1963

Semi-Markov chains

Denny Durfee Culbertson

The University of Montana

Let us know how access to this document benefits you.

Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation

https://scholarworks.umt.edu/etd/8340

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
ACKNOWLEDGMENTS

I wish to thank Professor Howard Reinhardt for his guidance and instruction throughout the preparation of this thesis. Also, I should like to express my appreciation to Professor William Ballard and Merle Manis for their careful reading of the manuscript and their helpful suggestions for its improvement.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>SOME LEMMAS FROM NUMBER THEORY</td>
<td>3</td>
</tr>
<tr>
<td>RECURRENT EVENTS AND MARKOV CHAINS</td>
<td>8</td>
</tr>
<tr>
<td>SEMI-MARKOV CHAINS</td>
<td>21</td>
</tr>
<tr>
<td>REFERENCES CITED</td>
<td>35</td>
</tr>
</tbody>
</table>
INTRODUCTION

Often problems in Markov chains can be extended in such a way that one may think of waiting times between changes of states. The purpose of this paper is to show how this can be done in some cases by redefining the states so that we still have a Markov chain and the calculation of unknowns can be done in terms of Markov chain theory of the original states and the probability function describing the distribution of waiting times in the various states.

In the first part of this paper we review the notions of Markov chains and prove the theorems we need. Following Feller [3] we define recurrent events and develop an expression for the probability that an event occurs in terms of the probability that it occurs for the first time in some trial (Theorem 1). Lemmas 2 and 3 are taken in part from Niven and Zuckerman [4]. Armed with Theorem 1 and the definition and classification of Markov chains, we give Feller's proof of Theorem 4 which characterizes irreducible, aperiodic Markov chains in terms of their transition probabilities and their mean recurrence times.

We then leave Feller and follow Anselone [2] who develops the notion of semi-Markov chains. First of all we extend the Markov chain by introducing the idea of waiting times in the states. We use the waiting times to define a semi-Markov chain. We then define the sub-state chain by pairing each state from the semi-Markov chain with the time the process will remain in that state. The substate chain is a Markov chain. It is this extension of the semi-Markov chain that

1
yields the information we seek, that is, characterizes our problem. In a straightforward way we arrive at Anselone's results in Theorems 11 and 12, which are more or less analogous to Theorems 2 and 4 for Markov chains.
SOME LEMMAS FROM NUMBER THEORY

We begin by stating and proving three lemmas from number theory.

Lemma 1: From any set \( \{a_i\} \) of positive integers with greatest common divisor one, it is possible to choose a finite subset with greatest common divisor one.

Proof: Choose \( a_1 \), the smallest element of \( \{a_i\} \). If \( a_1 \neq 1 \), choose \( a_2 \), the smallest element of \( \{a_i\} \) such that \( a_1 \mid a_2 \) (i.e., \( a_1 \) does not divide \( a_2 \)). Let \( g_1 = (a_1,a_2) \). If \( g_1 \neq 1 \), choose \( a_3 \), the smallest element of \( \{a_i\} \) such that \( g_1 \mid a_3 \). Let \( g_2 = (a_1,a_2,a_3) \). Then \( g_2 \leq g_1 \), but since \( g_2 \mid a_3 \) and \( g_1 \mid a_3 \), \( g_2 < g_1 \). We continue in this fashion, obtaining a monotone decreasing sequence of positive integers. There exists, for some positive integer \( n \), a \( g_n = (a_1,a_2,\ldots,a_{n+1}) = 1 \), for if \( g_n > 1 \) we can obtain a \( g_{n+1} < g_n \).

Lemma 2: If \( a \), \( b \), and \( c \) are positive integers such that \((a,b)\mid c\), and \((a,b)c > ab\), then there exists at least one positive solution to \( ax + by = c \).

Proof: There exist integers \( x_0 \) and \( y_0 \) such that \( ax_0 + by_0 = (a,b) \). All integral solutions \( r,s \) of \( ax + by = c \) can be written in the form

\[
r = cx_0/(a,b) + bt/(a,b)
\]

and

\[
s = cy_0/(a,b) - at/(a,b).
\]
For a solution to be positive it is necessary and sufficient that
\[-\frac{cx_0}{b} < t < \frac{cy_0}{a}.

We have \(cax_0 + cby_0 = (a,b)c > ab\) and, dividing by \(ab\), we obtain
\[\frac{cx_0}{b} + \frac{cy_0}{a} > 1\]
or,
\[-\frac{cx_0}{b} < 1 - \frac{cx_0}{b} < \frac{cy_0}{a}.

Thus the length of the interval \(-\frac{cx_0}{b}, \frac{cy_0}{a}\) is greater than one so that there exists at least one integer \(t\) such that
\[-\frac{cx_0}{b} < t < \frac{cy_0}{a}.

Lemma 3: If \(\{a_i\}\) is a finite sequence of distinct positive integers with greatest common divisor one, and \(k\) is an integer such that
\[k > \prod_{i=1}^{n} a_i,\]
then there exist positive integers \(x_i\) such that \(k = \sum_{i=1}^{n} a_i x_i\).

Proof: From the equation \(\sum_{i=1}^{n} a_i x_i = k\) we first derive the equation \(a_1x_1 + b_1y_1 = k\) in such a way that \(b_1 > 0\), \((a_1,b_1) = 1\), and \(a_1b_1 < k\). This will imply, by Lemma 2, that there exists at least one positive integral solution to \(a_1x_1 + b_1y_1 = k\). Using an \(x_1\) from one of these solutions, we construct the desired solution for the equation
\[\sum_{i=1}^{n} a_i x_i = k\]
Suppose $\sum_{i=1}^{n} a_i x_i = k$. Let $\beta = -a_n/(a_{n-1}, a_n)$ and $\delta = a_{n-1}/(a_{n-1}, a_n)$.

Then $(\beta, \delta) = 1$ so that there exist $\alpha$ and $\gamma$ such that $\alpha \delta - \beta \gamma = 1$.

Let $u = \delta x_{n-1} - \beta x_n$ and $v = \gamma x_{n-1} + \alpha x_n$ so that $x_{n-1} = \alpha u + \beta v$ and $x_n = \gamma u + \delta v$. We claim that

$$\sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{n-2} a_i x_i + (a_{n-1} \alpha + a_n \gamma) u.$$ 

Noting that $(a_{n-1} \alpha + a_n \gamma) = \delta (a_{n-1}, a_n) \alpha - \beta (a_{n-1}, a_n) \gamma$ 

$$= (\delta \alpha - \beta \gamma ) (a_{n-1}, a_n)$$

$$= (a_{n-1}, a_n)$$

we have $a_{n-1} \alpha + a_n \gamma > 0$. Since $(a_1, a_2, \ldots, a_{n-2}, (a_{n-1}, a_n)) = (a_1, a_2, \ldots, a_n)$ we have obtained an equation whose coefficients are positive integers with greatest common divisor one, and the number of coefficients is one less. We may continue in this manner obtaining

$$a_1 x_1 + b_1 y_1 = k$$

where $b_1 = (a_2, a_3, \ldots, a_n) > 0$ and $(a, b) = (a_1, a_2, \ldots, a_n) = 1$. Also

$$a_1 b_1 < \prod_{i=1}^{n} a_i < k.$$ 

Hence $a_1 x_1 + b_1 y_1 = k$ has at least one positive integral solution.

