilon_n < 1, \\ \varepsilon_n \to 0 \quad \text{as } n \to \infty, \\ \varepsilon_{n+1} \leq \varepsilon_n \quad \text{for all } n. \end{gathered} \tag{1.2}$$

The convention is $\varepsilon^\infty = 0$. The sequence (ε_n) and the matrix (v_{ij}) are not uniquely determined by the transition probabilities. For example, we can always replace ε_n by $\varepsilon'_n = \varepsilon^r_n$ and v_{ij} by $v'_{ij} = v_{ij}/r$, with $0 < r < \infty$. Whenever we speak of a chain (X_n) satisfying (1.1) and (1.2), we will always refer to a fixed choice of (ε_n) and (v_{ij}). We call a chain satisfying these conditions a *chain with regularly diminishing transitions*. Let us bear in mind, however, that the frequency of transitions from i to j does *not* diminish if $v_{ij} = 0$. Transitions from i to j are impossible if $v_{ij} = \infty$. Note that assumption (1.1) is weaker than its counterpart in Connors and Kumar (1989). We could relax the condition of monotonicity of (ε_n) similarly to the cited authors. However, the apparent generalization in unnecessary. If $\varepsilon_{n+m} \le k\varepsilon_n$ for all n, $m \ge 0$, then ε_n can be replaced by $\tilde{\varepsilon}_n = \max_{m\ge 0} \varepsilon_{n+m}$ in (1.1).

REMARK 1.1. If a chain (X_n) satisfies condition (A) of Doeblin, then it satisfies (1.1) and (1.2).

Let us recall Doeblin's condition: For all states i and j,

$$\text{(A)}\qquad \mathbf{P}(X_{n+1} = j \mid X_n = i)\begin{cases} \text{either} & \ge \delta > 0 \quad \text{for all } n, \\ \text{or} & = 0 \quad \text{for all } n. \end{cases}$$

This corresponds to condition (1.1) with v_{ij}'s equal either to 0 or to ∞. The sequence (ε_n) can be chosen arbitrarily.

REMARK 1.2. If a chain (X_n) satisfies (1.1) and (1.2), then it satisfies condition (B) of Doeblin.

The second Doeblin condition stipulates that

$$\text{(B)}\qquad \mathbf{P}(X_{n+1} = j \mid X_n = i)\begin{cases} \text{either} & \ge \delta > 0 \quad \text{for all } n, \\ \text{or} & \to 0 \quad \text{as } n \to \infty. \end{cases}$$

REMARK 1.3. Let (X_n) satisfy (1.1) and (1.2). For $d \ge 1$, the periodically sampled chain $(\hat{X}_n = X_{nd})$ also satisfies these conditions, with $\hat{\varepsilon}_n = \varepsilon_{nd}$ and

$$\hat{v}_{ij} = \min\left\{ \sum_{1\le s\le d} v_{i(s-1)i(s)} : i = i(0), i(1), \dots, i(d) = j \right\}.$$

We will not use Remark 1.3 in this paper. Let us omit the easy proof. Another standard construction will be more important. Consider the Cartesian product of two *independent* copies (X^1_n) and (X^2_n) of the chain (X_n). Elements of the product space $\mathscr{S} \times \mathscr{S}$ are denoted by underlined letters. The following obvious fact will be exploited in Section 5.

REMARK 1.4. If (X_n) satisfies (1.1), then the double chain $[\underline{X}_n = (X^1_n, X^2_n)]$ also satisfies (1.1) with the same sequence (ε_n) and with the matrix of powers $(v_{\underline{ij}})$ given by

$$v_{\underline{ij}} = v_{i_1 j_1} + v_{i_2 j_2} \qquad \text{where } \underline{i} = (i_1, i_2), \underline{j} = (j_1, j_2).$$

Assumptions (1.1) and (1.2) will stand throughout the paper. We will be concerned with asymptotic behavior of trajectories of (X_n). Speaking more precisely, our goal will be to describe the tail σ-field

$$\mathscr{T} = \bigcap_{n\geq 0} \sigma(X_k, k \geq n).$$

Let $(\Omega, \mathscr{F}, \mathbf{P})$ be the probability space on which our chain lives. We will use curly brackets $\{\cdots\}$ for events $\{\omega \in \Omega : \cdots\}$ and $\mathbf{1}(\cdots)$ for their indicator functions. Events are denoted by capital roman letters. If $\mathbf{P}(B\backslash A) = 0$, we will write $A \subset B$ a.s. (almost surely). If (A_n) is a sequence of events, $\{A_n \text{ i.o.}\}$ (infinitely often) stands for $\bigcap_k \bigcup_{n\geq k} A_n$ and $\{A_n \text{ ult.}\}$ (ultimately), for $\bigcup_k \bigcap_{n\geq k} A_n$. Let

$$\mathscr{F}_n = \sigma(X_k, 0 \leq k \leq n).$$

By (sub)martingale we will mean, by default, a (sub)martingale with respect to $(\mathscr{F}_n)$. Sets of states are denoted by script letters. Inclusions between subsets of $\mathscr{S}$ will be denoted by $\mathscr{A} \subseteq \mathscr{B}$ or $\mathscr{A} \subsetneq \mathscr{B}$. Whenever we are concerned with convergence of a series, we will freely write $\sum_n$ instead of $\sum_{n=k}^{\infty}$, leaving k unspecified. By convention, $N = \min\{n : \cdots\}$ will be understood to be ∞ when the dotted condition is false for every n.

2. Orders of recurrence. For i, $j \in \mathscr{S}$, define random variables α_i and α_{ij} by

$$\alpha_i = \sup\left\{h \geq 0 : \sum_n \varepsilon_n^h \mathbf{1}(X_n = i) = \infty\right\}, \tag{2.1}$$

$$\alpha_{ij} = \sup\left\{h \geq 0 : \sum_n \varepsilon_n^h \mathbf{1}(X_n = i, X_{n+1} = j) = \infty\right\}, \tag{2.2}$$

where, by convention, we set $\sup\varnothing = -\infty$. We call α_i and α_{ij} the *orders of recurrence* of, respectively, state i and the transition from i to j. They are the *pathwise* orders, introduced by Borkar (1992). *Mean* recurrence orders, defined earlier by Connors and Kumar (1989), will not be exploited here. Let

$$\varrho = \sup\left\{h \geq 0 : \sum_n \varepsilon_n^h = \infty\right\}. \tag{2.3}$$

Following Connors and Kumar, we call ϱ the *order of cooling* of (ε_n). Unless otherwise stated, assume

$$0 \leq \varrho < \infty \quad \text{and} \quad \sum_n \varepsilon_n^{\varrho} = \infty. \tag{2.4}$$

The other possibility is

$$0 < \varrho \leq \infty \quad \text{and} \quad \sum_n \varepsilon_n^{\varrho} < \infty. \tag{2.4$'$}$$

In this section, we consider the two cases (2.4) and (2.4′) in parallel. On the other hand, (2.4) will stand in Sections 4, 5 and 6.

PROPOSITION 2.1. *If* (2.4) *is true, then for all i, almost surely,*

$$\sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) = \infty \quad \text{whenever } \alpha_i \geq 0. \tag{2.5}$$

If (2.4′) *is true, then for all i, almost surely,*

$$\sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) < \infty \quad \text{whenever } \alpha_i \geq 0. \tag{2.5′}$$

Similar statements hold for α_{ij}.

Let us defer the proof of Proposition 2.1. It will be given at the end of this section.

The rule of *propagation* for the pathwise orders of recurrence was established by Borkar (1992). A slightly strengthened version of his result is the following theorem.

THEOREM 2.1. *Let* α_i *and* α_{ij} *be the recurrence orders defined by* (2.1) *and* (2.2). *Let* ϱ *be given by* (2.3). *If* (2.4) *holds, then almost surely,*

$$\alpha_{ij} = \begin{cases} \alpha_i - v_{ij}, & \text{if } \alpha_i \geq v_{ij}, \\ -\infty, & \text{if } \alpha_i < v_{ij}, \end{cases} \qquad i, j \in \mathscr{S}. \tag{2.6}$$

If (2.4′) *holds, then almost surely,*

$$\alpha_{ij} = \begin{cases} \alpha_i - v_{ij}, & \text{if } \alpha_i > v_{ij}, \\ -\infty, & \text{if } \alpha_i \leq v_{ij}, \end{cases} \qquad i, j \in \mathscr{S}. \tag{2.6′}$$

To prove Theorem 2.1, we will need some auxiliary results, which will be given later in this section. The theorem will follow immediately once we prove Lemma 2.2 and Proposition 2.1.

An obvious consequence of definition (2.1) is

$$\alpha_i \geq 0 \quad \text{or} \quad \alpha_i = -\infty. \tag{2.7}$$

The so-called *balance* equations were introduced by Connors (1988) for the mean orders of recurrence. Borkar (1992) noticed they hold for the pathwise orders too. The proof remains essentially the same. We are going to supplement the equations with a condition not necessarily true for the mean orders. Note, in passing, that the following statements hold for *all* trajectories of the chain.

THEOREM 2.2 [Connors (1988); Borkar (1992)]. *Let* α_i *and* α_{ij} *be the recurrence orders defined by* (2.1) *and* (2.2). *Let* ϱ *be given by* (2.3). *We have*

$$\max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} \alpha_{ij} = \max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} \alpha_{ji} \quad \text{for all } \mathscr{A}, \varnothing \subsetneq \mathscr{A} \subsetneq \mathscr{S}, \tag{2.8}$$

$$\max_{i\in\mathscr{S}} \alpha_i = \varrho, \tag{2.9}$$

$$\max_{\substack{i\in\mathscr{A}\\ j\notin\mathscr{A}}} \alpha_{ij} \geq 0 \quad \textit{whenever} \quad \max_{i\in\mathscr{A}} \alpha_i \geq 0 \textit{ and } \max_{j\notin\mathscr{A}} \alpha_j \geq 0. \tag{2.10}$$

PROOF. To prove (2.8), consider the successive times when the process moves from $\mathscr{A}$ to $\mathscr{S}\setminus\mathscr{A}$ and comes back: $M_0 = -1$, $N_k = \min\{n > M_k : X_n \notin \mathscr{A}\}$ and $M_{k+1} = \min\{n > N_k : X_n \in \mathscr{A}\}$. We have $M_k < N_k < M_{k+1}$ unless $N_k = \infty$. Thus,

$$\begin{aligned}
\sum_{\substack{i\in\mathscr{A}\\ j\notin\mathscr{A}}} \sum_n \varepsilon_n^h \mathbf{1}(X_n = i, X_{n+1} = j) &= \sum_n \varepsilon_n^h \mathbf{1}(X_n \in \mathscr{A}, X_{n+1} \notin \mathscr{A}) \\
&= \sum_k \varepsilon_{N_k-1}^h \leq \sum_k \varepsilon_{M_k-1}^h \\
&= \sum_n \varepsilon_n^h \mathbf{1}(X_n \notin \mathscr{A}, X_{n+1} \in \mathscr{A}) \\
&= \sum_{\substack{j\notin\mathscr{A}\\ i\in\mathscr{A}}} \sum_n \varepsilon_n^h \mathbf{1}(X_n = j, X_{n+1} = i).
\end{aligned} \tag{2.11}$$

Consequently, $\max_{i\in\mathscr{A}, j\notin\mathscr{A}} \alpha_{ij} \leq \max_{i\in\mathscr{A}, j\notin\mathscr{A}} \alpha_{ji}$, which is tantamount to (2.8). Assertion (2.9) follows from the fact that $\sum_n \varepsilon_n^h = \sum_i \sum_n \varepsilon_n^h \mathbf{1}(X_n = i)$ for all h. To prove (2.10), notice that if $\max_{i\in\mathscr{A}} \alpha_i \geq 0$ and $\max_{j\notin\mathscr{A}} \alpha_j \geq 0$, then the process visits $\mathscr{A}$ and $\mathscr{S}\setminus\mathscr{A}$ infinitely often. Therefore, it must move from $\mathscr{A}$ to $\mathscr{S}\setminus\mathscr{A}$ infinitely many times, so $\max_{i\in\mathscr{A}, j\notin\mathscr{A}} \alpha_{ij} \geq 0$. □

In general, the mean orders of recurrence do not have property (2.10). In fact, it is just for this reason that the pathwise orders are easier to handle. Consider the following example, taken from Connors and Kumar (1989).

EXAMPLE 2.1. Let $\mathscr{S} = \{1,2,3\}$. Set the transition probabilities $p_{ij}(n) = \mathbf{P}(X_{n+1} = j \mid X_n = i)$ as

$$(p_{ij}(n)) = \begin{pmatrix} 1-n^{-2} & 0 & n^{-2} \\ 0 & 1-n^{-1} & n^{-1} \\ 1-a & a & 0 \end{pmatrix},$$

where $0 < a < 1$. Condition (1.1) is satisfied for $n > 1$ with $\varepsilon_n = n^{-1}$ and

$$(v_{ij}) = \begin{pmatrix} 0 & \infty & 2 \\ \infty & 0 & 1 \\ 0 & 0 & \infty \end{pmatrix}.$$

We have $\varrho = 1$ and (2.4) holds. The mean orders of recurrence are

$$\beta_1 = 1, \qquad \beta_2 = a, \qquad \beta_3 = -\infty;$$

see Connors and Kumar [(1989), Example 4]. For each a, we obtain a different solution to the system of conditions (2.6), (2.7), (2.8) and (2.9). If we knew only the matrix (v_{ij}) and ϱ, we could not tell which of these solutions gives correct values of the mean recurrence orders. On the other hand, the pathwise orders of recurrence are

$$\alpha_1 = 1, \qquad \alpha_2 = -\infty, \qquad \alpha_3 = -\infty \quad \text{a.s.}$$

This is because these are the only assignments that satisfy (2.6), (2.7), (2.8), (2.9) *and* (2.10).

The remaining part of this section is devoted to proofs of Proposition 2.1 and Theorem 2.1. We will need the following *sub*martingale analog of a martingale theorem used by Borkar (1992).

