DEDICATION

To my parents

Georgia and Anastasio

Στοὺς γονεῖς μου,

Georgia καὶ Anastasio
ACKNOWLEDGEMENTS

The author would like to express his gratitude to his thesis supervisor, Dr. Abraham Boyarsky of Concordia University. Dr. Boyarsky's encouragement and guidance were most helpful.
TABLE OF CONTENTS

Page

TITLE PAGE............................................................. i
SIGNATURE PAGE..................................................... ii
ABSTRACT.............................................................. iii
DEDICATION........................................................... iv
ACKNOWLEDGEMENTS.................................................. v
TABLE OF CONTENTS.................................................. vi

Chapter

I. INTRODUCTION....................................................... 1
   1.1. Markov Maps................................................ 1
   1.2. The Outline of the Thesis................................... 5

II. FROBENIUS-PERRON OPERATOR AND NON-NEGATIVE MATRICES........... 7
   2.1. Representation of the Frobenius-Perron Operator
       for Markov Maps........................................... 7
   2.2. Properties of $M_t$.......................................... 15

III. INVARIANT DENSITIES FOR A CLASS OF PIECEWISE
     LINEAR TRANSFORMATIONS....................................... 25
   3.1. The Perron-Frobenius Theorem for Irreducible
       Matrices.................................................. 25
   3.2. The Matrix $M_t$ and its Eigenfunctions.................... 33
TABLE OF CONTENTS (continued)

Chapter  

IV. MATRICES INDUCED BY PIECEWISE LINEAR TRANSFORMATION AND THEIR IN Variant VECTORS  .......................................................... 44  

4.1. Preliminaries ......................................................... 44  

4.2. Elements from Graph Theory ..................................... 45  

Directed Graphs ....................................................... 45  

Rooted Trees .......................................................... 46  

Matrices and Graphs .................................................. 47  

4.3. Permutation Invariant Matrices ................................. 50  

Induced 0-1 Matrices ................................................ 50  

A Special Labelling Procedure .................................... 57  

Characterization of $P_n$-Invariant Vectors .................. 62  

Examples .................................................................... 69  

4.4. Extension of $P_n$-Invariant Vectors ......................... 73  

4.5. A Special Construction ........................................... 79  

4.6. Concatenation Results .......................................... 87  

4.7. Permutation Results .............................................. 99

V. MARKOV MAPS AND IRREDUCIBLE (OR PRIMITIVE) MATRICES ........................................................................... 102  

5.1. Preliminaries ......................................................... 102  

5.2. Matrix Representation of a Markov Map .................. 104  

5.3. Irreducibility and Primitivity of Transition Matrices .... 110

REFERENCES .................................................................. 130
CHAPTER I

INTRODUCTION

1.1. MARKOV MAPS

A function $f$ is said to be piecewise continuous on an interval $I = [a, b]$ if there exists a finite partition $J = \{I_i\}_{i=1}^n$, $I_i = (a_{i-1}, a_i)$, $a_0 = a$, $a_n = b$ such that:

(i) $f$ is continuous on each open subinterval $I_i$, $1 \leq i \leq n$;

(ii) $f$ approaches a finite limit as the end points of each subinterval are approached from within the subinterval.

In other words, $f$ is piecewise continuous on $[a, b]$ if it is continuous, except for a finite number of jump discontinuities.

A simple example of a piecewise continuous function is the following:

![Graph of a piecewise continuous function](image)

Fig. 1.1

Moreover, $f$ is said to be piecewise $C^k$ if for each $i$, $1 \leq i \leq n$, the restriction $f_i$ of $f$ on the open interval $(a_{i-1}, a_i)$ is a $C^k$. 


function which can be extended to the closed interval \([a_{i-1}, a_i]\), need not be continuous at the partition points \(a_i\).

In the following we shall use the terminology "map" or "transformation", instead of a function.