Now we wish to show that for suitable choice of $x_1$

$$a_2 x_2 + b_2 y_2 = k - a_1 x_1$$

has a positive integral solution and, in general, that

$$a_rx_r + b_r y_r = k - \sum_{i=1}^{r-1} a_i x_i$$

where $r \geq 2$ and $b_r = (a_{r+1}, a_{r+2}, \ldots, a_n)$ has a positive integral solution.
Since the $a_1$'s are all distinct we may suppose that $a_n$ is the least $a_1$. Also, from the positive integral solutions of $a_1x_1 + b_1y_1 = k$, we shall choose the smallest $x_1$. This means that $x_1 \leq b_1$ since otherwise $a_1(x_1 - b_1) + b(y_1 + a_1) = k$ and $0 < x_1 - b_1 < x_1$.

Having reduced $\sum_{i=1}^{n} a_ix_i = k$ to $a_1x_1 + b_1y_1 = k$, we take as an induction hypothesis that

$$a_sx_s + b_sy_s = k - \sum_{i=1}^{s-1} a_ix_i$$

has positive integral solutions for $s = 1, 2, \ldots, r-1 < n$. Each $x_s$ is taken as the least positive solution. In order to find positive integral solutions to

$$a_rx_r + b_ry_r = k - \sum_{i=1}^{r-1} a_ix_i$$

we need to show that

$$a_rb_r < (a_r, b_r)(k - \sum_{i=1}^{r-1} a_ix_i).$$

For this it will suffice that

$$a_rb_r + \sum_{i=1}^{r-1} a_ix_i < k.$$  

Since $x_i \leq b_i$ for $i = 1, 2, \ldots, r - 1$ we have

$$\sum_{i=1}^{r-1} a_ix_i \leq \sum_{i=1}^{r-1} a_ib_i.$$  

Also, since $(a_{i+1}, a_{i+2}, \ldots, a_n) \leq (a_{i+2}, a_{i+3}, \ldots, a_n)$ -- i.e., $b_i \leq b_{i+1}$ -- we have

$$b_ra_r + \sum_{i=1}^{r-1} a_ix_i \leq \sum_{i=1}^{r} a_ib_i \leq b_r \sum_{i=1}^{r} a_i.$$  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
If any of the $a_i = 1$, it is $a_n$, and since $r < n$ we have

$$\sum_{i=1}^{r} a_i \leq \prod_{i=1}^{r} a_i.$$ 

Hence

$$b_r \sum_{i=1}^{r} a_i \leq b_r \prod_{i=1}^{r} a_i \leq \prod_{i=1}^{n} a_i < k.$$

So we have

$$a_r b_r + \sum_{i=1}^{r-1} a_i x_i < k.$$

By hypothesis we have

$$b_{r-1} y_{r-1} = k - \sum_{i=1}^{r-1} a_i x_i$$

so that $b_{r-1}$ divides $k - \sum_{i=1}^{r-1} a_i x_i$. Since $b_{r-1} = (a_r, b_r)$, we have satisfied the conditions of Lemma 2. Therefore,

$$a_r x_r + b_r y_r = k - \sum_{i=1}^{r-1} a_i x_i$$

has at least one positive integral solution.
RECURRENT EVENTS AND MARKOV CHAINS

Suppose we have a sequence of experiments each with possible outcomes \( E_1, E_2, \ldots, E_n, \ldots \). We speak of an attribute \( \mathcal{E} \) of some finite sequence of trials. That is, any finite sequence of trials either possesses the attribute \( \mathcal{E} \) or it does not. To say that \( \mathcal{E} \) occurs at the \( n^{th} \) place of the sequence \( E_{j_1}, E_{j_2}, \ldots, E_{j_n} \) means that this sequence possesses the attribute \( \mathcal{E} \). For example, if the outcomes \( E_j \) are the positive integers, a particular attribute \( \mathcal{E} \) might be "an even integer occurs on the fifth trial." In particular, we wish to speak of recurrent events.

**Definition 1:** The attribute \( \mathcal{E} \) defines a recurrent event provided that:

1) \( \mathcal{E} \) occurs in the \( n^{th} \) place and the \( (n + m)^{th} \) place of the sequence \( (E_{j_1}, E_{j_2}, \ldots, E_{j_{n+m}}) \) means that \( \mathcal{E} \) occurs in the last place of each of the two subsequences \( (E_{j_1}, E_{j_2}, \ldots, E_{j_n}) \) and \( (E_{j_{n+1}}, E_{j_{n+2}}, \ldots, E_{j_{n+m}}) \).

2) Whenever this happens

\[
P \{ E_{j_1}, E_{j_2}, \ldots, E_{j_{n+m}} \} = P \{ E_{j_1}, E_{j_2}, \ldots, E_{j_n} \} \cdot P \{ E_{j_{n+1}}, E_{j_{n+2}}, \ldots, E_{j_{n+m}} \}.
\]

We shall adopt the following notation:

\[
u_n = P \{ \mathcal{E} \text{ occurs on the } n^{th} \text{ trial} \},
\]

\[
u_0 = 1.
\]
\[ f_n = P \{ \mathcal{E} \text{ occurs for the first time on the } n^{\text{th}} \text{ trial} \}, \]
\[ f_0 = 0, \text{ and} \]
\[ f = \sum_{n=1}^{\infty} f_n. \]

We note that
\[ u_n = f_1 u_{n-1} + f_2 u_{n-2} + \ldots + f_n u_0, \quad n \geq 1. \]

**Definition 2:** A recurrent event \( \mathcal{E} \) is called **persistent** if \( f = 1 \) and **transient** if \( f < 1 \).

**Theorem 1:** Suppose \( 0 \leq g_n \leq 1, \sum_{n=1}^{\infty} g_n = 1, g_0 = 0, \) and
\[ x_n = g_1 x_{n-1} + g_2 x_{n-2} + \ldots + g_n x_0, \quad n = 1, 2, \ldots. \]
If \( \gcd \{ n \mid g_n > 0 \} = 1 \), then
\[ \lim_{n \to \infty} x_n = (\sum_{n=1}^{\infty} n g_n)^{-1} \]
if \( \sum_{n=1}^{\infty} n g_n \) is finite, and
\[ \lim_{n \to \infty} x_n = 0 \]
if \( \sum_{n=1}^{\infty} n g_n \) diverges.