LEMMA 2.1. *If* (S_n) *is a submartingale,* $S_n = \sum_{k \le n}(Y_k - Z_k)$ *with* $0 \le Y_n \le 1$ *and* $Z_n \ge 0$, *then*

$$\left\{\sum_n Z_n = \infty\right\} \subset \left\{\sum_n Y_n = \infty\right\} \quad a.s.$$

PROOF. Fix a and consider the stopped submartingale $(S_{\min(n,N)})$, where $N = \min\{n : S_n > a\}$. Because $S_{\min(n,N)} \le a+1$, from the classical submartingale convergence theorem [Durrett (1991), 4.2.10] we infer that $S_{\min(n,N)} \to S > -\infty$ a.s. Consequently, $\{\sup S_n \le a\} \subset \{\inf S_n > -\infty\}$ a.s. Letting $a \to \infty$, we get $\{\sup S_n < \infty\} \subset \{\inf S_n > -\infty\}$ a.s. If $\sum Z_n = \infty$ and $\inf_n \sum_{k \le n}(Y_k - Z_k) > -\infty$, then $\sum Y_n = \infty$. If $\sup_n \sum_{k \le n}(Y_k - Z_k) = \infty$, then $\sup_n \sum_{k \le n} Y_k = \infty$ and hence $\sum Y_n = \infty$. □

Under an assumption slightly stronger than our (1.1), Borkar (1992) proved the following result.

LEMMA 2.2 [Borkar (1992)]. *Almost surely, the following implications hold:*

(2.12)
$$\begin{aligned} &\text{if } \alpha_i \ge v_{ij} \text{ and } \sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) = \infty, \\ &\text{then } \alpha_{ij} = \alpha_i - v_{ij} \text{ and } \sum_n \varepsilon_n^{\alpha_{ij}} \mathbf{1}(X_n = i, X_{n+1} = j) = \infty; \end{aligned}$$

(2.13)
$$\begin{aligned} &\text{if } \alpha_i > v_{ij} \text{ and } \sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) < \infty, \\ &\text{then } \alpha_{ij} = \alpha_i - v_{ij} \text{ and } \sum_n \varepsilon_n^{\alpha_{ij}} \mathbf{1}(X_n = i, X_{n+1} = j) < \infty; \end{aligned}$$

(2.14) *if* $\alpha_i = v_{ij}$ *and* $\sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) < \infty$, *then* $\alpha_{ij} = -\infty$;

(2.15) *if* $\alpha_i < v_{ij}$, *then* $\alpha_{ij} = -\infty$.

PROOF. First we will prove that for fixed $h \geq v_{ij}$,

$$\left\{\sum_n \varepsilon_n^h \mathbf{1}(X_n = i) = \infty\right\} = \left\{\sum_n \varepsilon_n^{h-v_{ij}} \mathbf{1}(X_n = i, X_{n+1} = j) = \infty\right\} \quad \text{a.s.} \tag{2.16}$$

To begin with, note that $\{\sum \varepsilon_n^h \mathbf{1}(A_n) = \infty\} = \{\sum \varepsilon_{n+1}^h \mathbf{1}(A_n) = \infty\}$ a.s. for $h \geq 0$ and every sequence of events (A_n). To see this, write the two series as $\sum_k \varepsilon_{N_k}^h$ and $\sum_k \varepsilon_{N_k+1}^h$, where $N_0 = -1$, $N_{k+1} = \min\{n > N_k : \mathbf{1}(A_n) = 1\}$. Notice that $\varepsilon_{N_{k+1}}^h \leq \varepsilon_{N_k+1}^h \leq \varepsilon_{N_k}^h$, by (1.2). Now, the inclusion $\subset$ in (2.16) follows from Lemma 2.1 with

$$S_{n+1} = \sum_{k \leq n} [c\varepsilon_{k+1}^{h-v_{ij}} \mathbf{1}(X_k = i, X_{k+1} = j) - \varepsilon_{k+1}^h \mathbf{1}(X_k = i)].$$

This is a submartingale, because

$$\mathbf{E}(S_{n+1} - S_n \mid \mathscr{F}_n) = \varepsilon_{n+1}^{h-v_{ij}} \mathbf{1}(X_n = i)(c\mathbf{P}(X_{n+1} = j \mid \mathscr{F}_n) - \varepsilon_{n+1}^{v_{ij}}) \geq 0,$$

by the Markov property and assumption (1.1). The inclusion $\supset$ in (2.16) can be obtained in a similar way, if we consider the submartingale (S'_n) defined by

$$S'_{n+1} = \sum_{k \leq n} [c\varepsilon_k^h \mathbf{1}(X_k = i) - \varepsilon_k^{h-v_{ij}} \mathbf{1}(X_k = i, X_{k+1} = j)].$$

To infer (2.12)–(2.15) from (2.16), let us use the fact that the tail σ-field $\mathscr{T}$ consists of finitely many atoms. This fact is true for every nonhomogeneous Markov chain with a finite number of states; see Cohn (1970). From now on, we will consider a fixed atomic set T of $\mathscr{T}$. It is enough to show that implications (2.12)–(2.15) are true almost surely *on* T. The recurrence orders α_i are, of course, $\mathscr{T}$-measurable random variables, so they are a.s. constant on T. Moreover, random event $T \cap \{\alpha_i \geq 0, \sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) = \infty\}$ belongs to $\mathscr{T}$, so it must be a.s. equal either to T or to $\varnothing$. The statements to follow should be understood as holding *almost surely on* T and this phrase will not be repeated. If $\alpha_i \geq v_{ij}$ and $\sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) = \infty$, then we can apply (2.16) with $h = \alpha_i$ to get $\sum_n \varepsilon_n^{\alpha_i - v_{ij}} \mathbf{1}(X_n = i, X_{n+1} = j) = \infty$ and, consequently, $\alpha_{ij} \geq \alpha_i - v_{ij}$. To show that $\alpha_{ij} \leq \alpha_i - v_{ij}$, it is enough to use (2.16) with $h > \alpha_i$. Thus, (2.12) is proved. Proofs of (2.13)–(2.15) are similar. Note that we have restricted our considerations to a fixed atom T in order to legitimize putting $h = \alpha_i$ in (2.16) (on T, we can treat α_i as constant). $\square$

Now we are in a position to prove Proposition 2.1.

PROOF OF PROPOSITION 2.1. We will prove that (2.4) implies (2.5). Just as in the preceding proof, fix an atom T of $\mathscr{T}$. We will show that (2.6) holds for all i almost surely on T. For every i, the recurrence order α_i is a.s. constant on T. The series $\sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i)$ either converges a.s. on T or diverges a.s. on T

(provided that $\alpha_i \geq 0$ a.s. on T). From now on, we consider random variables restricted to T and we omit the phrase "a.s. on T." Define the set of states

$$\mathscr{A} = \left\{ i \in \mathscr{S} : \alpha_i \geq 0, \sum_n \varepsilon_n^{\alpha_i} \mathbf{1}(X_n = i) = \infty \right\}.$$

Under (2.4), the set $\mathscr{A}$ is nonempty because there exists some state m such that $\sum \varepsilon_n^{\varrho} \mathbf{1}(X_n = m) = \infty$ and $\alpha_m = \varrho \geq 0$. Suppose, contrary to (2.5), that there is some $j \notin \mathscr{A}$ with $\alpha_j \geq 0$. Recall Theorem 2.2. Set

$$h = \max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} \alpha_{ij} = \max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} \alpha_{ji}.$$

From (2.10) we infer that $h \geq 0$. By Lemma 2.2, $h = \max_{i \in \mathscr{A}, j \notin \mathscr{A}} (\alpha_i - v_{ij})$, because (2.12)–(2.15) show that α_{ij} is equal to $\alpha_i - v_{ij}$ whenever $\alpha_{ij} > -\infty$. Moreover, (2.12) implies that

$$\sum_n \varepsilon_n^h \mathbf{1}(X_n = i, X_{n+1} = j) = \infty \quad \text{for some } i \in \mathscr{A}, j \notin \mathscr{A}$$

(choose i, j such that $\alpha_{ij} = \alpha_i - v_{ij} = h$). On the other hand, (2.13)–(2.15) imply that

$$\sum_n \varepsilon_n^h \mathbf{1}(X_n = j, X_{n+1} = i) < \infty \quad \text{for every } j \notin \mathscr{A}, i \in \mathscr{A}.$$

Now, recall the inequality (2.11). We have just shown that its left-hand side diverges while its right-hand side converges, a contradiction. Therefore, (2.5) must hold. □

Now, Theorem 2.1 follows from Lemma 2.2 combined with Proposition 2.1.

3. Solving the propagation-balance equations. In this section, the conditions stipulated by Theorems 2.1 and 2.2 will be treated in a purely algebraic way. Let us forget about the chain (X_n) and random variables defined by (2.1) and (2.2). We regard (v_{ij}) just as a given matrix ($0 \leq v_{ij} \leq \infty$) such that for every i there is some j with $v_{ij} = 0$. In much the same way, ϱ is treated as a given number, $0 \leq \varrho < \infty$. We look at (α_i) as a system of numbers (or $-\infty$). Assume that (2.6) defines (α_{ij}). The goal is to find systems satisfying conditions (2.7), (2.8), (2.9) and (2.10). Our approach will be similar to that of Connors and Kumar (1989). Condition (2.10) makes all the difference. In general, (2.6), (2.7), (2.8) and (2.9) admit infinitely many solutions, most of them with no probabilistic meaning. If we add (2.10) to the set of conditions, only finitely many solutions will remain. We will show later (in Section 4) that all these solutions have clear interpretation in terms of the stochastic process. The *first* elegant algorithm of Connors and Kumar will be suitably modified to produce all solutions to our set of conditions. Contrary to the cited authors, we do *not* assume that the chain is irreducible. We call the chain *irreducible* if the graph of possible transitions is connected. In terms of the matrix (v_{ij}), the irreducibility assumption can be expressed as

$$i \to j \quad \text{for all } i, j \in \mathscr{S}, \tag{3.1}$$

where

$$i \to j, \quad \text{if } i = j, \text{ or there exists a sequence of states } i = i(0), i(1), \ldots, i(r) = j \text{ such that } v_{i(s-1)i(s)} < \infty \text{ for } 1 \le s \le r. \tag{3.2}$$

Under the irreducibility assumption, the basic algorithm would not be simpler (proofs of some lemmas would). We will need assumption (3.1) only in Lemmas 3.9 and 3.10 at the end of this section and later in Section 6.

To begin with, let us rewrite our conditions in an equivalent form, to stress a "local" character of solutions. If we set $\mathscr{R} = \{i : \alpha_i \ge 0\}$, then (2.6), (2.7), (2.8), (2.9) and (2.10) imply $\alpha_i = -\infty$ for $i \notin \mathscr{R}$,

$$\alpha_i \ge 0 \quad \text{if } i \in \mathscr{R}, \tag{3.3}$$

$$\max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R} \setminus \mathscr{A}}} (\alpha_i - v_{ij}) = \max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R} \setminus \mathscr{A}}} (\alpha_j - v_{ji}), \quad \text{for all } \mathscr{A}, \emptyset \subsetneq \mathscr{A} \subsetneq \mathscr{R}, \tag{3.4}$$

$$\max_{i \in \mathscr{R}} \alpha_i = \varrho, \tag{3.5}$$

$$\max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R} \setminus \mathscr{A}}} (\alpha_i - v_{ij}) \ge 0, \quad \text{for all } \mathscr{A}, \emptyset \subsetneq \mathscr{A} \subsetneq \mathscr{R}, \tag{3.6}$$

$$\max_{\substack{i \in \mathscr{R} \\ j \notin \mathscr{R}}} (\alpha_i - v_{ij}) < 0. \tag{3.7}$$

Indeed, (3.3) and (3.5) are obvious. To get (3.7), we argue as follows. For all $j \notin \mathscr{R}$, we have $\alpha_{ji} = -\infty$, by (2.6) and (2.7). Therefore, $\alpha_{ij} = -\infty$ for all $i \in \mathscr{R}$, $j \notin \mathscr{R}$, by (2.8) (with $\mathscr{A} = \mathscr{R}$). In view of (2.6), $\alpha_{ij} = -\infty$ means just that $\alpha_i - v_{ij} < 0$. Therefore, (3.7) holds. Now, if we take (3.7) into account, we can see that (3.4) is a consequence of (2.8), whereas (3.6) is a consequence of (2.10).

The advantage of (3.3)–(3.7) is that these conditions involve only $(\alpha_i, i \in \mathscr{R})$. Conversely, suppose (3.3)–(3.7) hold for a system $(\alpha_i, i \in \mathscr{R})$. If we put $\alpha_i = -\infty$ for $i \notin \mathscr{R}$, the extended system $(\alpha_i, i \in \mathscr{S})$ will satisfy (2.7), (2.8), (2.9) and (2.10), provided that (α_{ij}) are given by (2.6).

In the sequel, we will have to work with systems that need not satisfy (3.7), but satisfy the following weaker condition

$$\max_{\substack{i \in \mathscr{R} \\ j \notin \mathscr{R}}} (\alpha_i - v_{ij}) \le \max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R} \setminus \mathscr{A}}} (\alpha_i - v_{ij}) \quad \text{for all } \mathscr{A}, \emptyset \subsetneq \mathscr{A} \subsetneq \mathscr{R}. \tag{3.8}$$

By definition, a system $(\alpha_i, i \in \mathscr{R})$ is a *coalition* if (3.3), (3.4), (3.6) and (3.8) hold. Say $\mathscr{R}$ is the *domain* of the coalition and $\max_{i\in\mathscr{R}} \alpha_i$ is its *height*. The set $\mathscr{M} = \{j \in \mathscr{R} : \alpha_j = \max_{i\in\mathscr{R}} \alpha_i\}$ will be called the *top* of the coalition. The coalition is said to be *closed* if (3.7) holds, too. Thus, the problem is to find all closed coalitions of height ϱ. Notice that if $\mathscr{R}$ is a set of one element, then (3.4), (3.6) and (3.8) are vacuously satisfied. Thus, in particular, all singleton systems $(\alpha_i = \varrho, \{i\})$ are coalitions of height ϱ.