**Definition 1.1.1:** Let \(I = [a, b]\) be an interval on the real line and let us consider a finite partition \(J\) of \(I\). Let us denote by \(Q\) the partition points of \(J\). We say that a piecewise continuous transformation \(\tau: I \to I\) takes partition points into partition points if \(\tau(Q) \subseteq Q\). If \(\tau\) is discontinuous at the partition points \(a_i\), we require \(\tau(a_i^-)\) and \(\tau(a_i^+}\) to be in \(Q\).

The above partition \(J\) is called a Markov partition for \(\tau\), and the transformation \(\tau\) is called a Markov map with respect to the partition \(J\).

**Definition 1.1.2:** Let \(I_i = (a_{i-1}, a_i), 1 \leq i \leq n\). We say that the partition \(J\) has the communication property ("irreducibility property" in Markov chain terminology), under the piecewise continuous transformation \(\tau: I \to I\), if for any \(I_i, I_j \in J\) there exist integers \(p\) and \(q\) such that:

\[
I_i \subseteq \tau^p(I_j) \quad \text{and} \quad I_j \subseteq \tau^q(I_i),
\]

where \(\tau^k\) is the \(k\)th iterate of \(\tau\).

Let \(I = [a, b]\) and let \(\tau: I \to I\) be a measurable non-singular transformation. By "non-singular" we mean that \(m(\tau^{-1}(A)) = 0\) whenever \(m(A) = 0\) for \(A\) a measurable set (\(m\) denotes the Lebesgue measure). 

---

\(^{1}\)A continuous function with continuous derivatives up to \(k\)th order.
measure on $I$ and $\tau^{-1}(A) = \{x \in I : \tau(x) \in A\}$. Also let $\mu$ be a 
measure defined on the $\sigma$-algebra of $I$. We say that $\mu$ is invariant 
under $\tau$ if for all measurable sets $A \subseteq I$ we have $\mu(A) = \mu(\tau^{-1}(A))$. 
$\mu$ is absolutely continuous if there exists a function 
$f: I \to [0, \infty)$, $f \in L_1 = L_1(I)$, the space of integrable functions 
on $I$, such that:

$$\mu(A) = \int_A f(x) \, dx,$$

for every Lebesgue measurable set $A \subseteq I$. We refer to $f$ as the 
invariant density (of $\mu$) under $\tau$.

It is known [14] that if $\inf_{x \in I} |\tau'(x)| > 1$, and $\tau$ is a piecewise 
$C^2$ transformation, then it admits an absolutely continuous invariant 
measure. It is also known [3] that a certain subclass of the 
aforementioned class possesses a unique absolutely continuous invariant 
measure. Moreover, [14], the invariant density $f$ of $\mu$ under $\tau$ 
is a fixed point of the Frobenius-Perron operator $P_\tau: L_1 \to L_1$ 
defined by:

$$P_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}([0,x])} f(s) \, ds.$$ 

On the other hand, it is known [10] that the fixed points of the 
Frobenius-Perron operator $P_\tau$ are piecewise constant functions. 
Therefore, the unique invariant density under $\tau$, for $\tau$ belonging to 
that particular subclass, is a piecewise constant function.

Generally, it is very difficult to find the unique invariant 
density for $\tau$. In this thesis, we shall show that in the case where 
$\tau$ is piecewise linear, the unique invariant density under $\tau$, is found
by simply solving a system of linear equations. Moreover, we shall find the invariant densities of some particular piecewise linear transformations and we shall characterize some classes of piecewise constant functions, which can serve as the unique invariant densities for constructible linear transformations.

Remark: Since the results of the theorems remain valid when we use linear transformations (i) from \([a,b]\) into \([a,b]\), \([a,b]\) any interval on the real line and (ii) from \([0,1]\) into \([0,1]\), we shall use, mostly, linear transformations from \([0,1]\) into \([0,1]\).

Let us now give some matrix theoretic definitions.

Definition 1.1.3: A square matrix \( A = (a_{ij}) \), \( i, j = 1, 2, \ldots, n \), is called non-negative, \( A \geq 0 \) (or positive, \( A > 0 \)) if all the entries of \( A \) are non-negative (or positive).