**Proof:** Let
\[ r_n = \sum_{i=1}^{n} g_{n+1} = \sum_{i=n+1}^{\infty} g_i \text{ and } \mu = \sum_{n=1}^{\infty} n g_n \]
so that
\[ \mu = \sum_{n=1}^{\infty} r_n. \] We have then,
\[ r_n = r_0 - \sum_{i=1}^{n} g_i \]
and therefore \( r_{n-1} - r_n = g_n \). Substituting into
\[ x_n = g_1 x_{n-1} + g_2 x_{n-2} + \ldots + g_n x_0, \]
we obtain
\[ x_n = (r_0 - r_1)x_{n-1} + (r_1 - r_2)x_{n-2} + \cdots + (r_{n-1} - r_n)x_0. \]
Thus
\[ r_0x_n + r_1x_{n-1} + \cdots + r_nx_0 = r_0x_{n-1} + r_1x_{n-2} + \cdots + r_{n-1}x_0. \]
This shows inductively that
\[ 1 = r_0x_0 = \ldots = \sum_{i=0}^{n} r_i x_{n-1} \]
for all \( n \). Now \( x_1 = g_1x_0 \leq 1 \) since \( x_0 = 1 \) and \( g_1 \leq 1 \).
Suppose that \( x_0, x_1, \ldots, x_k \) are all at most one. Then
\[ x_{k+1} = g_1x_k + \ldots + g_{k+1}x_0 \leq g_1 + g_2 + \ldots + g_{k+1} \leq 1. \]
Therefore, there exists a \( \lambda = \limsup x_n \), i.e., for every \( \varepsilon > 0 \), there exists an \( M \) such that \( n > M \) implies that \( x_n < \lambda + \varepsilon \). Also, there exists a sequence \( \{n_v\} \) such that \( \lim_{v \to \infty} x_{n_v} = \lambda \). Choose an integer \( j > 0 \) such that \( g_j > 0 \). Then we assert that \( \lim_{v \to \infty} x_{n_v-j} = \lambda \).
Suppose that this were not true. Then for any \( \varepsilon > 0 \) and each \( N \) there exists an \( n_v > N \) such that either \( x_{n_v-j} < \lambda - \varepsilon \) or \( x_{n_v-j} > \lambda + \varepsilon \).
If \( N > M \), then the latter is impossible so that there exists a \( \lambda' \) such that \( x_{n_v-j} < \lambda' < \lambda \). Since \( \lim_{v \to \infty} x_{n_v} = \lambda \), if we take \( N \) large enough we also have \( x_{n_v} > \lambda - \varepsilon \). For every \( \delta > 0 \) there exists an \( R > j \) such that \( r_n < \delta \) for all \( n > R \), since \( \sum_{n=1}^{\infty} g_n \) converges. Since \( g_0 = 0 \) and \( x_k \leq 1 \), we take \( \delta = \varepsilon \) so that
\[ x_{n_v} \leq g_0x_{n_v} + g_1x_{n_v-1} + \cdots + g_{R+1}x_{n_v-R} + \varepsilon \]
for \( n > R \). Also \( n > M + R \) implies that \( x_n < \lambda + \varepsilon \) so that
\[
x_{n_v} < (g_0 + g_1 + \cdots + g_{j-1} + g_{j+1} + \cdots + g_R)(\lambda + \varepsilon) + g_j\lambda' + \varepsilon
\leq (1 - g_j)(\lambda + \varepsilon) + g_j\lambda' + \varepsilon
< \lambda + 2\varepsilon - g_j(\lambda - \lambda').
\]
Choose $\epsilon$ such that $3\epsilon < g_j(\lambda - \lambda')$ so that $x_{n\nu} < \lambda - \epsilon$, a contradiction of $x_{n\nu} > \lambda - \epsilon$.

Similarly we see that $g_j > 0$ and $\lim_{n\nu \to \infty} u_{n\nu} = \lambda$ implies that $x_{n\nu - 2j} \to \lambda$, $x_{n\nu - 3j} \to \lambda$, \ldots.

Consider, first, the case where $g_{\perp} > 0$. We take $j = 1$ and conclude that $x_{n\nu - k} \to \lambda$ for all $k$. Since $r_0 x_{n\nu} + r_1 x_{n\nu - 1} + \ldots + r_n x_0 = 1$, we have

$$r_0 x_{n\nu} + r_1 x_{n\nu - 1} + \ldots + r_n x_{n\nu - N} \leq 1$$

for $n = n\nu$. For fixed $N$, $x_{n\nu - k} \to \lambda$ for all $k \leq N$, so that

$$(r_0 + r_1 + \ldots + r_N) \leq 1.$$ 

Since $N$ is arbitrary, $\lambda / \mu \leq 1$ or $\lambda \leq 1 / \mu$. If $\sum_{n=0}^{\infty} r_n$ diverges, then $\lim_{n \to \infty} x_n = 0$. If $\mu < \infty$, let $\gamma = \lim \inf x_n$. The same argument as above shows that for every sequence $\{n_{\nu}\}$ for which $\lim_{n\nu \to \infty} x_{n\nu} = \gamma$ we have $\lim_{n\nu \to \infty} x_{n\nu - k} = \gamma$ for all $k$. If $N$ is large enough so

$$\sum_{n=N}^{\infty} r_n < \epsilon,$$

then

$$1 \leq r_0 x_{n\nu} + \ldots + r_n x_{n\nu - N} + \epsilon,$$

so that

$$1 \leq (r_0 + r_1 + \ldots + r_N) \gamma + \epsilon.$$ 

Hence $1 / \mu \leq \gamma$, so that $\lambda \leq 1 / \mu \leq \gamma$. But $\lim \inf x_n \leq \lim \sup x_n$. Therefore, $\lim_{n \to \infty} x_n = 1 / \mu$.

Consider, now, the case where $g_{\perp} = 0$. By Lemma 1, we can choose from the set of integers $j$ for which $g_j > 0$ a finite collection $\{a_i\}$, $i = 1, 2, \ldots, n$, such that $g.c.d. \{a_i\} = 1$. We know that when
By Lemma 3, \( \sum_{i=1}^{n} a_i y_i = k \), where \( k \) is a positive integer, has positive integral solutions provided that \( \prod_{i=1}^{n} a_i < k \). Hence \( x_{ny-k} \rightarrow \lambda \). The remainder of the proof follows as in the preceding case.

In the theory of Markov chains we consider outcomes whose probabilities depend only upon the outcome of the preceding trial. Hence, knowing the outcome of any particular trial, say \( E_k \), we may neglect any further information about earlier states in making a probability statement about \( E_{k+1} \).

**Definition 3:** A sequence of trials with possible outcomes (states) \( E_1, E_2, \ldots \) is called a Markov chain provided that the probabilities of sample sequences are given by

\[
P\{(E_{j0}, E_{j1}, \ldots, E_{jn})\} = a_{j0} p_{j0j1} p_{j1j2} \cdots p_{jn-1jn}
\]

in terms of an initial probability distribution \( \{a_k\} \) for the states \( E_k \) at time zero and transition probabilities \( p_{jk} = P\{E_k \mid E_j\} \) (i.e., the probability that \( E_k \) occurs, given that \( E_j \) occurred on the preceding trial).

Suppose we let \( p_{jk}^{(n)} \) designate the probability that \( E_k \) occurs on the \( n^{th} \) trial after \( E_j \) occurred. Thus we see that \( p_{jj}^{(n)} = v_n \) if a recurrent event occurs on the zeroth trial. \( (E = E_j) \).