It will be helpful to express conditions (3.3)–(3.8) in terms of a reachability relation. Given a system $(\alpha_i, i \in \mathscr{R})$ (not necessarily a coalition), write for $i \in \mathscr{R}$ and $j \in \mathscr{S}$,

(3.9) $i \to_h j$, if $i = j$ and $\alpha_i \geq h$, or there exists a sequence of states $i = i(0), i(1), \dots, i(r) = j$ such that $i(s-1) \in \mathscr{R}$ and $\alpha_{i(s-1)} - v_{i(s-1)i(s)} \geq h$ for $1 \leq s \leq r$.

If there is a danger of ambiguity, we will write explicitly "$i \to_h j$ with respect to (w.r.t.) $(\alpha_i, i \in \mathscr{R})$." If $i \to_h j$, we say j is *reachable at height* h from i. We follow Hajek (1988) in using this term. His definition and ours will be compared in Section 6. A sequence of states with the property described in (3.9) will be called a path at height h. Note that all states of a path but its end, j, must belong to $\mathscr{R}$.

LEMMA 3.1. *System* $(\alpha_i, i \in \mathscr{R})$ *satisfies* (3.4) *if and only if the relation defined by* (3.9) *with respect to this system has the property:*

(3.10) *for every* h *and for all* $i, j \in \mathscr{R}, i \to_h j$ *implies* $j \to_h i$.

Equivalently,

(3.11) *for all* $i, j \in \mathscr{R}, h = \alpha_i - v_{ij} > -\infty$ *implies* $j \to_h i$.

PROOF. The fact that (3.10) is equivalent to (3.11) is obvious. Suppose (3.11) is true. Let $\emptyset \subsetneq \mathscr{A} \subsetneq \mathscr{R}$ and $h = \max_{i\in\mathscr{A}, j\in\mathscr{R}\setminus\mathscr{A}}(\alpha_i - v_{ij})$. We will show that $\max_{i\in\mathscr{A}, j\in\mathscr{R}\setminus\mathscr{A}}(\alpha_j - v_{ji}) \geq h$, which is tantamount to (3.4). If $h = -\infty$, there is nothing to prove. If $h > -\infty$, choose $i \in \mathscr{A}$ and $j \in \mathscr{R}\setminus\mathscr{A}$ such that $h = \alpha_i - v_{ij}$ and consider a path from j to i at height h. Select the first entry this path into $\mathscr{A}$, to get $k \in \mathscr{R}\setminus\mathscr{A}$, $l \in \mathscr{A}$ such that $\alpha_k - v_{kl} \geq h$. Conversely, to deduce (3.11) from (3.4), fix i, $j \in \mathscr{R}$ and write $h = \alpha_i - v_{ij}$. Consider the set $\mathscr{A} = \{j\} \cup \{k \in \mathscr{R} : j \to_h k\}$. We have $\max_{k\in\mathscr{A}, l\in\mathscr{R}\setminus\mathscr{A}}(\alpha_k - v_{kl}) < h$, by definition of $\mathscr{A}$. If $i \in \mathscr{R}\setminus\mathscr{A}$, this would contradict (3.4). Thus, $i \in \mathscr{A}$. □

LEMMA 3.2. *System* $(\alpha_i, i \in \mathscr{R})$ *satisfies* (3.3) *and* (3.6) *if and only if the relation defined by* (3.9) *with respect to this system has the property:*

(3.12) $i \to_0 j$ *for all* $i, j \in \mathscr{R}$.

Condition (3.7) *holds iff*

(3.13) $i \not\to_0 k$ *for all* $i \in \mathscr{R}$ *and* $k \notin \mathscr{R}$.

Condition (3.8) *holds iff*

(3.14) *for every* h, $i \to_h k$, $i \in \mathscr{R}$ *and* $k \notin \mathscr{R}$ *imply* $i \to_h j$ *for all* $j \in \mathscr{R}$.

PROOF. Obvious. □

Conditions (3.10)–(3.14) are sometimes much easier to check than (3.4) and (3.6)–(3.8).

LEMMA 3.3. *Suppose* $(\alpha_i, i \in \mathscr{R})$ *and* $(\alpha'_i, i \in \mathscr{R}')$ *are coalitions. If* $\mathscr{R} \cap \mathscr{R}' \neq \emptyset$, *then there is* $c \in \mathbf{R}$ *such that* $\alpha'_i = \alpha_i + c$ *for all* $i \in \mathscr{R} \cap \mathscr{R}'$.

PROOF. Choose constants b and b' such that $\alpha_i + b = \alpha'_i + b'$ for some $i \in \mathscr{R} \cap \mathscr{R}'$. By symmetry, it is enough to show that $\alpha_i + b \leq \alpha'_i + b'$ for all $i \in \mathscr{R} \cap \mathscr{R}'$. Let

$$\mathscr{A} = \{i \in \mathscr{R} \cap \mathscr{R}' : \alpha_i + b > \alpha'_i + b'\}, \qquad \mathscr{A}' = \{i \in \mathscr{R} \cap \mathscr{R}' : \alpha_i + b \leq \alpha'_i + b'\}.$$

We know that $\mathscr{A}' \neq \emptyset$ and we claim that $\mathscr{A} = \emptyset$. Suppose the contrary. Then,

$$\begin{aligned}
&\max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} (\alpha'_i + b' - v_{ij}) \\
&\qquad = \max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R}' \setminus \mathscr{A}}} (\alpha'_i + b' - v_{ij}) \quad \text{by (3.8) for } (\alpha'_i, i \in \mathscr{R}') \\
&\qquad < \max_{\substack{i \in \mathscr{A} \\ j \in \mathscr{R}' \setminus \mathscr{A}}} (\alpha_i + b - v_{ij}) \quad \text{by definition of } \mathscr{A} \\
&\qquad \leq \max_{\substack{i \in \mathscr{R} \setminus \mathscr{A}' \\ j \in \mathscr{R}' \setminus \mathscr{A}}} (\alpha_i + b - v_{ij}) \\
&\qquad = \max_{\substack{i \in \mathscr{R} \setminus \mathscr{A}' \\ j \in \mathscr{A}'}} (\alpha_i + b - v_{ij}) \quad \text{by (3.8) for } (\alpha_i, i \in \mathscr{R}) \\
&\qquad = \max_{\substack{i \in \mathscr{R} \setminus \mathscr{A}' \\ j \in \mathscr{A}'}} (\alpha_j + b - v_{ji}) \quad \text{by (3.4) for } (\alpha_i, i \in \mathscr{R}) \\
&\qquad \leq \max_{\substack{i \notin \mathscr{A}' \\ j \in \mathscr{A}'}} (\alpha_j + b - v_{ji}) \\
&\qquad\quad \vdots \\
&\qquad \leq \max_{\substack{i \in \mathscr{A} \\ j \notin \mathscr{A}}} (\alpha'_i + b' - v_{ij}) \quad \text{through a similar chain of inequalities.}
\end{aligned}$$

This is impossible and the proof is complete. Note that we have tacitly used (3.6); the foregoing strict inequality is justified in view of the fact that its left-hand side is finite. This fact follows from (3.6), because $\mathscr{A} \neq \emptyset$ and $\mathscr{R}' \setminus \mathscr{A} \neq \emptyset$, by assumption. □

LEMMA 3.4. *If $(\alpha_i, i \in \mathscr{R})$ and $(\alpha'_i, i \in \mathscr{R})$ are coalitions (with the same domain $\mathscr{R}$) of the same height ϱ, then $\alpha'_i = \alpha_i$ for all $i \in \mathscr{R}$.*

PROOF. Obvious. □

LEMMA 3.5. *Suppose $(\alpha_i, i \in \mathscr{R})$ is a closed coalition of height ϱ and $(\alpha'_i, i \in \mathscr{R}')$ is a coalition of the same height ϱ. Then either $\mathscr{R} \cap \mathscr{R}' = \emptyset$ or $\mathscr{R}' \subseteq \mathscr{R}$. If $\mathscr{R}' \subseteq \mathscr{R}$, then there is $c \geq 0$ such that $\alpha'_i = \alpha_i + c$ for all $i \in \mathscr{R}'$.*

PROOF. Suppose $\mathscr{R} \cap \mathscr{R}' \neq \emptyset$. By Lemma 3.3, we can choose constants b and b' such that $\alpha_i + b = \alpha'_i + b'$ for all $i \in \mathscr{R} \cap \mathscr{R}'$. If $\mathscr{R}' \setminus \mathscr{R}$ and $\mathscr{R} \setminus \mathscr{R}'$ were nonempty, we would have

$$
\begin{aligned}
&\max_{\substack{i \in \mathscr{R} \cap \mathscr{R}' \\ j \in \mathscr{R} \setminus \mathscr{R}'}} (\alpha'_i + b' - v_{ij}) \\
&\quad \leq \max_{\substack{i \in \mathscr{R} \cap \mathscr{R}' \\ j \in \mathscr{R}' \setminus \mathscr{R}}} (\alpha'_i + b' - v_{ij}) \quad \text{by (3.8) for } (\alpha'_i, i \in \mathscr{R}') \\
&\quad = \max_{\substack{i \in \mathscr{R} \cap \mathscr{R}' \\ j \in \mathscr{R}' \setminus \mathscr{R}}} (\alpha_i + b - v_{ij}) \\
&\quad < \max_{\substack{i \in \mathscr{R} \cap \mathscr{R}' \\ j \in \mathscr{R} \setminus \mathscr{R}'}} (\alpha_i + b - v_{ij}) \quad \text{by (3.6) and (3.7) for } (\alpha_i, i \in \mathscr{R}),
\end{aligned}
$$

which is impossible. Thus, $\mathscr{R}' \subseteq \mathscr{R}$ or $\mathscr{R} \subseteq \mathscr{R}'$. If $\mathscr{R}' \subseteq \mathscr{R}$, then

$$
\varrho + b' = \max_{i \in \mathscr{R}'} (\alpha'_i + b') = \max_{i \in \mathscr{R}'} (\alpha_i + b) \leq \varrho + b,
$$

so $b' \leq b$ and we can set $c = b - b'$ to conclude. If $\mathscr{R} \subseteq \mathscr{R}'$, then $b \leq b'$, by a similar argument. From this, we can infer that $\mathscr{R} = \mathscr{R}'$. Indeed, $\mathscr{R}' \setminus \mathscr{R} \neq \emptyset$ would imply

$$
b > \max_{\substack{i \in \mathscr{R} \\ j \in \mathscr{R}' \setminus \mathscr{R}}} (\alpha_i + b - v_{ij}) \geq \max_{\substack{i \in \mathscr{R} \\ j \in \mathscr{R}' \setminus \mathscr{R}}} (\alpha'_i + b' - v_{ij}) \geq b',
$$

by (3.7) for $(\alpha_i, i \in \mathscr{R})$ and (3.6) for $(\alpha'_i, i \in \mathscr{R}')$. □

LEMMA 3.6. *If $(\alpha_i, i \in \mathscr{R})$ and $(\alpha'_i, i \in \mathscr{R}')$ are closed coalitions of the same height ϱ, then either $\mathscr{R} \cap \mathscr{R}' = \emptyset$ or $\mathscr{R} = \mathscr{R}'$ and $\alpha_i = \alpha'_i$ for all i.*

PROOF. Obvious. □

The following two lemmas will directly lead to the algorithm for computing all closed coalitions. Given a system $(\alpha_i, i \in \mathscr{R})$ and sets $\mathscr{A} \subseteq \mathscr{R}$, $\mathscr{B} \subseteq \mathscr{S}$, we will write

$$\mathscr{A} \rightarrow_h \mathscr{B} \quad \text{if } i \rightarrow_h j \text{ for some } i \in \mathscr{A}, j \in \mathscr{B},$$

with respect to $(\alpha_i, i \in \mathscr{R})$. Recall that relation $i \rightarrow_h j$ was defined for j not necessarily in $\mathscr{R}$.

LEMMA 3.7. *Let $(\alpha_i, i \in \mathscr{R}_1), \dots, (\alpha_i, i \in \mathscr{R}_m)$ be coalitions of height ϱ such that their domains $\mathscr{R}_1, \dots, \mathscr{R}_m$ are disjoint. Set $h_k = \max_{i \in \mathscr{R}_k, j \notin \mathscr{R}_k}(\alpha_i - v_{ij}) \geq 0$, $k = 1, \dots, m$, and suppose*

$$\mathscr{R}_1 \rightarrow_{h_1} \mathscr{R}_2 \rightarrow_{h_2} \cdots \mathscr{R}_m \rightarrow_{h_m} \mathscr{R}_1,$$

where the kth arrow is understood with respect to $(\alpha_i, i \in \mathscr{R}_k)$. If we let $\tilde{\mathscr{R}} = \mathscr{R}_1 \cup \cdots \cup \mathscr{R}_m$ and

$$\tilde{\alpha}_i = \alpha_i - (h_k - \tilde{h}) \quad \text{for } i \in \mathscr{R}_k,$$

where $\tilde{h} = \min(h_1, \dots, h_m)$, then the system $(\tilde{\alpha}_i, i \in \tilde{\mathscr{R}})$ is a coalition of height ϱ.