Definition 1.1.4: A square matrix \( A = (a_{ij}) \), \( i, j = 1, 2, \ldots, n, n \geq 2 \), is called reducible if there is a permutation matrix \( P \) such that

\[
P^T A P = \begin{bmatrix}
A_{11} & A_{12} \\
0 & A_{22}
\end{bmatrix}
\]

where \( A_{11}, A_{22} \) are square matrices of order less than \( n \), and \( T \) denotes transpose. Otherwise, \( A \) is called irreducible.

Definition 1.1.5: A \( n \times n \) square non-negative matrix \( A \) is primitive if there exists an integer \( k > 0 \), such that \( A^k > 0 \), where \( A^k = (a_{ij}^{(k)}) \) denotes the \( k \)th power of \( A \).

\[\text{Permutation matrix is a square matrix where the entries are either zero or one and which has precisely a single 1 in each row and in each column.}\]
Definition 1.1.6: A $n \times n$ square non-negative matrix $A$ is said to be row stochastic if each row sum is equal to 1, i.e.,

$$\sum_{j=1}^{n} a_{ij} = 1, \quad i = 1, 2, \ldots, n.$$ 

Definition 1.1.7: A non-negative matrix $A$ is contiguous if all the positive entries in each row are equal and there are no zero entries between them.

1.2. OUTLINE OF THE THESIS

In Chapter II, we explore the property that the Frobenius-Perron operator $P_\tau$, when restricted to piecewise constant functions, can be represented as a product of a vector and of an $n \times n$ matrix $M_\tau$. The matrix $M_\tau$ is referent as the matrix induced by the transformation $\tau$. We define a certain class of transformations $\mathcal{T}$, and study some properties of $M_\tau$ for $\tau \in \mathcal{T}$. Finally, we prove a theorem (Theorem 2.2.1) which gives necessary and sufficient conditions for an irreducible matrix to be primitive. This result will be used frequently in Chapter IV.

We start Chapter III by proving the famous Perron-Frobenius Theorem (Theorem 3.1.1) and using that we show (Theorem 3.2.1) that the matrix $M_\tau$, induced by a piecewise linear transformation $\tau \in \mathcal{T}$ has 1 as the eigenvalue of maximum absolute value. This permits us to show (Theorem 3.2.2) that it is possible to construct the unique invariant density for a piecewise linear transformation $\tau \in \mathcal{T}$ when $\tau$ satisfies two simple conditions. Moreover, we shall show that it is possible to construct a transformation $\tau$ under which a given function is invariant.
In Chapter IV, we define a class (denoted $A_n$) of non-negative square matrices $A$, which are primitive, row stochastic and contiguous, and we define a 1-1 correspondence between $A_n$ and $g(A_n)$, where $g(A_n)$ is a subspace of the space of 0-1 matrices. Since each $A \in g(A_n)$ induces an equivalent class of transformation $\tau \in \mathcal{T}$ (Chapter II), we are able to find $A$-invariant vectors which are invariant densities under $\tau$.

First, we consider the class $P_n$ of all the permutation invariant matrices $A$, $A \in A_n$ (section 4.3) and we associate to each $A \in P_n$ a digraph $G$ (Lemma 4.3.4). Then, following a special labelling procedure, we construct (Theorem 4.3.1) the $A$-invariant vector.

In section 4.4, we extend the $P_n$-invariant vectors in a number of directions.

In section 4.5, starting with some vectors $\pi$, we construct matrices $A$, which are in $A_n$ but not in $P_n$, such that $\pi P = \pi$.

In section 4.6, we construct new $A_n$-invariant vectors from those known to be $A_n$-invariant.

Finally, in section 4.7, we investigate some properties of $A_n$-invariant vectors when $A \not\in P_n$.