**Definition 4:** A Markov chain is irreducible provided that for all \((j,k)\) there exists an \( n \) such that \( p_{jk}^{(n)} > 0 \). (Every state can be reached from every other state).
A state $E_j$ is said to be periodic with period $t > 1$ if $p_{jj}^{(n)} = 0$ whenever $n$ is not divisible by $t$ and $t$ is the smallest integer with this property.

**Definition 5:** An aperiodic Markov chain is a Markov chain in which no states are periodic.

**Definition 6:** A state $E_j$ of an aperiodic Markov chain is a transient state provided that $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$.

**Definition 7:** A state $E_j$ of an aperiodic Markov chain is a persistent null state provided that $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and $\lim_{n \to \infty} p_{jj}^{(n)} = 0$.

**Definition 8:** A state $E_j$ of an aperiodic Markov chain is an ergodic state provided that it is neither transient nor null.

**Theorem 2:** If a state $E_j$ is ergodic, then

$$\lim_{n \to \infty} p_{ij}^{(n)} = f_{ij} = \sum_{n=1}^{\infty} n f_j^{(n)}$$

where $f_{ij}$ is the probability that, starting from state $E_i$, the system ever reaches state $E_j$, and $f_j^{(n)}$ is the probability that state $E_j$ is reached for the first time on the $n^{th}$ trial. The probability $f_j^{(n)}$ plays the role of $f_i$ defined on page 9.

**Proof:** Let $h_{ij}^{(n)}$ represent the probability that, starting from state $E_i$, the system reaches state $E_j$ for the first time on the $n^{th}$ step.

Clearly, $f_{ij} = \sum_{n=1}^{\infty} h_{ij}^{(n)}$. Also,
where we define \( f_j(0) = 0, h_{ij}(0) = 1 \). Thus

\[
\sum_{n=0}^{\infty} p_{ij}(n)s^n = \sum_{n=0}^{\infty} h_{ij}(n)s^n + \sum_{n=0}^{\infty} f_j(n)s^n \sum_{n=0}^{\infty} p_{ij}(n)s^n.
\]

We now define the following generating functions:

\[
P_i(s) = \sum_{n=0}^{\infty} p_{ij}(n)s^n,
\]

\[
F(s) = \sum_{n=0}^{\infty} f_j(n)s^n, \quad \text{and}
\]

\[
H_i(s) = \sum_{n=0}^{\infty} h_{ij}(n)s^n.
\]

By comparison with the geometric series these series converge at least in the open interval \((-1, 1)\). Also \(|F(s)| < 1\) on the open interval \((-1, 1)\).

\[
P_i(s) = H_i(s) + F(s)P_i(s) \quad \text{or} \quad P_i(s) = H_i(s)\left[1 - F(s)\right]^{-1}.
\]

Since \(F(s)\) has a power series expansion and \(|F(s)| < 1\) on \((-1, 1)\) we can write \(\left[1 - F(s)\right]^{-1} = K(s) = \sum_{n=0}^{\infty} k(n)s^n\). Rewriting what we have,

\[K(s) = 1 + F(s)K(s),\]

and in particular, equating coefficients of \(s^n\),

\[k(n) = f_j(1)k(n-1) + f_j(2)k(n-2) + \ldots + f_j(n)k(0)\].

By Theorem 1 we have, then,

\[\lim_{n \to \infty} k(n) = 1/\sum_{n=1}^{\infty} n f_j(n)\]

Since

\[P(s) = H(s)\left[1 - F(s)\right]^{-1} = H(s)K(s),\]
we have

\[ p_{ij}(n) = k(n)h_{ij}(o) + k(n-1)h_{ij}(1) + \ldots + k(o)h_{ij}(n). \]

For my fixed \( r \),

\[ \lim_{n \to \infty} k(n-r)h_{ij}(r) = h_{ij}(r)/\sum_{n=1}^{\infty} n f_j(n), \]

and

\[ \lim_{n \to \infty} (k(n)h_{ij}(o) + \ldots + k(n-N)h_{ij}(N)) \]

\[ = \sum_{n=1}^{N} h_{ij}(n)/\sum_{n=1}^{\infty} n f_j(n). \]

Therefore, given \( \epsilon > 0 \) there exists an \( N \) large enough so that

\[ |k(n)h_{ij}(o) + \ldots + k(n-N)h_{ij}(N) - \sum_{n=1}^{N} h_{ij}(n)/\sum_{n=1}^{\infty} n f_j(n)| < \epsilon/3. \]

Further, since \( \sum_{n=0}^{\infty} h_{ij}(n) < \infty \) and the \( k(n) \) are bounded we have, for large enough \( N \),

\[ \sum_{n=N+1}^{\infty} k(n-r)h_{ij}(r) < \epsilon/3, \]

or, in other words,

\[ |p_{ij}(n) - (k(n)h_{ij}(o) + \ldots + k(n-N)h_{ij}(N))| < \epsilon/3. \]

If we further choose \( N \) sufficiently large that

\[ |H(1)/\sum_{n=1}^{\infty} n f_j(n) - \sum_{n=1}^{N} h_{ij}(n)/\sum_{n=1}^{\infty} n f_j(n)| < \epsilon/3, \]

we have, for \( n > N \)

\[ |p_{ij}(n) - H(1)/\sum_{n=1}^{\infty} n f_j(n)| < \epsilon. \]

Therefore,

\[ \lim_{n \to \infty} p_{ij}(n) = f_{ij}/\sum_{n=1}^{\infty} n f_j(n). \]
Theorem 3: In an irreducible aperiodic Markov chain all states belong to the same one of the three classes defined above.

Proof: Let $E_j$ be a fixed non-transient state and let $E_k$ be some other state that can be reached from it in no less than $N$ steps, and let $p_{jk}^{(N)} > 0$. A return from $E_k$ to $E_j$ must have positive probability since the chain is irreducible. That is, for some $M$, $p_{kj}^{(M)} > 0$. Clearly

$$p_{jk}^{(N)} p_{kk}^{(n)} p_{kj}^{(M)} \leq p_{jj}^{(N+n+M)}$$

and

$$p_{kj}^{(M)} p_{jj}^{(n)} p_{jk}^{(N)} \leq p_{kk}^{(M+n+N)}$$

for all $n$.

Therefore, $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ implies $\sum_{n=1}^{\infty} p_{kk}^{(n)} = \infty$, and

$$\lim_{n \to \infty} p_{jj}^{(n)} = 0 \implies \lim_{n \to \infty} p_{kk}^{(n)} = 0.$$ 

Hence if $E_j$ is persistent null, so is $E_k$.

Suppose $E_j$ is a transient state. Then any other state, say $E_k$, that can be reached from it must be transient also, for if it were not, by the above, $E_j$ would have to be non-transient also.

Since $E_j$ persistent null implies any other state $E_k$ is persistent null, and $E_j$ transient implies any other state $E_k$ is transient, it follows that $E_j$ ergodic implies any other state $E_k$ is ergodic.