PROOF. Let us recall the conditions appearing in the definition of coalition and invoke Lemmas 3.1 and 3.2. We are going to check that conditions (3.10), (3.12), (3.14) and (3.5) hold for $(\tilde{\alpha}_i, i \in \tilde{\mathscr{R}})$. Rewrite the sequence of arrows as

$$\mathscr{R}_1 \rightarrow_{\tilde{h}} \mathscr{R}_2 \rightarrow_{\tilde{h}} \cdots \mathscr{R}_m \rightarrow_{\tilde{h}} \mathscr{R}_1,$$

all arrows with respect to $(\tilde{\alpha}_i, i \in \tilde{\mathscr{R}})$. If $i \in \mathscr{R}_k$ and $j \in \mathscr{R}_{k+1}$, then $i \rightarrow_{\tilde{h}} j$ with respect to $(\tilde{\alpha}_i)$, because $(\alpha_i, i \in \mathscr{R}_k)$ has property (3.14). The same applies to $i \in \mathscr{R}_m$ and $j \in \mathscr{R}_1$. Thus, $i \rightarrow_{\tilde{h}} j$ for all $i, j \in \tilde{\mathscr{R}}$. On the other hand, $i \not\rightarrow_h j$ with respect to $(\tilde{\alpha}_i)$ for $h > \tilde{h}$, whenever $i \in \mathscr{R}_k$ and $j \notin \mathscr{R}_k$, by definition of h_k. This shows that (3.14) is true for $(\tilde{\alpha}_i)$. To verify that (3.10) holds for $(\tilde{\alpha}_i, i \in \tilde{\mathscr{R}})$, consider two cases. If $i \in \mathscr{R}_k$ and $j \in \mathscr{R}_l$, $k \neq l$, then $i \rightarrow_{\tilde{h}} j$, $j \rightarrow_{\tilde{h}} i$ and $i \not\rightarrow_h j$, $j \not\rightarrow_h i$ for every $h > \tilde{h}$. If i and j belong to the same $\mathscr{R}_k$ for some k, use the fact that $(\alpha_i, i \in \mathscr{R}_k)$ satisfies (3.10). Properties (3.12) and (3.5) are ensured by the choice of $\tilde{h}$. □

LEMMA 3.8. *Let $(\alpha_i, i \in \mathscr{R}_1), \dots, (\alpha_i, i \in \mathscr{R}_m), (\alpha_i, i \in \mathscr{R}_{m+1})$ be coalitions of height ϱ such that their domains, $\mathscr{R}_1, \dots, \mathscr{R}_{m+1}$, are disjoint. Set $h_k = \max_{i \in \mathscr{R}_k, j \notin \mathscr{R}_k}(\alpha_i - v_{ij}) \geq 0$, $k = 1, \dots, m$, and suppose*

$$\mathscr{R}_1 \rightarrow_{h_1} \mathscr{R}_2 \rightarrow_{h_2} \cdots \mathscr{R}_m \rightarrow_{h_m} \mathscr{R}_{m+1},$$

where the kth arrow is understood with respect to $(\alpha_i, i \in \mathscr{R}_k)$. If $(\alpha_i, i \in \mathscr{R}_{m+1})$ is closed, then $\mathscr{R}_1 \cup \cdots \cup \mathscr{R}_m$ is disjoint from domains of all closed coalitions of height ϱ.

PROOF. Suppose the assertion of the lemma is false. Select the greatest k, $1 \leq k \leq m$, such that $\mathscr{R}_k$ is not disjoint from domains of all closed coalitions of height ϱ. By Lemma 3.5, there is a closed coalition $(\alpha'_i, i \in \mathscr{R}')$ of height ϱ such that $\mathscr{R}_k \subseteq \mathscr{R}'$. We claim that $\mathscr{R}' \backslash \mathscr{R}_k \neq \emptyset$. Indeed, the coalition with domain $\mathscr{R}_k$ is not closed, because $h_k \geq 0$. It is impossible that $\mathscr{R}_k = \mathscr{R}'$, because this would imply the corresponding two coalitions coincide, by Lemma 3.4. Now, Lemma 3.5 gives $\alpha_i = \alpha'_i + c$ for all $i \in \mathscr{R}_k$. We have $h_k = \max_{i \in \mathscr{R}_k, j \notin \mathscr{R}_k}(\alpha'_i + c - v_{ij}) \geq c > \max_{i \in \mathscr{R}_k, j \notin \mathscr{R}'}(\alpha'_i + c - v_{ij})$, because $(\alpha'_i, i \in \mathscr{R}')$ satisfies (3.6) and (3.7). This contradicts the assumption $\mathscr{R}_k \to_{h_k} \mathscr{R}_{k+1}$. Indeed, $\mathscr{R}_{k+1} \subseteq \mathscr{S} \backslash \mathscr{R}'$. If $k < m$, the last inclusion follows from Lemma 3.5 in view of the choice of k. If $k = m$, it follows from Lemma 3.6, due to the fact that $(\alpha_i, i \in \mathscr{R}_{m+1})$ is closed. □

Now we are in a position to describe the algorithm. We will need some more definitions and notational conventions. All coalitions produced in the course of computations will have the same height ϱ. Therefore, Lemma 3.4 allows us to abbreviate "coalition $(\alpha_i, i \in \mathscr{R})$" to "coalition $\mathscr{R}$." Say a coalition is *open* (at height ϱ) if it is disjoint from all closed coalitions of height ϱ. The fact that a given coalition is closed depends on the coalition itself and can be verified if falsified directly. On the other hand, the fact that a given coalition is open depends on its "environment." Each of the coalitions to be dealt with in the algorithm will be assigned to one of the following three categories: closed coalitions, coalitions known to be open and nonclosed coalitions that are not known to be open. Coalitions of the third category will be called *unlabeled*. An unlabeled coalition may be open or may be a proper subset of a closed coalition (Lemma 3.5). The idea behind the algorithm is simple. We aggregate disjoint coalitions until we get coalitions that are either closed or open.

ALGORITHM

Input. (v_{ij}) and ϱ.

Output. The family of all closed coalitions of height ϱ.

Start. Begin with a family of all singleton systems $(\alpha_i = \varrho, \{i\})$. They are coalitions, of course. Find out which of them are closed.

Step. Consider a family (indexed by t) of coalitions $(\alpha_i, i \in \mathscr{R}_t)$ of height ϱ, with disjoint domains $\mathscr{R}_t$ such that $\bigcup_t \mathscr{R}_t = \mathscr{S}$. Select a nonclosed unlabeled coalition, say $\mathscr{R}_{t_1}$. Set $h_1 = \max_{i \in \mathscr{R}_{t_1}, j \notin \mathscr{R}_{t_1}}(\alpha_i - v_{ij}) \geq 0$ and find another coalition $\mathscr{R}_{t_2}$ such that $\mathscr{R}_{t_1} \to_{h_1} \mathscr{R}_{t_2}$. If $\mathscr{R}_{t_2}$ is unlabeled, repeat the same procedure with $\mathscr{R}_{t_2}$ and so on. The resulting sequence

$$\mathscr{R}_{t_1} \to_{h_1} \mathscr{R}_{t_2} \to_{h_2} \cdots$$

either contains a cycle or ends at a labeled coalition (closed or open).

- If a cycle obtains, apply Lemma 3.7. Replace those coalitions that form the cycle with one larger coalition. Check whether the new coalition is closed or not. If it is not closed, leave it unlabeled.
- If a labeled coalition appears at the end of the sequence of arrows, label all its predecessors in this sequence as open coalitions.

Repeat.

Stop. If every coalition of current generation is either closed or labeled as open, stop.

THEOREM 3.1. *In a finite number of steps, the algorithm produces all closed coalitions of height ϱ.*

PROOF. Each step decreases the number of coalitions that are unlabeled, so the algorithm must stop. Lemma 3.7 shows that the procedure of producing new coalitions is correct. Notice that once a coalition has been labeled, it will not be altered by the algorithm. Thus, Lemma 3.8 shows by induction that the procedure of labeling open coalitions is correct. The very form of the stop criterion guarantees the desired result. □

The rest of this section contains results that will not be used until Section 6. The algorithm consisted of building larger coalitions from smaller ones. Now, let us try the other way around and sketch another approach. Begin with a large coalition of big height τ and look for its subsets that are closed coalitions of height $\varrho < \tau$. This makes sense if a large coalition is known. We will encounter such a situation in Section 6. Now, recall a result of Connors and Kumar (1989).

LEMMA 3.9 (Connors and Kumar). *If* (3.1) *holds, then for any sufficiently big $\tau > 0$ there exists a coalition $(\lambda_i, i \in \mathscr{S})$, with the domain equal to the whole space of height τ.*

PROOF. In consecutive steps of the algorithm, whenever a closed coalition with a domain smaller than $\mathscr{S}$ appears, add a sufficiently big constant to all α_i's. The coalition will cease to be closed. □

Given τ, the system (λ_i) spoken of in Lemma 3.9 is unique, by Lemma 3.4. Connors and Kumar called (λ_i) the solution to the "modified balance equations." If the chain is irreducible, all closed coalitions of height ϱ can be identified in terms of the system (λ_i).

LEMMA 3.10. *Suppose $(\lambda_i, i \in \mathscr{S})$ is a coalition of height τ and $\varrho < \tau$. If $(\alpha_i, i \in \mathscr{R})$ is a closed coalition of height ϱ and $m \in \mathscr{R}$ is such that $\alpha_m = \varrho$, then*

(3.15) *for all* $j \in \mathscr{S}, m \to_{\lambda_m - \varrho} j$ *w.r.t.* (λ_i) *implies* $\lambda_j \leq \lambda_m$,

$$\begin{aligned} &\mathscr{R} = \{ j \in \mathscr{S} : m \to_{\lambda_m - \varrho} j \textit{ w.r.t. } (\lambda_i) \} \\ &\textit{and for all } j \in \mathscr{R}, \qquad \alpha_j = \lambda_j - \lambda_m + \varrho. \end{aligned} \tag{3.16}$$

Conversely, if m *satisfies* (3.15) *and the system* $(\alpha_i, i \in \mathscr{R})$ *is defined by* (3.16), *then it is a closed coalition of height* ϱ.

PROOF. If $(\alpha_i, i \in \mathscr{R})$ is a coalition, then Lemma 3.5 shows that $\alpha_i = \lambda_i - c$, all $i \in \mathscr{R}$. Fix m such that $\alpha_m = \varrho$. We obtain $c = \lambda_m - \varrho$. If $i \in \mathscr{R}$ and $\lambda_i - v_{ij} \geq \lambda_m - \varrho$, then $\alpha_i - v_{ij} \geq 0$ and $j \in \mathscr{R}$, by (3.13). Thus, $m \to_{\lambda_m - \varrho} j$ w.r.t. (λ_i) implies $j \in \mathscr{R}$ and $\alpha_j \leq \alpha_m$, so $\lambda_j \leq \lambda_m$. If $j \in \mathscr{R}$, then $m \to_0 j$ w.r.t (α_i) by (3.12) and consequently, $m \to_{\lambda_m - \varrho} j$ w.r.t. (λ_i). We have verified (3.15) and (3.16). The converse part is easy. □

Property (3.15) characterizes states belonging to tops of closed coalitions of height ϱ. Note that if m satisfies (3.15), $m \to_{\lambda_m - \varrho} m'$ and $\lambda_m = \lambda_{m'}$, then m' also satisfies (3.15) and both the states belong to the top of one closed coalition.

We think Lemma 3.10 has some intuitive appeal. We do not think an algorithm based on it would be efficient, even if we take (λ_i) for granted.

4. Recurrence and transience. Now, return to our Markov chain (X_n) and its asymptotics. We are going to exploit algebraic results of the previous section and to show their probabilistic meaning. Let (v_{ij}) and (ε_n) be the matrix and the sequence appearing in (1.1). Let ϱ be the order of cooling of (ε_n), given by (2.3). Assume (2.4) holds. This assumption is not essential and it is made merely to simplify notation. The case (2.4′) could be treated in a similar way and the results to follow would need only minor modifications.

THEOREM 4.1. (a) *There is a decomposition* $\mathscr{S} = \bigcup_t \mathscr{R}_t \cup \mathscr{C}$ *of the state space into disjoint sets and a corresponding decomposition* $\Omega = \bigcup_t U_t$ *of the probability space into disjoint events such that*

$$U_t = \{X_n \in \mathscr{R}_t \quad i.o.\} = \{X_n \in \mathscr{R}_t \quad ult.\} \qquad a.s.,$$

$$\{X_n \in \mathscr{C} \quad i.o.\} = \varnothing \qquad a.s.$$

Moreover,

$$U_t = \{X_n = i \quad i.o.\} \qquad a.s. \text{ for every } i \in \mathscr{R}_t.$$

(b) *On each of the events* U_t, *the recurrence orders defined by* (2.1) *are a.s. constant. Their a.s. values, say* $\alpha_i(U_t)$, *are determined by the fact that*

$$(\alpha_i(U_t), i \in \mathscr{R}_t)$$

is one of the closed coalitions of height ϱ, *produced by the algorithm. In particular, the domain and the top of the coalition can be expressed as*

$$\mathscr{R}_t = \{i : \alpha_i(U_t) \geq 0\}, \qquad \mathscr{M}_t = \{i : \alpha_i(U_t) = \varrho\}.$$

PROOF. For almost every $\omega \in \Omega$, the systems $(\alpha_i(\omega), i \in \mathscr{S})$ and $(\alpha_{ij}(\omega), i, j \in \mathscr{S})$ satisfy (2.6), (2.7), (2.8), (2.9) and (2.10), by Theorems 2.1 and 2.2. We proved in Section 3 that there are finitely many, say s, distinct solutions to this set of conditions. Each solution is a closed coalition of height ϱ (extended by setting $-\infty$ outside its domain). All the solutions are produced by the Algorithm. Now, set $U_t = \{\omega \in \Omega : (\alpha_i(\omega))$ is the tth solution $\}$ for $1 \le t \le s$. Simply by construction, the values of recurrence orders are constant on these events and form closed coalitions of height ϱ. The domains of the coalitions under consideration are disjoint by Lemma 3.6. Now note that $\{\omega : X_n(\omega) = i \quad \text{i.o.}\} = \{\omega : \alpha_i(\omega) \ge 0\}$ to conclude the proof. □

Note that we have defined as many events U_t as there are closed coalitions of height ϱ. Without some assumptions on the initial distribution of the chain, we cannot claim that all these events have positive probability. The following proposition will clarify this point.