In Chapter V, we use Markov maps to check the irreducibility or primitivity for constructible matrices. More explicitly, we give a new matrix representation for a Markov map, and we show that when a Markov map $\tau$ satisfies a simple condition, then a given large, non-negative, square matrix, induced by $\tau$, can be reduced to a smaller one preserving irreducibility or primitivity.
CHAPTER II
FROBENIUS-PERRON OPERATOR AND
NON-NEGATIVE MATRICES

2.1. REPRESENTATION OF THE FROBENIUS-PERRON OPERATOR FOR MARKOV MAPS

Let \( \tau: [0,1] \to [0,1] \) be a measurable, non-singular transformation.

We define the Frobenius-Perron operator \( P_\tau: L_1 \to L_1 \) by the formula:

\[
P_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}([0,x])} f(s)ds,
\]

where \( L_1 \) is the space of integrable functions on \([0,1]\).

It is well-known [14] that the operator \( P_\tau \) is linear and continuous. Moreover, it satisfies the following conditions:

a) \( P_\tau \) is positive: \( f \geq 0 \Rightarrow P_\tau f \geq 0 \).

b) \( P_\tau \) preserves integrals:

\[
\int_0^1 P_\tau f dm = \int_0^1 f dm, \quad f \in L_1.
\]

c) \( P_\tau^n = P_\tau^n \).

d) \( P_\tau f = f \) if and only if the measure \( dm = f dm \) is invariant under \( \tau \).

In this section, working with piecewise constant functions \( f \), we shall show that the Frobenius-Perron operator can be represented as a product of a vector \( f \) and of a \( n \times n \) matrix \( M_\tau \), depending on \( \tau \). To show that we need the following result.
Theorem 2.1.1: Let $0 = b_0 < b_1 < \ldots < b_n = 1$ be a partition of $[0,1]$. Let $\psi_i \in C^1([b_{i-1},b_i])$ and monotone. Assume also that each $\phi_i$ can be extended as a monotone $C^1$ function on $[0,1]$. Then for

$$\phi = \sum_{i=1}^{n} \phi_i \chi_{B_i},$$

where $B_i = (b_{i-1}, b_i)$,

$$p_{\phi} f(x) = \sum_{i=1}^{n} f(\psi_i(x)) \sigma_i(x) \chi_{J_i}(x),$$

where $\psi_i = \psi_i^{-1}$, $\sigma_i = |\psi'_i|$ and $J_i = \psi_i(B_i)$ and $\chi_A$ is the characteristic function of the set $A$.

Proof: Let $A_i(x) = \psi_i^{-1}([0,x]) \cap B_i = \psi_i([0,x]) \cap B_i$. Then,

since $\phi = \sum \phi_i \chi_{B_i}$, we get

$$\phi^{-1}([0,x]) = \bigcup_{i=1}^{n} A_i(x),$$

where $A_i(x)$'s are disjoint, since $B_i$'s are disjoint. Thus

$$p_{\phi} f(x) = \frac{d}{dx} \int_{\phi^{-1}([0,x])} f(s) ds$$

$$\bigg/ \phi^{-1}([0,x])$$

$$= \sum_{i=1}^{n} \frac{d}{dx} \int_{A_i(x)} f(s) ds. \quad (2.1)$$

Considering now the integral $\int_{A_i(x)} f(s) ds$, we have:

$$\int_{A_i(x)} f(s) ds = \int_{\psi_i(x)}^{\psi_i(0)} f(s) \chi_{B_i}(s) ds.$$

By condition (a) above, we want $\int_{A_i(x)} f \geq 0$ when $f \geq 0$. Now, since $\psi_i$ is monotone, $\psi_i$ is monotone and $\phi_i$ and $\psi_i$ are either both increasing or both decreasing. Therefore,
\[
\frac{\psi_i'(x)}{|\psi_i'(x)|} = \frac{\psi_i'(y)}{|\psi_i'(y)|} \quad \text{for each } x, y \in [0, 1].
\]