Definition 9: A probability distribution $\{v_k\}$ is called stationary with respect to $\{p_{ij}\}$ provided that $v_j = \sum_{i=1}^{\infty} v_i p_{ij}$.

Theorem 4: An irreducible aperiodic Markov chain belongs to one of the following two classes:

1) The states are all transient or all null states, in which case there exists no stationary distribution.
2) The states are all ergodic, in which case \( \{ u_k \} \) is a unique stationary distribution where \( u_k = \lim_{n \to \infty} p_{jk}^{(n)} > 0 \).

Proof: Suppose all states are ergodic. Then for fixed \( j \) and \( n \),

\[
\sum_{k=1}^{\infty} p_{jk}^{(n)} = 1 \quad \text{so that} \quad \sum_{k=1}^{N} u_k \leq 1.
\]

In the equation

\[
p_{jk}^{(m+1)} = \sum_{v=1}^{\infty} p_{jv}^{(m)} p_{vk}
\]

let \( m \) approach infinity. Then \( \lim_{m \to \infty} p_{jk}^{(m+1)} = u_k \) so that

\[
\lim_{m \to \infty} p_{jv}^{(m)} p_{vk} = p_{vk} \lim_{m \to \infty} p_{jv}^{(m)} = p_{vk} u_v.
\]

For all finite \( t \) we have

\[
p_{jk}^{(m+1)} = \sum_{v=1}^{\infty} p_{jv}^{(m)} p_{vk} \geq \sum_{v=1}^{t} p_{jv}^{(m)} p_{vk}.
\]

Letting \( m \) approach \( \infty \),

\[
u_k = \sum_{v=1}^{t} u_v p_{vk}^*\]

and

\[
u_k = \sum_{v=1}^{\infty} u_v p_{vk}^*.
\]

We suppose that

\[
u_k > \sum_{v=1}^{\infty} u_v p_{vk}
\]

and sum both sides over \( k \) obtaining

\[
\sum_{k=1}^{\infty} u_k > \sum_{k=1}^{\infty} \sum_{v=1}^{\infty} u_v p_{vk}^*.
\]

Since we have absolute convergence we can interchange the order.
of summation so that
\[ \sum_{k=1}^{\infty} u_k > \sum_{v=1}^{\infty} \sum_{k=1}^{\infty} p_{vk} u_v = \sum_{v=1}^{\infty} u_v, \]
a contradiction. Hence
\[ u_k = \sum_{v=1}^{\infty} u_v p_{vk}. \]

Let
\[ v_k = u_k / \sum_{j=1}^{\infty} u_j. \]

Then
\[ v_k = \sum_{j=1}^{\infty} u_j p_{jk} / \sum_{j=1}^{\infty} u_j \]
\[ = \sum_{j=1}^{\infty} (u_j / \sum_{j=1}^{\infty} u_j) p_{jk} \]
\[ = \sum_{j=1}^{\infty} v_j p_{jk} \]
so that \( \{v_k\} \) is a stationary distribution.

Let \( \{v_k\} \) be any stationary distribution. Then,
\[ v_j = \sum_{i=1}^{\infty} v_i p_{ij} \]
and
\[ \sum_{j=1}^{\infty} p_{jk} (1) v_j = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{jk} (1) v_i p_{ij} \]
so that
\[ v_k = \sum_{i=1}^{\infty} v_i p_{ik} (2). \]

We proceed, inductively, by supposing
\[ v_j = \sum_{i=1}^{\infty} v_i p_{ij} (m). \]
Then
\[ \sum_{j=1}^{\infty} p_{jk} v_j = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} p_{ijk} v_i u_j \]

Absolute convergence allows us to interchange limits so that we have

\[ v_k = \sum_{i=1}^{\infty} v_i u_j \]

Letting \( n \) approach infinity in

\[ v_r = \sum_{v=1}^{\infty} v_v u_r = u_r \]

Hence the distribution is unique.

If the states are transient or null states and \( \{ v_k \} \) is a stationary distribution, then

\[ v_r = \sum_{v=1}^{\infty} v_v p_{vr} \]

and \( \lim_{n \to \infty} p_{vr}^{(n)} = 0 \), so that no stationary distribution exists.

Example: Three chess players, Adams, Berlyov, and Schultz have a tournament in which the last player to win a game is champion. Adams and Schultz don't get on together so they never play one another. Therefore, whenever Adams or Schultz is champion he remains champion until dethroned by Berlyov. Berlyov, however, never wins two games in succession. Whenever Adams is champion, Berlyov has probability three-fourths in favor of defeating him. Whenever Berlyov has the choice of opponents he chooses to play Adams three-fourths of the time. Whenever Schultz is champion, Berlyov regains the championship only one-fourth of the time.

If we let the integers \( a, b, \) and \( s \) represent the states of Adams,
Berlyov, and Schultz, respectively, being champion, we may write the transition matrix as follows:

\[
\begin{bmatrix}
a & b & s \\
1/4 & 3/4 & 0 \\
3/4 & 0 & 1/4 \\
0 & 1/4 & 3/4
\end{bmatrix}
= \begin{bmatrix}
P_{aa} & P_{ab} & P_{as} \\
P_{ba} & P_{bb} & P_{bs} \\
P_{sa} & P_{sb} & P_{ss}
\end{bmatrix}
\]

This is an example of a Markov chain in which all states are ergodic. It is easy to verify that the vector whose entries are \( u_k \) is

\[
a \ b \ c \\
(1/3 \ 1/3 \ 1/3).
\]
WE adopt the following notation:

\[ T = \{0, 1, 2, \ldots\} \]

1 represents some fixed set of integers, and \( A \) denotes a Markov chain with random variables \( A_k \) and the \( A_k \) take on values in \( I \). Initial probabilities will be denoted by \( P\{A_0 = i\} \) and transition probabilities by

\[ P_{im} = P\{A_{k+1} = m | A_k = i\} \]

We would like to introduce the notion of random waiting times in successive states.

We assume that the random variables \( (A_k, B_k), k \in T \), define a Markov chain \( A' \) with transition probabilities

\[ P\{A_{k+1} = m, B_{k+1} = n | A_k = i, B_k = j\} = p_{im} a_{mn}. \]

**Definition 10:** For any fixed sequence of events in \( A' \), we define \( t_k \) by \( t_0 = 0 \) and recursively by \( t_{k+1} = t_k + B_k \).

The \( B_k \)'s are interpreted as waiting times in the states \( A_k \). For any sequence of events we interpret the successive elements of \( T \) as "running time," and label the time of the \( k^{\text{th}} \) transition in \( A \) by \( t_k \). That is, we label the time of the transition from \( A_{k-1} \) to \( A_k \) by \( t_k \).

Clearly

\[ 0 = t_0 < t_1 < \cdots < t_k < t_{k+1} < \cdots \]

We have the following conditions:

\[ P\{B_k = j | A_k = i\} = a_{ij}, \]
\[
\sum_{j=1}^{\infty} a_{ij} = 1, \\
0 \leq a_{ij}, \text{ and} \\
a_{10} = 0.
\]

The following theorem shows us that the introduction of waiting times still leaves us with a Markov chain.