PROPOSITION 4.1. *Suppose $\mathscr{R}_t$ is the domain of one of closed coalitions of height ϱ. If $\mathbf{P}(X_0 \in \mathscr{R}_t) > 0$, then*

$$\mathbf{P}(X_n \in \mathscr{R}_t \textit{ for all } n \ge 0) > 0$$

and, consequently, $\mathbf{P}(U_t) > 0$.

The proof involves notion of periodicity, so let us defer it. In the next section we will prove a stronger result—Proposition 5.1.

Theorem 4.1(a) and Proposition 4.1 allow us to say the domains $\mathscr{R}_t$ of closed coalitions of height ϱ are *recurrent classes* and $\mathscr{E} = \mathscr{S} \setminus \bigcup_t \mathscr{R}_t$ is the *set of transient states*. Loosely speaking, the process is a.s. eventually attracted to one of the recurrent classes and visits all its states infinitely often. Note that the statement of Theorem 4.1(a) makes no reference to the recurrence orders, yet its proof heavily depends on them. To explain the role played by our standing assumptions (1.1) and (1.2), we will give an example of Markov chain of Doeblin's type (B), for which the conclusion of Theorem 4.1(a) is *false*.

EXAMPLE 4.1. Let $\mathscr{S} = \{0, 1, 2, 3\}$ and let the transition probabilities $\mathbf{P}(X_{n+1} = j | X_n = i) = p_{ij}(n)$ be

$$p_{10}(n) = p_{20}(n) = 1, \qquad p_{30}(n) = 1 - n^{-2},$$

$$p_{33}(n) = n^{-2}, \qquad p_{03}(n) = 1 - n^{-1},$$

$$p_{01}(n) = \begin{cases} 0, & \text{if } n \text{ is even,} \\ n^{-1}, & \text{if n is odd,} \end{cases} \qquad p_{02}(n) = \begin{cases} n^{-1}, & \text{if } n \text{ is even,} \\ 0, & \text{if } n \text{ is odd.} \end{cases}$$

This is clearly not a chain with regularly diminishing transitions. It is easy to see that the tail σ-field of this chain has two atoms:

$$A_1 = \{X_{2n+1} = 0 \quad \text{i.o.}\} = \{X_n = 1 \quad \text{i.o.}\},$$

$$A_2 = \{X_{2n} = 0 \quad \text{i.o.}\} = \{X_n = 2 \quad \text{i.o.}\}.$$

If $\mathscr{S}_1 = \{0,1,3\}$ and $\mathscr{S}_2 = \{0,2,3\}$, then $A_1 = \{X_n \in \mathscr{S}_1 \text{ ult.}\}$ and $A_2 = \{X_n \in \mathscr{S}_2 \text{ ult.}\}$. However, the sets $\mathscr{S}_1$ and $\mathscr{S}_2$ are not disjoint.

The preceding example sheds some light on the role of Lemma 3.6 in the proof of Theorem 4.1. In fact, this lemma is the only result of Section 3 needed to prove that there *exists* a decomposition of $\mathscr{S}$ into recurrent classes $\mathscr{R}_t$ and the transient set $\mathscr{C}$. The remaining part of Section 3 shows how to *identify* the decomposition in algebraic terms.

5. Atoms of the tail sigma field. It will turn out that chains with regularly diminishing transitions display essentially no more patterns of asymptotic behavior than homogeneous ones. The events U_t defined in the preceding section are not necessarily atoms of the tail σ-field, yet their decomposition into atoms is surprisingly simple. Our main tool will be the method of *two particles* due to Doeblin, in conjunction with our theorems concerning the recurrent classes. For our purposes, the Doeblin's trick can be summarized as follows. Recall that the chain (X_n) is defined on $(\Omega, \mathscr{F}, \mathbf{P})$. Let (X_n^1) and (X_n^2) be two copies of (X_n). Let $\underline{X}_n = (X_n^1, X_n^2)$ be defined on $\Omega \times \Omega$ by $\underline{X}_n(\omega_1, \omega_2) = (X_n^1(\omega_1), X_n^2(\omega_2))$. Equip $\Omega \times \Omega$ with the product measure $\mathbf{P} \times \mathbf{P}$ to get the *double* chain, consisting of two *independent* copies of the original one. The following lemma is valid for arbitrary Markov chains. It is probably well known, yet we will give the proof for completeness.

LEMMA 5.1. *Assume $A \in \mathscr{T}$ and $\mathbf{P}(A) > 0$. If for some $i \in \mathscr{S}$ we have*

$$A \times A \subset \{\underline{X}_n = (X_n^1, X_n^2) = (i,i) \quad \text{i.o.}\} \ \text{a.s.} \quad [\mathbf{P} \times \mathbf{P}], \tag{5.1}$$

then A is an atom of the tail σ-field $\mathscr{T}$.

PROOF. Suppose, contrary to our claim, that there exists $B \in \mathscr{T}$, with $B \subset A$, such that $\mathbf{P}(B) > 0$ and $\mathbf{P}(A \backslash B) > 0$. Let $\mathscr{F}_\infty = \sigma(X_n, n \geq 0)$. By the Markov property and the Levy's 0–1 law [see Durrett (1991), 4.5.6],

$$\mathbf{P}(B \mid X_n) = \mathbf{P}(B \mid \mathscr{F}_n) \to \mathbf{P}(B \mid \mathscr{F}_\infty) = \mathbf{1}(B) \quad \text{a.s.} \tag{5.2}$$

For clarity, write the two copies of Ω as Ω^1 and Ω^2. Now $\underline{X}_n = (X_n^1, X_n^2)$ lives on $\Omega^1 \times \Omega^2$. For $j = 1,2$, let B^j and A^j be the copies of B and A, respectively, on Ω^j. By (5.1) and (5.2) we may find $\omega_1 \in B^1$ and $\omega_2 \in A^2 \backslash B^2$ such that

$$\underline{X}_n(\omega_1, \omega_2) = (i,i) \quad \text{i.o.} \tag{5.3}$$

$$\mathbf{P}(B^1 | X_n^1)(\omega_1) \to 1 \quad \text{as } n \to \infty,$$

and

$$\mathbf{P}(B^2 | X_n^2)(\omega_2) \to 0 \quad \text{as } n \to \infty.$$

Let (n_k) be the sequence of moments when $\underline{X}_n(\omega_1, \omega_2)$ visits (i,i), that is $n_0 = -1$, $n_{k+1} = \min\{n > n_k : X_n^1(\omega_1) = i = X_n^2(\omega_2)\}$. This is an infinite increasing sequence of integers, in view of (5.3). Because for $j = 1,2$ and all k,

$$\mathbf{P}(B^j \mid X^j_{n_k})(\omega_j) = \mathbf{P}(B^j \mid X^j_{n_k} = i) = \mathbf{P}(B \mid X_{n_k} = i),$$

we have contradictory convergence statements for $\mathbf{P}(B \mid X_{n_k} = i)$. □

Now, return to the chain (X_n) with regularly diminishing transitions and invoke Theorem 4.1. Of course, we can restrict our description of atoms of $\mathscr{T}$ to those contained in one of the events U_t. Fix t, write U for U_t and $\mathscr{R}$ for the corresponding recurrent class $\mathscr{R}_t$. Let $(\alpha_i, i \in \mathscr{R})$ be the corresponding closed coalition, that is, α_i's are values of the recurrence orders on U. Recall the reachability relation (3.9) introduced in Section 3. Now we have to keep track of the lengths of underlying paths. If there is a sequence $i = i(0), i(1), \dots, i(r) = j$ (consisting of *precisely* $r+1$ states), which is a path at height h leading from i to j, we will write $i \to_h j$ (r steps). Write $i \to_h i$ (0 steps) if $\alpha_i \geq h$. The system $(\alpha_i, i \in \mathscr{R})$ is fixed and notation $i \to_h j$ will refer to it whenever $i, j \in \mathscr{S}$, that is, whenever we deal with the *single* chain. It will be helpful to look at $\mathscr{R}$ as a set of nodes of a graph. We say that *direct transition* from i to j is *recurrent* if $i \to_0 j$ (1 step), that is, if $\alpha_{ij} = \alpha_i - v_{ij} \geq 0$ or, equivalently, $\{X_n = i, X_{n+1} = j \text{ i.o.}\} = U$. Equip $\mathscr{R}$ with arcs (i, j) corresponding to recurrent direct transitions. We say the resulting graph has *period* d, if for all $i \in \mathscr{R}$, $i \to_0 i$ (r steps) implies that r is a multiple of d. If d is the largest number with this property, we call it the *proper* period. If the graph has proper period d, the class $\mathscr{R}$ can be decomposed into disjoint *periodic subclasses* $\mathscr{P}^{(p)}$, $0 \leq p \leq d-1$, such that $i \in \mathscr{P}^{(p)}$ and $i \to_0 j$ (r steps) implies $j \in \mathscr{P}^{(p+r)}$, where $p+r$ is understood modulo d. The preceding definitions and facts are borrowed from the classical theory of homogeneous chains, so we omit details. Note, however, that the setting is different. The graph of recurrent direct transitions is not equal here to the graph of possible direct transitions.

THEOREM 5.1. *Let $\mathscr{R}$ be a recurrent class and $U = \{X_n \in \mathscr{R} \text{ ult.}\}$. Suppose the graph of recurrent direct transitions in $\mathscr{R}$ has proper period $d \geq 1$ and let $\mathscr{P}^{(p)}$, $0 \leq p \leq d-1$, be the periodic subclasses. The events*

$$A^{(p)} = \{X_{nd} \in \mathscr{P}^{(p)} \text{ i.o.}\} = \{X_{nd} \in \mathscr{P}^{(p)} \text{ ult.}\}, \quad 0 \leq p \leq d-1,$$

if nonnull, are atoms of the tail σ-field $\mathscr{T}$ and $\bigcup_{0 \leq p \leq d-1} A^{(p)} = U$ a.s.

PROOF. It is obvious that $\{X_{nd} \in \mathscr{P}^{(p)} \text{ i.o.}\} = \{X_{nd} \in \mathscr{P}^{(p)} \text{ ult.}\}$ and $\bigcup_{0 \leq p \leq d-1} A^{(p)} = U$ a.s. For definiteness, we will consider, say $A^{(0)}$, and show it is an atom of $\mathscr{T}$, provided that $\mathbf{P}(A^{(0)}) > 0$.

Consider the double chain $(\underline{X}_n)$. By Remark 1.4, this is a chain with regularly diminishing transitions. Elements and subsets of the product space

$\mathscr{S} \times \mathscr{S}$ will be denoted by underlined letters. We have

$$A^{(0)} \times A^{(0)} \subset \{\underline{X}_n \in \underline{\mathscr{P}} \ \text{ ult.}\} \quad \text{a.s., where } \underline{\mathscr{P}} = \bigcup_{0 \le p \le d-1} \mathscr{P}^{(p)} \times \mathscr{P}^{(p)},$$

because $A^{(0)} = \{X_{nd} \in \mathscr{P}^{(0)}, X_{nd+1} \in \mathscr{P}^{(1)}, \ldots, X_{nd+d-1} \in \mathscr{P}^{(d-1)} \ \text{ult.}\}$ a.s. The theorem will be proved if we show that $\underline{\mathscr{P}}$ contains precisely one recurrent class $\underline{\mathscr{R}}$ of the double chain and there exists $m \in \mathscr{R}$ such that $(m, m) \in \underline{\mathscr{R}}$. Indeed,

$$\{\underline{X}_n \in \underline{\mathscr{P}} \ \text{ ult.}\} = \{\underline{X}_n \in \underline{\mathscr{R}} \ \text{ ult.}\} = \{\underline{X}_n = (m, m) \ \text{ i.o.}\} \quad \text{a.s.}$$

will follow from Theorem 4.1, and application of Lemma 5.1 will conclude the proof.

Let us first construct a closed coalition $(\alpha_{\underline{i}}, \underline{i} \in \underline{\mathscr{R}})$ of height ϱ. Our tools here will be Lemmas 3.1 and 3.2. For $\underline{i} \in \underline{\mathscr{P}}$, set

$$\alpha_{\underline{i}} = \alpha_{i_1} + \alpha_{i_2} - \varrho.$$

We check that the system $(\alpha_{\underline{i}}, \underline{i} \in \underline{\mathscr{P}})$ satisfies (3.11). Indeed, fix $\underline{i}, \underline{j} \in \underline{\mathscr{P}}$ and note that $\alpha_{\underline{i}} - v_{\underline{ij}} = h_1 + h_2 - \varrho$, where $h_1 = \alpha_{i_1} - v_{i_1 j_1}$ and $h_2 = \alpha_{i_2} - v_{i_2 j_2}$, by Remark 1.4. Now use the fact that $(\alpha_i, i \in \mathscr{R})$ satisfies (3.11). We have

$$\begin{array}{llll} i_1 \to_{h_1} j_1 & (1 \text{ step}), & j_1 \to_{h_1} i_1 & (r \text{ steps, say}), \\ i_2 \to_{h_2} j_2 & (1 \text{ step}), & j_2 \to_{h_2} i_2 & (s \text{ steps, say}). \end{array}$$

We may assume $r \ge 1$ and $s \ge 1$. We can repeat cyclically the paths from i_1 to j_1 to i_1 and from i_2 to j_2 to i_2, obtaining

$$j_1 \to_{h_1} i_1 \ ((r+1)(s+1) - 1 \text{ step}),$$

$$j_2 \to_{h_2} i_2 \ ((r+1)(s+1) - 1 \text{ step}).$$

Now, as the number of steps on each coordinate is the same, we can compose the two paths into a single path in the product space. Note that all states of this path belong to $\underline{\mathscr{P}}$, because one step moves each coordinate from, say, $\mathscr{P}^{(p)}$ to $\mathscr{P}^{(p+1)}$. Therefore, we get $\underline{j} \to_{h_1+h_2-\varrho} \underline{i}$. We have verified property (3.11). Now we wish to arrange that (3.12) and (3.13) be fulfilled. To this end, let us restrict the domain of the system $(\alpha_{\underline{i}})$. Choose m belonging to the *top* $\mathscr{M}$ of the coalition $(\alpha_i, i \in \mathscr{R})$, that is such that $\alpha_m = \varrho$. Set

$$\underline{\mathscr{R}} = \{\underline{i} \in \underline{\mathscr{P}} : (m, m) \to_0 \underline{i}\}.$$

System $(\alpha_{\underline{i}}, \underline{i} \in \underline{\mathscr{R}})$ satisfies (3.12) and (3.13) by construction [it is easy to see that no states outside $\underline{\mathscr{P}}$ can be reached at height 0 from (m, m)]. It is also clear that (3.11) remains true if the domain $\underline{\mathscr{P}}$ is restricted to $\underline{\mathscr{R}}$. Property (3.5) is obvious. We have built a closed coalition of height ϱ with (m, m) belonging to its domain $\underline{\mathscr{R}}$. We are left with the task of showing that this is the only such coalition inside $\underline{\mathscr{P}}$.