**Theorem 5:** \[P\{A_{k+1} = m \mid A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = j\} = \left\{P \ A_{k+1} = m \mid A_k = i\right\}.\]

(Note that \(A_0, A_1, \ldots, A_k = i\) is an abbreviation of \(A_0 = i_0, A_1 = i_1, \ldots, A_k = i\).)

**Proof:** \[P\{A_{k+1} = m \mid A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = j\} = \sum_{n=1}^{\infty} P\{A_{k+1} = m, B_{k+1} = n \mid A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = j\} = \sum_{n=1}^{\infty} P\{A_{k+1} = m, B_{k+1} = n \mid A_k = i, B_k = j\} = \sum_{n=1}^{\infty} p_{im} a_{mn} = P_{im} = P\{A_{k+1} = m \mid A_k = i\}.\]

**Theorem 6:** \[P\{B_{k+1} = m \mid A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = s\} = P\{B_{k+1} = m \mid A_{k+1} = i\}.\]

**Proof:** \[P\{B_{k+1} = m \mid A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = s\} = \frac{P\{B_{k+1} = m, A_{k+1} = i, (A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = s)\}}{P\{A_{k+1} = i \mid (A_0, A_1, \ldots, A_k = i, B_0, B_1, \ldots, B_k = s)\}} = \frac{P\{B_{k+1} = m, A_{k+1} = i \mid (A_0, A_1, \ldots, A_k = j, B_0, B_1, \ldots, B_k = s)\}}{P\{A_{k+1} = i \mid A_k = j\}} = \frac{P\{A_{k+1} = i, B_{k+1} = m \mid A_k = j, B_k = s\}}{P\{A_{k+1} = i \mid A_k = j\}} \cdot \]
\[ P_{ji} a_{im} / p_{ji} = a_{im} = P \{ B_k = m | A_k = i \} \]

**Definition 11**: For any fixed sequence of events in \( A \), the sequence \( \{ x_t \} \), \( t \in T \), defined by \( x_t = A_k \) provided that \( t_k \leq t < t_{k+1} \), is called a semi-Markov chain.

**Theorem 7**: For any subset \( I_0 \) of \( I \)

\[ \lim_{t \to \infty} P \{ x_t \in I_0 | x_0 = 1 \} = \lim_{k \to \infty} P \{ A_k \in I_0 | A_0 = 1 \} \]

**Proof**: Let \( \xi_t = P \{ x_t \in I_0 | x_0 = 1 \} \) and \( \alpha_k = P \{ A_k \in I_0 | A_0 = 1 \} \), and note that \( \{ \alpha_k \} \) is a subsequence of \( \{ \xi_t \} \). Therefore, if \( \{ \xi_t \} \) converges to \( L \), then \( \{ \alpha_k \} \) converges to \( L \).

Suppose that \( \{ \alpha_k \} \) converges to \( L \). That is, for all \( \epsilon > 0 \), there exists a positive integer \( N \) such that \( N \leq k \) implies that \( | \alpha_k - L | < \epsilon \). If \( \xi_t = \alpha_k \), then \( \min \alpha_j \leq \xi_t \leq \max \alpha_j \), whenever \( k \leq j \) and \( t \leq i \), so that \( N \leq k \leq t \leq i \) implies that \( | \xi_i - L | < \epsilon \). Hence, if \( \{ \alpha_k \} \) converges to \( L \), then \( \{ \xi_t \} \) converges to \( L \).

Therefore \( \lim_{t \to \infty} \xi_t = \lim_{k \to \infty} \alpha_k \).

**Definition 12**: For a fixed sequence \( \{ A_k \} \) we define a substate chain \( \{ (x_t, y_t) \} \), \( t \in T \), where \( y_t \) is given by:

1) \( y_t = 0 \) if \( t = t_k \), and
2) \( y_t = t_{k+1} - t \) if \( t_k < t < t_{k+1} \).

This means that if \( y_t = 0 \), then the semi-Markov process has just reached \( x_t \). If \( y_t = m > 0 \), then \( m \) represents the time the process will remain in \( x_t \).

The transition probabilities for the substate chain are given by

\[ P \{ x_{t+1} = m, y_{t+1} = n | x_t = i, y_t = j \} = q_{ijmn} \]
where $i, m \in I$ and $j, n \in T$. We note the following cases.

1. If $j = n = 0$, then $q_{ijmn} = a_{ij+1} p_{im}$. This is simply the probability of waiting in state $i$ for one unit of time times the probability of going from $i$ to $m$.

2. If $j = 1$ and $n = 0$, then $q_{ijmn} = p_{im}$. Knowing that this is the last unit of waiting time before a change of state, $j = 1$, we have the probability of going from $i$ to $m$.

3. If $i = m$ and $n \geq 1$ and $j = 0$, then $q_{ijmn} = a_{i,n+1}$. Given that the process is in state $i = m$, this is the probability that it waits here $n + 1$ units of time.

4. If $j \geq 2$ and $n = j - 1$ and $i = m$, then $q_{ijmn} = 1$. If the process is going to wait in state $i = m$ $j \geq 2$ times clearly the next state of the substate chain is $(i, j - 1)$.

5. Otherwise, $q_{ijmn} = 0$.

6. $P\{x_0 = i, y_0 = j\} = \left\{ \begin{array}{ll} P\{x_0 = i\} & \text{if } j = 0 \\ 0 & \text{if } j > 0 \end{array} \right\}$.

Since the transition probabilities for the substate chain are well defined in terms of probabilities that depend only upon the previous state, the substate chain is a Markov chain.

Example: Turning to our chess players again, suppose we introduce the following complication: Berlyov never plays more than one game in a day, and on any given day he is as likely to play as not.

A state in the substate chain might be the event that Schultz wins the championship by defeating Berlyov on the eighth day after the tournament began. This is represented by $(x_8, y_8) = (s, o)$. The probability that he will be champion the following day only is
represented by
\[ P \{ (s,1) \mid (s,0) \} = q_{s0s1} = a_{s1} P_{ss} = (1/2)(3/4) = 3/8. \]
The probability that Berlyov regains the championship the day following
that is represented by
\[ P \{ (b,0) \mid (s,1) \} = q_{s1b0} = p_{sb} = 1/4. \]
The probability that Berlyov will then retain the championship for
three days is represented by
\[ P \{ (b,3) \mid (b,0) \} = q_{b0b3} = a_{bl} = (1/2)^3 = 1/16. \]
Letting \( i \in I \) and \( j \in T \) we adopt the following notation:
\[ b_{ij} = \sum_{n>j} a_{in}, \]
\[ e_i = E \{ B_k \mid A_k = i \} = \sum_{j=0}^{\infty} b_{ij} = \sum_{j=0}^{\infty} j a_{ij}, \]
\[ m_i = \lim_{k \to \infty} P \{ A_k = i \}, \]
\[ m_{ij}(t) = P \{ (x_t,y_t) = (i, j) \}, \]
\[ m_{ij} = \lim_{t \to \infty} P \{ (x_t,y_t) = (i, j) \}, \]
and
\[ I^* = \{ (i, j) \mid i \in I, b_{ij} > 0 \}. \]
Notice that
\[ P \{ (x_t,y_t) \in I^* \mid t \in T \} = 1. \]
Recall that \( A \) represents the Markov chain defined on page 21. We let
\( A^* \) represent the substate chain where \( (x_t,y_t) \in I^* \).