Suppose $(\alpha'_{\underline{i}}, \underline{i} \in \underline{\mathscr{R}}')$ is another closed coalition of height ϱ. Because the recurrence orders for the double chain cannot exceed the recurrence orders for each coordinate, $\alpha'_{\underline{i}} \leq \min(\alpha_{i_1}, \alpha_{i_2})$. Thus, the top of $\underline{\mathscr{R}}'$ must be a subset of $\mathscr{M} \times \mathscr{M}$, where $\mathscr{M}$ is the top of $\mathscr{R}$. We will argue that $\underline{\mathscr{R}}'$ cannot be a subset of $\underline{\mathscr{P}}$ (in fact, $\underline{\mathscr{R}}'$ must be disjoint from $\underline{\mathscr{P}}$), because the coalition $\underline{\mathscr{R}}$ constructed previously contains $(\mathscr{M} \times \mathscr{M}) \cap \underline{\mathscr{P}}$. We claim that

$$\begin{aligned} &\text{if } m_1 \in \mathscr{M} \cap \mathscr{P}^{(p)} \text{ and } i_2, j_2 \in \mathscr{P}^{(p)}, \\ &\text{then } (m_1, i_2) \to_0 (m_1, j_2) \text{ w.r.t. } (\alpha_{\underline{i}}, \underline{i} \in \underline{\mathscr{P}}). \end{aligned} \tag{5.4}$$

Indeed, we have

$$m_1 \to_\varrho m_1 \quad (sd \text{ steps) for some } s \geq 1,$$

by property (3.10) for the system $(\alpha_i, i \in \mathscr{R})$. Of course, s can be replaced by its multiple, because the path can be repeated cyclically. For the second component, we have

$$i_2 \to_0 j_2 \quad (rd \text{ steps) for all } r \text{ sufficiently large},$$

because d is the proper period of the graph of recurrent transitions. We can choose r to be a multiple of s. Composing coordinates along two paths of the same length, in the same way as in the preceding part of the proof, we get (5.4). The rest is easy. Let $m_1, m_2 \in \mathscr{M} \cap \mathscr{P}^{(p)}$. Consider m appearing in the definition of $\underline{\mathscr{R}}$. If, say, $m \in \mathscr{M} \cap \mathscr{P}^{(0)}$, then moving p steps along an arbitrary path at height ϱ we get $(m, m) \to_\varrho (m'_1, m'_2)$ for some $m'_1, m'_2 \in \mathscr{M} \cap \mathscr{P}^{(p)}$. Apply (5.2) twice to obtain

$$(m, m) \to_\varrho (m'_1, m'_2) \to_0 (m'_1, m_2) \to_0 (m_1, m_2) \quad \text{w.r.t. } (\alpha_{\underline{i}}, \underline{i} \in \underline{\mathscr{P}}).$$

We have shown that $(m_1, m_2) \in \underline{\mathscr{R}}$. Therefore, $\underline{\mathscr{R}} \supseteq (\mathscr{M} \times \mathscr{M}) \cap \underline{\mathscr{P}}$. □

The following Proposition 5.1 is an extension of Proposition 4.1. Note that the terms "recurrent class" and "domain of a closed coalition of height ϱ" are synonyms.

PROPOSITION 5.1. *Suppose $\mathscr{P}^{(p)}$, $0 \leq p \leq d-1$, are periodic subclasses of a recurrent class $\mathscr{R}$. If, say, $\mathbf{P}(X_0 \in \mathscr{P}^{(0)}) > 0$, then*

$$\mathbf{P}(X_{nd+p} \in \mathscr{P}^{(p)} \textit{ for all } n \geq 0 \textit{ and } 0 \leq p \leq d-1) > 0.$$

Consequently, $\mathbf{P}(A^{(0)}) > 0$.

PROOF. For clarity, let us treat separately two cases.

Case 1. Assume $d = 1$, that is, the graph of recurrent direct transitions is aperiodic. We are to show that $\mathbf{P}(X_n \in \mathscr{R} \text{ for all } n \geq 0 | X_0 = i) > 0$ for all $i \in \mathscr{R}$. Imagine we have removed all other recurrent classes, leaving only $\mathscr{R}$ and its neighborhood. Consider the process confined to this set of states. More precisely, the new state space will be $\tilde{\mathscr{S}} = \mathscr{R} \cup \tilde{\mathscr{E}}$, where $\tilde{\mathscr{E}} = \{j \in \mathscr{S} : j \notin \mathscr{R}$ and there is an $i \in \mathscr{R}$ such that $v_{ij} < \infty\}$. Let $\mathbf{P}(\tilde{X}_{n+1} = j | \tilde{X}_n = i) = \mathbf{P}(X_{n+1} = j | X_n = i)$ for $i \in \mathscr{R}$ and $j \in \tilde{\mathscr{S}}$. If $i \in \tilde{\mathscr{E}}$, let $\mathbf{P}(\tilde{X}_{n+1} =$

$j|\tilde{X}_n = i) = c_{ij}$ for all n, where $c_{ij} > 0$ for $j \in \mathscr{R}$ and $c_{ij} = 0$ for $j \in \tilde{\mathscr{C}}$. The modified chain $(\tilde{X}_n)$ has the same transition rule inside $\mathscr{R}$. Moreover, $(\tilde{X}_n)$ satisfies the basic assumption (1.1) (with $\tilde{v}_{ij} = v_{ij}$ for $i \in \mathscr{R}, j \in \tilde{\mathscr{S}}$; $\tilde{v}_{ij} = 0$ for $i \in \tilde{\mathscr{C}}, j \in \mathscr{R}$ and $v_{ij} = \infty$ for $i, j \in \tilde{\mathscr{C}}$). It is easy to check that the modified chain has just one recurrent class, namely, $\mathscr{R}$. The new process behaves exactly like the original one until it hits $\tilde{\mathscr{C}}$. So, we are reduced to proving that $\mathbf{P}(\tilde{X}_n \in \mathscr{R}$ for all $n \geq 0|\tilde{X}_0 = i) > 0$ for all $i \in \mathscr{R}$. Suppose, contrary to our claim, that

$$\mathbf{P}(\tilde{X}_n \in \mathscr{R} \text{ for all } n \geq 0 \mid \tilde{X}_0 = i) = 0 \quad \text{for some } i \in \mathscr{R}. \tag{5.5}$$

This implies that for all k sufficiently large,

$$\mathbf{P}(\tilde{X}_n \in \mathscr{R} \text{ for all } n \geq k \mid \tilde{X}_k = j) = 0 \quad \text{for all } j \in \mathscr{R}, \tag{5.6}$$

because to each j there corresponds i such that the process can move with positive probability from $\tilde{X}_0 = i$ to $\tilde{X}_k = j$, without leaving $\mathscr{R}$. Here we have used the assumption of aperiodicity. Thus, we get

$$\mathbf{P}(\tilde{X}_n \in \tilde{\mathscr{C}} \text{ for some } n \geq k \mid \tilde{X}_0 = i) = 1 \tag{5.7}$$

and so $\mathbf{P}(\tilde{X}_n \in \tilde{\mathscr{C}}$ i.o.$|\tilde{X}_0 = i) = 1$, which is impossible. Theorem 4.1 can be applied to the modified chain $(\tilde{X}_n)$, yielding $\mathbf{P}(\tilde{X}_n \in \tilde{\mathscr{C}}$ i.o.$|\tilde{X}_0 = i) = 0$.

Case 2. Assume $d > 1$. The idea of the proof is similar, but we need another, more delicate modification of the chain (X_n). Let the new state space be $\tilde{\mathscr{S}} = \mathscr{R} \cup \tilde{\mathscr{C}}$, where $\tilde{\mathscr{C}} = \mathscr{R} \times \{1, \ldots, d-1\}$ and identify $\mathscr{R}$ with $\mathscr{R} \times \{0\}$. Let $RT = \{(i, j) \in \mathscr{R} \times \mathscr{R} : i \to_0 j \text{ (1 step)}\}$ be the set of arcs of the graph of recurrent direct transitions. For $i, j \in \mathscr{R}$, set $\mathbf{P}(\tilde{X}_{n+1} = j \mid \tilde{X}_n = i) = \mathbf{P}(X_{n+1} = j \mid X_n = i)$, if $(i, j) \in RT$, otherwise zero. Define probabilities of transition from $\mathscr{R}$ to $\tilde{\mathscr{C}}$ by $\mathbf{P}(\tilde{X}_{n+1} = (i, 1) \mid \tilde{X}_n = i) = 1 - \sum_{j \in \mathscr{R}} \mathbf{P}(\tilde{X}_{n+1} = j \mid \tilde{X}_n = i)$. Transitions inside $\tilde{\mathscr{C}}$ and from $\tilde{\mathscr{C}}$ to $\mathscr{R}$ are the following: $\mathbf{P}(\tilde{X}_{n+1} = (i, p+1) \mid \tilde{X}_n = (i, p)) = 1$, where $p < d$ and $p + 1$ is understood modulo d. We claim again that $\mathscr{R}$ is the unique recurrent class of the chain modified in this way. Using Lemmas 3.1 and 3.2, it is not hard to check that the closed coalition $(\alpha_i, i \in \mathscr{R})$ remains a closed coalition for the modified chain and becomes the only one. Now, the key fact is that the graph of all possible transitions of $(\tilde{X}_n)$ has proper period d. Behavior of (X_n) and $(\tilde{X}_n)$ is the same until the first nonrecurrent transition occurs. Therefore,

$$\begin{aligned}
&\mathbf{P}(X_{nd+p} \in \mathscr{P}^{(p)} \text{ for all } n \geq 0 \text{ and } 0 \leq p \leq d-1) \\
&\qquad \geq \mathbf{P}(X_0 \in \mathscr{P}^{(0)}, (X_n, X_{n+1}) \in RT \text{ for all } n \geq 0) \\
&\qquad = \mathbf{P}(\tilde{X}_0 \in \mathscr{P}^{(0)}, (\tilde{X}_n, \tilde{X}_{n+1}) \in RT \text{ for all } n \geq 0) \\
&\qquad = \mathbf{P}(\tilde{X}_{nd+p} \in \mathscr{P}^{(p)} \text{ for all } n \geq 0 \text{ and } 0 \leq p \leq d-1) \\
&\qquad = \mathbf{P}(\tilde{X}_0 \in \mathscr{P}^{(0)}, \tilde{X}_n \in \mathscr{R} \text{ for all } n \geq 0).
\end{aligned}$$

To show that $\mathbf{P}(\tilde{X}_n \in \mathscr{R}$ for all $n \geq 0 \mid \tilde{X}_0 = i) > 0$, $i \in \mathscr{P}^{(0)}$, we argue as in the first part of the proof, with the following adjustments. If (5.5) holds for some $i \in \mathscr{P}^{(0)}$, then we can only claim that (5.6) holds for all $j \in \mathscr{P}^{(k)}$, modulo d. Anyway, we get (5.7), because $\mathbf{P}(\tilde{X}_k \notin \mathscr{P}^{(k)} \mid \tilde{X}_0 \in \mathscr{P}^{(0)}) = 0$. □

6. Simulated annealing. Markov chains, generated by the classical *simulated annealing* algorithm [Kirkpatrick, Gelatt and Vecchi (1983); see van Laarhoven and Aarts (1987)] can be described as follows. Suppose $\mathscr{S}$ is a finite set and to each $i \in \mathscr{S}$, a number w_i is assigned. Let $(c_{ij}, i, j \in \mathscr{S})$ be a stochastic matrix. Consider Markov chain (X_n) with transition probabilities given by

$$\begin{aligned} \mathbf{P}(X_{n+1} = j \mid X_n = i) &= c_{ij}\varepsilon_n^{\max(w_j - w_i, 0)}, \qquad i \neq j, \\ \mathbf{P}(X_{n+1} = i \mid X_n = i) &= 1 - \sum_{j \neq i} \mathbf{P}(X_{n+1} = j \mid X_n = i), \end{aligned} \tag{6.1}$$

where the sequence (ε_n) satisfies (1.2). We will also assume that the order of cooling ϱ satisfies (2.4). The w_i's should be regarded as values of an objective function. The goal is to find m such that $w_m = w_*$, where $w_* = \min_{i \in \mathscr{S}} w_i$. Matrix (c_{ij}) describes random search. Formula (6.1) comprises both the generation of subsequent tentative solutions and the acceptance rule. The process defined by (6.1) is, of course, a chain with regularly diminishing transitions. We have

$$v_{ij} = \begin{cases} \max(w_j - w_i, 0), & \text{if } i \neq j \text{ and } c_{ij} > 0, \\ \infty, & \text{if } i \neq j \text{ and } c_{ij} = 0, \\ 0, & \text{if } i = j \text{ and } c_{ii} > 0, \\ \infty, & \text{if } i = j, c_{ii} = 0 \text{ and } w_k \leq w_i \text{ for all } k \\ & \quad \text{such that } c_{ik} > 0, \\ 0, & \text{if } i = j, c_{ii} = 0 \text{ and there is } k \\ & \quad \text{such that } w_k > w_i \text{ and } c_{ik} > 0. \end{cases} \tag{6.2}$$

To explain why the three cases with $i = j$ appear in (6.2), notice that (6.1) implies

$$\mathbf{P}(X_{n+1} = i \mid X_n = i) = c_{ii} + \sum_{k \neq i} c_{ik}[1 - \varepsilon_n^{\max(w_k - w_i, 0)}].$$

If $w_k > w_i$, then $1 - \varepsilon_n^{\max(w_k - w_i, 0)}$ exceeds 1/2 for large n; otherwise it is zero.