**Theorem 8:**
1) \( A^* \) non-null implies that \( \{ m_{ij} \mid (i, j) \in A^* \} \) is the
unique solution of the system
\[ v_{ij} > 0, \sum_{i,j=1}^{\infty} v_{ij} = 1, \sum_{i,j=1}^{\infty} v_{ij} q_{ijmn} = v_{mn}. \]
2) \( A^* \) null implies that this system has no solution.
Proof: This is a consequence of Theorem 4.

**Lemma 4:** If \( \lim_{t \to \infty} b_t = b \), and \( \sum_{n=0}^{\infty} a_n = a \), then \( \lim_{t \to \infty} \sum_{n=0}^{t} a_n b_{t-n} = ab \).

**Proof:** Given \( \epsilon > 0 \), we wish to show that if \( t \) is large enough then
\[
\left| \sum_{n=0}^{t} a_n b_{t-n} - ab \right| < \epsilon.
\]

We note that
\[
\left| \sum_{n=0}^{t} a_n b_{t-n} - ab \right| = \left| \sum_{n=0}^{t} a_n b_{t-n} - b \sum_{n=0}^{\infty} a_n \right|
\]
\[
= \left| \sum_{n=0}^{t} a_n (b_{t-n} - b) - b \sum_{n=t+1}^{\infty} a_n \right|
\]
\[
\leq \sum_{n=0}^{t} |a_n| |b_{t-n} - b| + b \sum_{n=t+1}^{\infty} |a_n|
\]
\[
= \sum_{n=0}^{N} |a_n| |b_{t-n} - b| + \sum_{n=N+1}^{\infty} |a_n| |b_{t-n} - b| + b \sum_{n=t+1}^{\infty} a_n.
\]

We choose \( N \) sufficiently large so that
\[
\sum_{n=N+1}^{t} |a_n| |b_{t-n} - b| < \frac{\epsilon}{3} \text{ if } t > N.
\]

This can be done since \( |b_{t-n} - b| \) is bounded and \( \sum_{n=0}^{\infty} |a_n| \) is convergent.

The choice of \( N \) is independent of \( t \) as long as \( t > N \).

Next we choose \( N_1 \geq N \) sufficiently large so that
\[
\left| b \sum_{n=t+1}^{\infty} a_n \right| < \frac{\epsilon}{3} \text{ if } t > N_1,
\]

and also large enough so that \( t - N \) is sufficiently large to insure that
\[
\sum_{n=0}^{N} |a_n| |b_{t-n} - b| < \frac{\epsilon}{3}.
\]
This can be done since \(|b_{t-n} - b|\) converges to zero for each fixed \(n\).

Therefore we have for the choice of \(t\)

\[
\left| \sum_{n=0}^{t} a_n b_{t-n} - ab \right| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.
\]

**Theorem 9:** \(m_{ij} = b_{ij} m_{i0}\).

**Proof:** Since

\[
P \{(x_t, y_t) = (i, j)\}
= P \{(x_{t-1}, y_{t-1}) = (i, j+1)\} P \{(x_t, y_t) = (i, j)\} P \{(x_{t-1}, y_{t-1}) = (i, j)\}
+ P \{(x_{t-1}, y_{t-1}) = (i, 0)\} P \{(x_t, y_t) = (i, j)\} P \{(x_{t-1}, y_{t-1}) = (i, 0)\}
= m_{i} j+1(t - 1) q_{i} j+1 i j + m_{i0} t - 1 q_{i0} i j
= m_{i} j+1(t - 1) + a_{i} j+1 m_{i0}(t - 1)
\]

for \(1 \leq j \leq t\), we have

\[
m_{ij}(t) = m_{i} j+1(t - 1) + a_{i} j+1 m_{i0}(t - 1).
\]

We note that

\[
m_{ij}(1) = m_{i} j+1(0) + a_{i} j+1 m_{i0}(0)
\]

where

\[
m_{i} j+1(0) = P (x_0, y_0) = (i, j+1) = 0.
\]

Since

\[
\lim_{t \to \infty} m_{i0}(t) = m_{i0}, 0 \leq m_{i0}(t) \leq 1 \text{ for all } t \in T, \text{ and}
\]

\[
\sum_{n=1}^{\infty} a_{i} j+n = b_{ij} \leq 1,
\]

we have, by Lemma 4,

\[
\lim_{t \to \infty} m_{ij}(t) = \lim_{t \to \infty} \sum_{n=1}^{t} a_{i} j+n m_{i0}(t - n) = b_{ij} m_{i0}.
\]
Theorem 10: If $A^*$ is non-null, then $\{m_{ij} | i \in I\}$ is the unique solution of the system

$$v_i > 0, \sum_{i=1}^{\infty} e_i v_i = 1, \sum_{i=1}^{\infty} v_i p_{im} = v_m.$$ 

There is no solution of $A^*$ is null.

Proof: Since $0 < m_{ij}$ and $1 \leq b_{ij}$, we have $0 < m_{i0}$.

Also

$$\sum_{i=1}^{\infty} e_i v_i = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} b_{ij} v_i = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} m_{ij} = 1.$$ 

Suppose $n = 0$, then

$$\sum_{i=1}^{\infty} \sum_{j=0}^{\infty} m_{ij} q_{ijmn} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_{ij} m_{i0} q_{ijm0}$$

$$= \sum_{i=1}^{\infty} b_{i0} m_{i0} a_{i1} p_{im} + \sum_{i=1}^{\infty} b_{i0} m_{i0} a_{i1} p_{im}$$

$$= \sum_{i=1}^{\infty} (a_{i1} + b_{i1}) m_{i0} p_{im}$$

$$= \sum_{i=1}^{\infty} b_{i0} m_{i0} p_{im}$$

$$= \sum_{i=1}^{\infty} m_{i0} p_{im}.$$ 

By Theorem 7 we have

$$\sum_{i=1}^{\infty} \sum_{j=0}^{\infty} m_{ij} q_{ijmn} = m_{m0}.$$ 

This solution is unique since another solution would contradict the uniqueness of

$$\sum_{i=1}^{\infty} \sum_{j=0}^{\infty} m_{ij} q_{ijmn}.$$
Theorem 11: Suppose $A$ (and therefore $A^*$) is aperiodic.

1) A non-null and $A^*$ non-null imply that

$$m_{ij} = b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n, \quad (i, j) \in \mathbb{I}^*.$$ 

2) A non-null and

$$\sum_{n=1}^{\infty} e_n m_n < \infty$$

imply that $A^*$ is non-null.

3) A non-null and $A^*$ null imply that

$$m_{ij} = \lim_{n \to \infty} b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n, \quad (i, j) \in \mathbb{I}^*.$$ 

4) A non-null and

$$\lim_{n \to \infty} \sum_{n=1}^{\infty} e_n m_n = \infty$$

imply that $A^*$ is null.