From Theorem 4.1, it is easy to derive necessary and sufficient conditions for reaching the set $\mathscr{G} = \{m : w_m = w_*\}$ of global minima with probability 1. We are going to state one of such conditions in a way resembling the result of Hajek (1988). It was Hajek who introduced the notion of "reachability at height h." Let us first make clear the distinction between his definition and ours. From now on, assume the search matrix (c_{ij}) is irreducible. Consequently, (3.1) holds. Fix τ big enough. Suppose

$$(\lambda_i, i \in \mathscr{S}) \text{ is the coalition of height } \tau. \tag{6.3}$$

Formula (6.3) should be regarded as a definition of the system of numbers (λ_i). Correctness of this definition is ensured by Lemmas 3.9 and 3.4. Note that the role of τ in (6.3) is only to "fix the origin of coordinates," informally speaking. By Lemma 3.1,

(6.4) for all $i, j \in \mathscr{S}$ and all $h, i \to_h j$ w.r.t. (λ_i) implies $j \to_h i$ w.r.t. (λ_i).

On the other hand, define another system $(\lambda'_i, i \in \mathscr{S})$. To simplify notation, assume $w_* = 0$. Set

$$\lambda'_i = \tau - w_i. \tag{6.3$'$}$$

Note that if $c_{ij} > 0$, then $w_i + v_{ij} = \max(w_i, w_j)$ and, therefore, $\lambda'_i - v_{ij} = \tau - \max(w_i, w_j)$. Thus,

$i \to_{\tau-h} j$, w.r.t. (λ'_i) if $i = j$ and $w_i \le h$, or there exists a sequence of states $i = i(0), i(1), \dots, i(r) = j$ such that $c_{i(s-1)i(s)} > 0$ for $1 \le s \le r$ and $w_{i(s)} \le h$ for $0 \le s \le r$.

We see that Hajek's relation of "reachability at height h" is precisely our $\to_{\tau-h}$ *with respect to* (λ'_i) given by (6.3′). Consequently, Hajek's "weak reversibility" condition is equivalent to the following statement:

(6.4′) for all $i, j \in \mathscr{S}$ and all $h, i \to_h j$ w.r.t. (λ'_i) implies $j \to_h i$ w.r.t. (λ'_i).

The difference between (6.4) and (6.4′) is that the former is a consequence of *definition* of the system (λ_i), whereas the latter is an *assumption*.

LEMMA 6.1. *Consider the simulated annealing chain with irreducible* (c_{ij}). *If the weak reversibility condition* (6.4′) *is fulfilled, then the system* (λ_i) *defined by* (6.3) *is equal to the system* (λ'_i) *given by* (6.3′).

PROOF. Follows immediately from Lemma 3.1. □

Note that combining Lemma 6.1 with Lemma 3.10 we get Theorem 5 in Connors and Kumar [(1989), the Potential Theorem].

Unfortunately, general conditions for reaching $\mathscr{G}$ a.s. turn out to be just as simple (or as complicated) as conditions for reaching *any* other subset of $\mathscr{S}$.

PROPOSITION 6.1. *Consider the simulated annealing chain with irreducible* (c_{ij}). *The following three statements are equivalent:*

(i) *For every i we have* $\mathbf{P}(X_n \in \mathscr{G}$ *for some* $n \mid X_0 = i) = 1$.

(ii) $\mathscr{G} \cap \mathscr{R}_t \neq \varnothing$ *for each recurrent class* $\mathscr{R}_t$.

(iii) *For every* $i \in \mathscr{S}$ *there is* $j \in \mathscr{G}$ *such that* $i \to_{\lambda_i - \varrho} j$ *with respect to the system* (λ_i) *given by* (6.3).

PROOF. The fact that (ii) is equivalent to (i) follows straightforward from Theorem 4.1 and Proposition 4.1. To prove that (iii) and (ii) are equivalent, use Lemma 3.10 from Section 3. Condition (iii) implies (ii), because if $m \to_{\lambda_m - \varrho} j$ and $m \in \mathscr{M}_t$, then $j \in \mathscr{R}_t$, by (3.16). To show that (iii) follows from (ii), suppose $i \not\to_{\lambda_i - \varrho} j$ for some i and all $j \in \mathscr{G}$. Consider the set $\mathscr{A} = \{j : i \to_{\lambda_i - \varrho} j\}$. $\mathscr{A}$ is nonempty because $i \in \mathscr{A}$. Select $m \in \mathscr{A}$ such that $\lambda_m = \max_{j \in \mathscr{A}} \lambda_j$. Then m satisfies (3.15) and so it belongs to some $\mathscr{M}_t$. We have $\mathscr{R}_t \cap \mathscr{G} = \varnothing$ for the corresponding recurrent class. □

If the weak reversibility assumption (6.4′) is fulfilled, we can replace (λ_i) by (λ'_i) in (iii) and we obtain the condition, under which Hajek proved convergence *in probability* of simulated annealing to the set of global minima. Let us stress that if (6.4′) fails, Proposition 6.1 remains true, yet system (λ_i) is no longer equal to (λ'_i) and it is no longer directly expressible in terms of the objective function (w_i). To illustrate this point, consider an example.

EXAMPLE 6.1. Let $\mathscr{S} = \{0, 1, 2, 3\}$ and $w_i = i$, for all i. The set of global minima is $\mathscr{G} = \{0\}$. Assume $\varepsilon_n = n^{-1}$, so $\varrho = 1$ and (2.4) holds. Assume that $c_{ii} = 0$ for all i. The following four cases correspond to four ways of defining the "neighborhood structure," that is, positions occupied by nonzero entries in the matrix (c_{ij}).

(a) Assume $c_{01} > 0$, $c_{12} > 0$, $c_{23} > 0$, $c_{10} > 0$, $c_{21} > 0$, $c_{30} > 0$ and $c_{ij} = 0$ for all other pairs $i \neq j$. Condition (6.4′) holds and we have

$$\lambda_0 = \lambda'_0 = \tau, \qquad \lambda_1 = \lambda'_1 = \tau - 1, \qquad \lambda_2 = \lambda'_2 = \tau - 2, \qquad \lambda_3 = \lambda'_3 = \tau - 3.$$

There is one recurrent class $\mathscr{R} = \{0, 1\}$. $\mathscr{G}$ is reachable with probability 1.

(b) Assume $c_{01} > 0$, $c_{12} > 0$, $c_{23} > 0$, $c_{21} > 0$, $c_{30} > 0$ and $c_{ij} = 0$ for all other pairs $i \neq j$. Condition (6.4′) fails:

$$\lambda_0 = \tau - 1, \qquad \lambda_1 = \tau, \qquad \lambda_2 = \tau - 1, \qquad \lambda_3 = \tau - 2.$$

Now, there is one recurrent class $\mathscr{R} = \{1, 2\}$. $\mathscr{G}$ is not reachable with probability 1.

(c) Assume $c_{01} > 0$, $c_{12} > 0$, $c_{23} > 0$, $c_{30} > 0$ and $c_{ij} = 0$ for all other pairs $i \neq j$. Condition (6.4′) fails:

$$\lambda_0 = \lambda_1 = \lambda_2 = \tau, \qquad \lambda_3 = \tau - 1.$$

The whole $\mathscr{S}$ becomes a recurrent class. $\mathscr{G}$ is reachable with probability 1.

(d) Assume $c_{02} > 0$, $c_{13} > 0$, $c_{23} > 0$, $c_{30} > 0$, $c_{31} > 0$ and $c_{ij} = 0$ for all other pairs $i \neq j$. Condition (6.4′) fails:

$$\lambda_0 = \tau, \qquad \lambda_1 = \tau, \qquad \lambda_2 = \tau - 1, \qquad \lambda_3 = \tau - 2.$$

There are two recurrent classes: $\mathscr{R}_1 = \{0\}$ and $\mathscr{R}_2 = \{1\}$. $\mathscr{G}$ is not reachable with probability 1. □

To conclude this section, let us point out an application of Theorem 5.1.

PROPOSITION 6.2. *Consider the simulated annealing chain with irreducible* (c_{ij}). *Assume the objective function* (w_i) *is not constant. The following three statements are equivalent.*

(i) *The chain is weakly ergodic.*
(ii) *There is only one recurrent class.*
(iii) *For every pair of states* $i, j \in \mathscr{S}$, *there is* $k \in \mathscr{S}$ *such that* $i \to_{\lambda_i - \varrho} k$ *and* $j \to_{\lambda_j - \varrho} k$ *with respect to the system* (λ_i) *given by* (6.3).

PROOF. By Theorem 8.1 in Cohn (1987), the chain is weakly ergodic iff the tail σ-field is $[\mathbf{P}]$ trivial for every initial distribution. Therefore, (i) implies (ii).

To show that (i) follows from (ii), use Theorem 5.1. We are going to check that the unique recurrent class, say $\mathscr{R}$, does not have periodic subclasses. Begin with the observation that there exist states $i \in \mathscr{R}$ and $j \in \mathscr{S}$ such that $c_{ij} > 0$ and $w_j > w_i$. Indeed, if it were false, $i \in \mathscr{R}$ and $i \to j$ [reachability in the sense defined by (3.2)] would imply $w_j \le w_i$ and hence $j \in \mathscr{R}$, by induction, in view of (6.2) and (2.6). Then we could use the irreducibility assumption (3.1) to deduce $w_i = \text{const}$, contrary to the assumption. If $c_{ij} > 0$ and $w_j > w_i$, then $v_{ii} = 0$, by (6.2). Therefore, we can find $i \in \mathscr{R}$ such that $v_{ii} = 0$, which implies that the graph of recurrent direct transition is aperiodic.

Now, we verify that (ii) is equivalent to (iii). Suppose (ii) holds. Let $\mathscr{M}$ be the top of the unique recurrent class $\mathscr{R}$. We claim that $i \to_{\lambda_i - \varrho} k$ for all $i \in \mathscr{S}$ and $k \in \mathscr{M}$. Indeed, if $i \not\to_{\lambda_i - \varrho} k$, then consider $\mathscr{A} = \{j : i \to_{\lambda_i - \varrho} j\}$ and select $m \in \mathscr{A}$ such that $\lambda_m = \max_{j \in \mathscr{A}} \lambda_j$, just as in the proof of Proposition 6.1. It is clear that m and k belong to tops of two distinct recurrent classes, contrary to (ii). Thus (ii) implies (iii). Converse is easy. If two recurrent classes exist, say $\mathscr{R}_1$ and $\mathscr{R}_2$, select $i \in \mathscr{M}_1$ and $j \in \mathscr{M}_2$. If (iii) were true, we could find k belonging simultaneously to $\mathscr{R}_1$ and $\mathscr{R}_2$, which is impossible. □

EXAMPLE 6.2. Let us return to the four cases in Example 6.1 to illustrate Proposition 6.2. The reachability relation is understood with respect to (λ_i).

(a) We have $i \to_{\lambda_i} 0$ for all i, because $0 \to_\tau 0$, $1 \to_{\tau-1} 0$, $2 \to_{\tau-2} 0$ and $3 \to_{\tau-3} 0$.
(b) We have $i \to_{\lambda_i - 1} 1$ for all i, because $0 \to_{\tau-2} 1$, $1 \to_\tau 1$, $2 \to_{\tau-1} 1$ and $3 \to_{\tau-2} 1$.
(c) We have $i \to_{\lambda_i - 1} 1$ for all i, because $0 \to_\tau 0$, $1 \to_{\tau-1} 0$, $2 \to_{\tau-1} 0$ and $3 \to_{\tau-1} 0$.
(d) Now, $\lambda_0 = \lambda_1 = \tau$; we have $0 \not\to_{\tau-1} i$ for $i \neq 0$ and $1 \not\to_{\tau-1} i$ for $i \neq 1$.

Under additional assumptions, strong ergodicity can be deduced from weak ergodicity. Well-known general theorems or the results of Anily and Federgruen (1987) can be used.

7. Concluding remarks. The literature on tail events of nonhomogeneous Markov chains is fairly large. Let us mention Cohn (1976, 1981, 1982, 1987) and Mukherjea (1985), who treated the problem in a more general setting than we have. There is, however, a price to be paid for the higher generality. The results of the cited authors are much less explicit than ours. Roughly speaking, they relate the structure of tail events to asymptotic behavior of multistep transition probabilities. If theorems of this kind are to be applied to simulated annealing, their assumptions turn out to be extremely difficult to check. For chains with regularly diminishing transitions, bounds on $\mathbf{P}(X_{n+k} = j | X_n = i)$ were derived by Tsitsiklis (1989). However, Tsitsiklis did not examine implications of the bounds for the structure of tail evens and his ad hoc definition of recurrent states does not seem to be appropriate. Chiang and Chow (1989) obtained precise results on asymptotics of $\mathbf{P}(X_n = i)$. They worked under assumptions that force the chain to be strongly ergodic and they were not concerned with classification of states. It is interesting to notice that quite different approaches lead to closely related results. Algorithm II of Tsitsiklis produces classes equal to our $\mathscr{M}_t$'s, *tops* of closed coalitions. The "height" function h of Chiang and Chow (with opposite sign) is the solution to "modified balance equations" of Connors and Kumar (1989) [$h(i) = \tau - \lambda_i$ in our Lemma 3.9]. Detailed discussion of these connections goes beyond the scope of this paper. Let us conclude with some conjectures. It is plausible that $\limsup_n \mathbf{P}(X_n = i) > 0$ if and only if $i \in \bigcup_t \mathscr{M}_t$, that is, if $\alpha_i = \varrho$ with positive probability. We think $\liminf_n \mathbf{P}(X_n = i) > 0$ if $i \in \mathscr{M}_t$ and the graph of recurrent transitions in $\mathscr{R}_t$ is aperiodic. These conjectures, if proved, would complete our classification of states, making a distinction between *positive recurrent* states and *null recurrent* states. It would follow that, in the case of aperiodicity, chains with regularly diminishing transitions have *bases* in the sense defined by Mukherjea (1984) or Cohn (1982).

APPENDIX

Hajek (1988) contrived a particularly nice example of a simulated annealing Markov chain with complex "neighborhood structure." Let us illustrate our Theorem 4.1 using this example. Transition probabilities of the ten chains to be considered are given by (6.1). Figures 1–10 correspond to various sequences (ε_n). Using the simulated annealing parlance, we can say the figures describe Markov chains with various "rates of cooling." The respective values of the order of cooling ϱ are indicated. We omit detailed description, because all relevant information can be shown graphically. Circles represent states, vertical coordinates of their centers are equal to values of objective function w. Arrows indicate possible transitions [corresponding to nonzero entries in the search matrix (c_{ij})]. The set of states, the objective function and the "neighborhood structure" shown by arrows are the same on each figure. Note that the weak reversibility condition (6.4$'$) holds. We actually used Lemmas 6.1 and 3.10 to identify recurrent classes. Recurrent states are blackened.

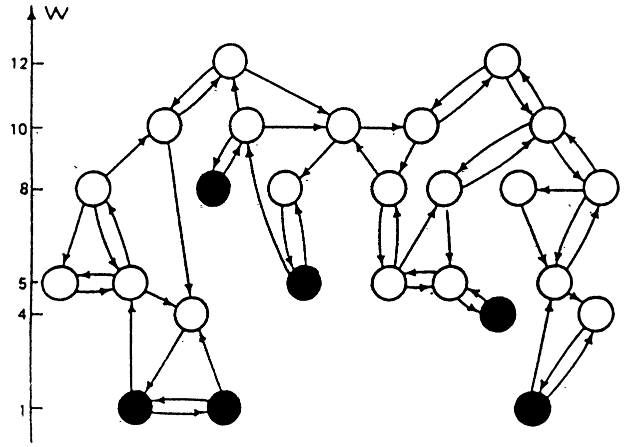

FIG. 1. $0 \leq \varrho < 1$.

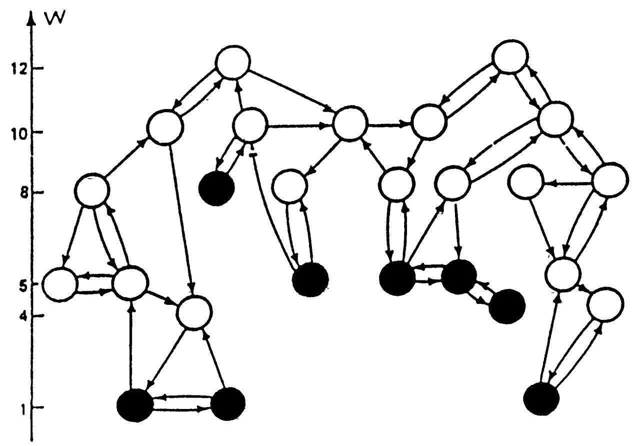

FIG. 2. $1 \leq \varrho < 2$.

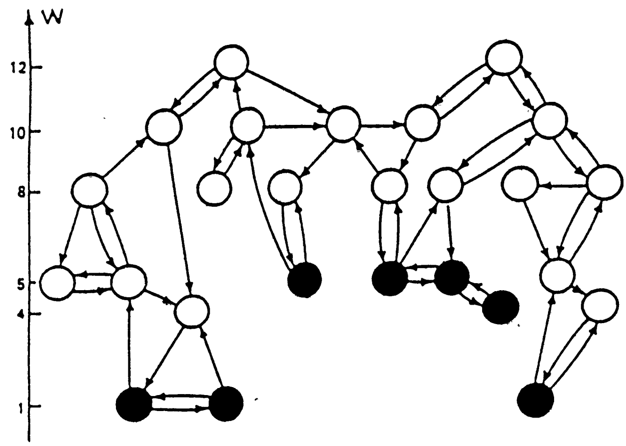

FIG. 3. $2 \leq \varrho < 3$.

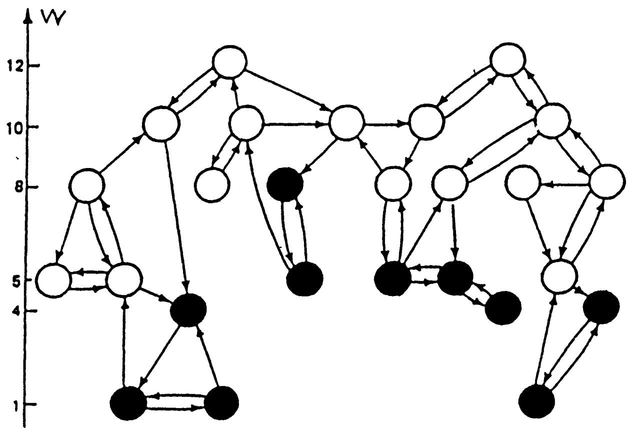

FIG. 4. $3 \leq \varrho < 4$.

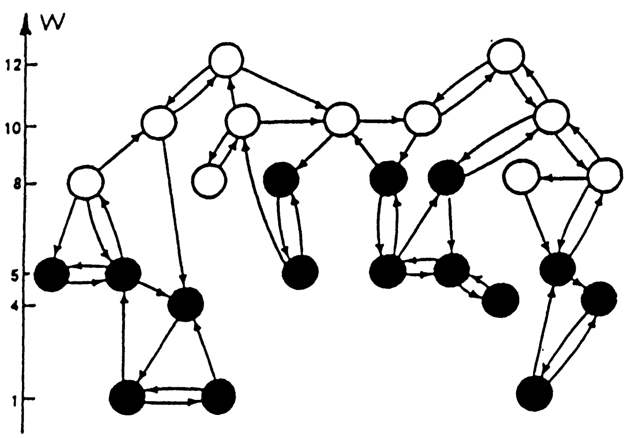

FIG. 5. $4 \leq \varrho < 5$.

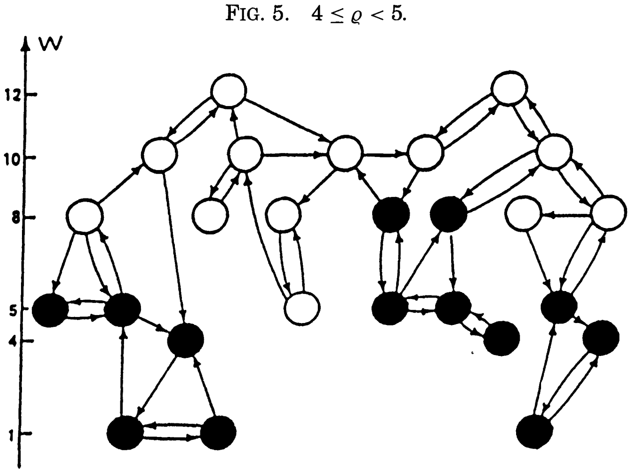

FIG. 6. $5 \leq \varrho < 6$.

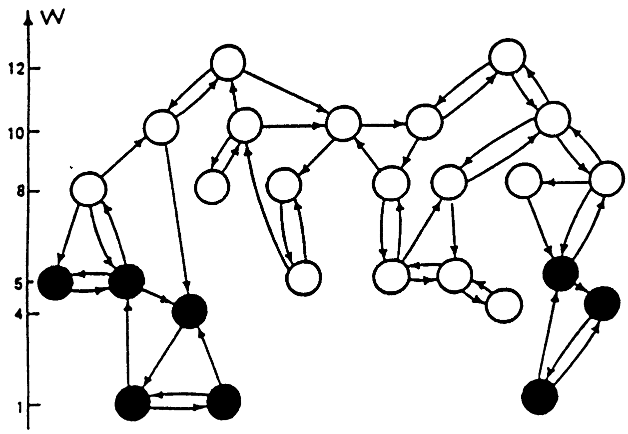

FIG. 7. $6 \leq \varrho < 7$.

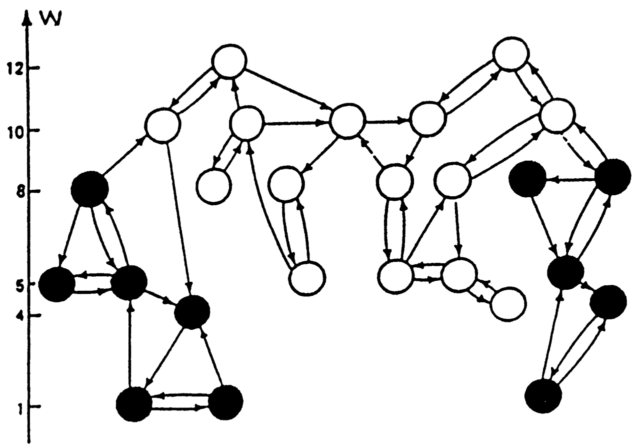

FIG. 8. $7 \leq \varrho < 9$.

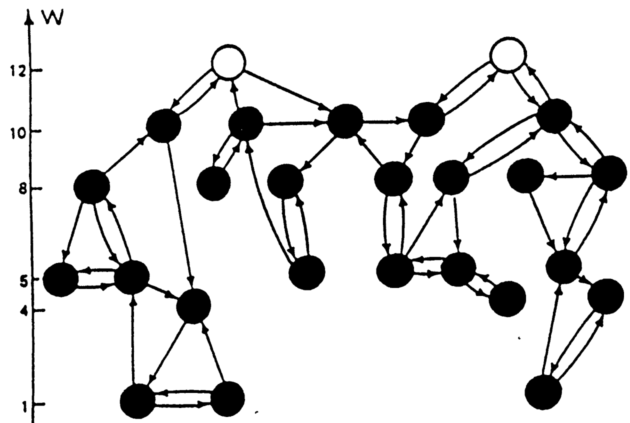

FIG. 9. $9 \leq \varrho < 11$.

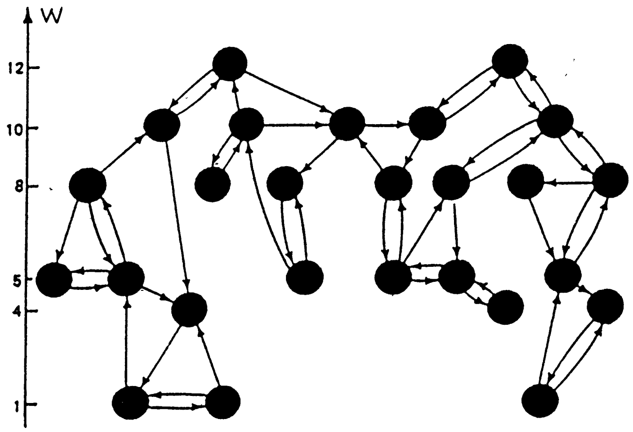

FIG. 10. $\varrho \geq 11$.

Acknowledgment. The authors are grateful to Associate Editor for many helpful suggestions and for pointing out several deficiencies in the first version of the paper.

REFERENCES

ANILY, S. and FEDERGRUEN, A. (1987). Simulated annealing methods with general acceptance probabilities. *J. Appl. Probab.* **24** 657–667.

BORKAR, V. S. (1992). Pathwise recurrence orders and simulated annealing. *J. Appl. Probab.* **29** 472–476.

CHIANG, T. S. and CHOW, Y. (1989). On the asymptotic behavior of some inhomogeneous Markov processes. *Ann. Probab.* **17** 1483–1502.

COHN, H. (1970). On the tail σ-algebra of the finite inhomogeneous Markov chains. *Ann. Math. Statist.* **41** 2175–2176.

COHN, H. (1976). Finite non-homogeneous Markov chains: asymptotic behavior. *Adv. in Appl. Probab.* **8** 502–516.

COHN, H. (1981). On a paper by Doeblin on non-homogeneous Markov chains. *Adv. in Appl. Probab.* **13** 388–401.

COHN, H. (1982). On a class of non-homogeneous Markov chains. *Math. Proc. Cambridge Phil. Soc.* **92** 527–534.

COHN, H. (1987). Products of stochastic matrices and applications. *Int. J. Math. Math. Sci.* **12** 209–233.

CONNORS, D. P. (1988). Balance of recurrence order in time-inhomogeneous Markov chains with application to simulated annealing. Ph.D. thesis, Univ. Illinois.

CONNORS, D. P. and KUMAR, P. R. (1989). Simulated annealing type Markov chains and their balance equations. *SIAM J. Control Optim.* **27** 1440–1461.

DURRETT, R. (1991). *Probability. Theory and Examples.* Wadsworth & Brooks/Cole, Pacific Grove, CA.

HAJEK, B. (1988). Cooling schedules for optimal annealing. *Math. Operat. Res.* **13** 311–329.

KIRKPATRICK, S., GELATT, C. D. and VECCHI, M. P. (1983). Optimization by simulated annealing. *Science* **220** 671–680.

MUKHERJEA, A. (1985). Non-homogeneous Markov chains: tail idempotents, tail sigma-fields and basis. *Proceedings of the 7th Conference on Probability Theory, 1982, Brasov, Romania.* IVNU Science Press, Utrecht.

TSITSIKLIS, J. (1989). Markov chains with rare transitions and simulated annealing. *Math. Operat. Res.* **14** 70–90.

VAN LAARHOVEN, P. J. M. and AARTS, E. H. L. (1987). *Simulated Annealing. Theory and Applications.* Riedel, Dordrecht.

INSTITUTE OF APPLIED MATHEMATICS
UNIVERSITY OF WARSAW
BANACHA 2
02-097 WARSAW
POLAND

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSAW
POLAND