5) A null implies that $A^*$ is null.

Proof: 1) A non-null and $m_1 = \lim_{k \to \infty} P\{A_k = 1\}, \quad i \in \mathbb{I}$ imply that $m_1 \neq 0.$

$A^*$ non-null and $m_{10} = \lim_{t \to \infty} P\{(x_t, y_t) = (1, 0)\}$ imply that $m_{10} \neq 0.$

Hence there exists a $\lambda > 0$ such that $m_1 = \lambda m_{10}.$ Since

$$\sum_{n=1}^{\infty} e_n m_{n0} = 1, \quad \lambda = \lambda \sum_{n=1}^{\infty} e_n m_{n0}.$$ 

Hence

$$m_1 = (\lambda \sum_{n=1}^{\infty} e_n m_{n0}) m_{10} = m_{ij} / b_{ij} \sum_{n=1}^{\infty} e_n (\lambda m_{n0}).$$

That is,

$$m_{ij} = b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n.$$
2) Since $A$ non-null implies that $m \neq 0$ we have that

$$m_i/\sum_{n=1}^{\infty} e_n m_n > 0.$$  

Also

$$\sum_{i=1}^{\infty} e_i (m_i/\sum_{n=1}^{\infty} e_n m_n) = \sum_{i=1}^{\infty} e_i m_i / \sum_{n=1}^{\infty} e_n m_n = 1,$$

and

$$\sum_{i=1}^{\infty} (m_i/\sum_{n=1}^{\infty} e_n m_n) p_i = m_n / \sum_{n=1}^{\infty} e_n m_n.$$  

Hence $A^*$ is non-null.

3) By 2) $A^*$ null implies that

$$\lim_{n \to \infty} \sum_{n=1}^{\infty} e_n m_n = \infty.$$  

Hence, since $b_{ij} \neq 0$ and $m_i \neq 0$, we have

$$\lim_{n \to \infty} b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n = 0.$$  

Therefore,

$$m_{ij} = \lim_{n \to \infty} b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n = 0.$$  

4) By 3) and 1) we have

$$m_{ij} = \lim_{n \to \infty} b_{ij} m_i / \sum_{n=1}^{\infty} e_n m_n = 0$$

so that $A^*$ is null.

5) Suppose $A^*$ is non-null. We have that $1 \leq e_i$ implies

$$0 < \sum_{n=1}^{\infty} m_{n0} \leq \sum_{n=1}^{\infty} e_n m_{n0} = 1.$$  

Hence, for all $i$

$$0 < m_{i0} / \sum_{n=1}^{\infty} m_{n0}.$$
Also
\[ \sum_{i=1}^{\infty} m_{i0} / \sum_{n=1}^{\infty} m_{n0} = \sum_{i=1}^{\infty} m_{i0} / \sum_{n=1}^{\infty} m_{n0} = 1, \]
and
\[ \sum_{i=1}^{\infty} (m_{i0} / \sum_{n=1}^{\infty} m_{n0}) P_{in} = m_n / \sum_{n=1}^{\infty} m_{n0}. \]

Hence A is non-null.

Since \( m_{ij}(t) = P\{ (x_t, y_t) = (i, j) \} \),
\[ P\{ x_t = i \} = \sum_{j=0}^{\infty} m_{ij}(t). \]

Also, \( b_{10} = 1 \) and
\[ b_{ij} = \sum_{n > j} a_{in} \]
and
\[ m_{ij}(t) = \sum_{n=1}^{t} a_i j+n m_{i0}(t-n), \quad 1 \leq j, \quad 1 \leq t, \]
give us
\[ \sum_{j=0}^{\infty} m_{ij}(t) = \sum_{n=0}^{t} b_{in} m_{i0}(t-n). \]

**Theorem 12:** Suppose A is aperiodic and \( \epsilon_i < \infty \).

1) A null implies that
\[ \lim_{t \to \infty} P\{ x_t = i \} = 0 = \sum_{j=0}^{\infty} m_{ij}. \]

2) A non-null and
\[ \lim_{n \to \infty} \sum_{n=1}^{\infty} e_n m_n = \infty \]
imply that
\[ \lim_{t \to \infty} P\{ x_t = i \} = 0 = \sum_{j=0}^{\infty} m_{ij}. \]
3) A non-null implies that

\[ \lim_{t \to \infty} P\{x_t = i\} = \frac{e_{i}}{\sum_{n=1}^{\infty} e_{n} m_{n}} = \sum_{j=0}^{\infty} m_{ij}. \]

Proof:

1) For all \( k, 0 \leq m_{10}(k) \leq 1 \) and \( \lim_{k \to \infty} m_{10}(k) = m_{10} \) and \( \sum_{n=0}^{\infty} b_{in} = e_{i} \)

imply that

\[ \lim_{t \to \infty} \sum_{n=0}^{\infty} b_{in} m_{10}(t - n) = e_{i} m_{10} \]

by Lemma 3. Also, A null implies \( A^{*} \) is null and \( m_{ij} = 0 \) for all \( j \geq 0 \)

so that

\[ \sum_{j=0}^{\infty} m_{ij} = 0. \]

Hence

\[ \sum_{j=0}^{\infty} b_{ij} m_{10} = 0 \]

implies that \( e_{i} m_{10} = 0 \). Therefore, we have

\[ \lim_{t \to \infty} P\{x_t = i\} = \lim_{t \to \infty} \sum_{n=0}^{t} b_{in} m_{10}(t - n) = e_{i} m_{10} = 0. \]

2) A non-null and

\[ \lim_{t \to \infty} \sum_{n=1}^{t} e_{n} m_{n} = \infty \]

imply that \( A^{*} \) is null and the above argument applies.

3) A non-null implies

\[ m_{ij} = \frac{b_{ij} m_{i}}{\sum_{n=1}^{\infty} e_{n} m_{n}} \]

so that

\[ \sum_{j=0}^{\infty} m_{ij} = \left( \frac{m_{i}}{\sum_{n=1}^{\infty} e_{n} m_{n}} \right) \sum_{j=0}^{\infty} b_{ij} = e_{i} m_{i} / \sum_{n=1}^{\infty} e_{n} m_{n}. \]
**Example:** Consider the example of the chess tournament. Suppose we wish to know the probability that Schultz is champion for three days. We make the following computations.

\[
\lim_{t \to \infty} P\{ (x_t, y_t) = (s, 2) \} = m_{s2},
\]

\[
e_1 = \sum_{j=0}^{\infty} j(1/2^{j+1}) = 1,
\]

\[
b_{s2} = \sum_{n>2} 1/2^{n+1} = 1 - 1/2 - 1/4 - 1/8 = 1/8,
\]

\[
\sum_{n=1}^{\infty} e_n m_n = \sum_{n=1}^{\infty} m_n = 1/3 + 1/3 + 1/3 = 1,
\]

so that

\[
m_{s2} = b_{s2} m_s / \sum_{n=1}^{\infty} e_n m_n = 1/24.
\]
REFERENCES CITED