We use this to establish the sign. Thus,
\[
\int_{A_i(x)} f(s) ds = \frac{\psi_i'(x)}{|\psi_i'(x)|} \int_{\psi_i(x)} f(s) x_{B_i}(s) ds.
\]

and
\[
\frac{d}{dx} \int_{A_i(x)} f(s) ds = \frac{\psi_i'(x)}{|\psi_i'(x)|} \frac{d}{dx} \int_{\psi_i(x)} f(s) x_{B_i}(s) ds
\]
\[
= \frac{\psi_i'(x)}{|\psi_i'(x)|} f(\psi_i(x)) x_{B_i}(\psi_i(x)) \psi_i(x)
\]
\[
= \frac{|\psi_i'(x)|^2}{|\psi_i'(x)|} f(\psi_i(x)) x_{B_i}(\psi_i(x))
\]
\[
= f(\psi_i(x)) \sigma_i(x) x_{B_i}(\psi_i(x)).
\]

We note that
\[
x_{B_i}(\psi_i(x)) = 1 \Rightarrow \psi_i(x) \in B_i
\]
\[
\Rightarrow x \in \varphi_i(B_i) = J_i
\]
\[
\Rightarrow x_{J_i}(x) = 1
\]

Therefore, \( x_{J_i}(x) = x_{B_i}(\psi_i(x)) \) and we obtain
\[
\frac{d}{dx} \int_{A_i(x)} f(s) ds = f(\psi_i(x)) \sigma_i(x) x_{J_i}(x). \tag{2.2}
\]

Now, by substituting (2.2) into (2.1) we get
\[
P_\phi f(x) = \sum_{i=1}^{n} f(\psi_i(x)) \sigma_i(x) x_{J_i}(x).
\]

Q.E.D.
Lemma 2.1.1: If \( f \in C^1[a,b] \) with \(|f'| > 0\) then \( f \) is monotone on \([a,b]\).

Proof: \( f \in C^1[a,b] \Rightarrow f' \in C[a,b] \)
and
\[ |f'| > 0 \text{ implies } -\infty < f' < 0 \text{ or } \ 0 < f' < \infty, \]
since \( f' \) is continuous, it is only possible to have \( f' < 0 \) for each \( x \in [a,b] \) or \( f' > 0 \) for each \( x \in [a,b] \). In either case, \( f \) is monotone on \([a,b]\).

Q.E.D.

Let now \( J \) be the interval \([0,1]\) with a fixed partition \( J \) given by:

\[ J = \{ 0 = a_0 < a_1 < \ldots < a_n = 1 \}. \]

For each \( i, 1 \leq i \leq n \), let \( I_i = (a_{i-1}, a_i) \) and let \( \tau: I \to I \) be a piecewise linear \( C^1 \) transformation with respect to \( J \), with \( \inf_{x} |\tau'(x)| > 1 \). Let \( F \) be the class of all functions which are piecewise constant on the above partition, that is,

\[ f \in F \iff f(x) = \sum_{i=1}^{n} f_i \chi_{I_i}(x), \]

with the requirement that \( \sum_{i=1}^{n} f_i = 1 \). Such an \( f \) will be also represented by the row vector \( f = (f_1, f_2, \ldots, f_n) \). The following theorem is one of the most important of this thesis.

Theorem 2.1.2: If \( \tau \) satisfies the above assumptions, then there exists an \( n \times n \) matrix \( M_\tau \) such that \( P \tau f = f M_\tau \) for every \( f \in F \).
Proof: Since $P_\tau$ is a linear operator, we have,

$$P_\tau f(x) = P_\tau \sum_{i=1}^{n} f_i \chi_{I_i}(x)$$

$$= \sum_{i=1}^{n} f_i P_\tau \chi_{I_i}(x).$$

By Lemma 2.1.1 and Theorem 2.1.1 we obtain:

$$P_\tau f(x) = \sum_{i=1}^{n} f_i \sum_{j=1}^{n} \chi_{I_i} (\tau_j^{-1}(x)) \left| \frac{d\tau_j^{-1}(x)}{dx} \right| \chi_{\tau_j(I_j)}(x)$$

$$= \sum_{i=1}^{n} f_i \sum_{j=1}^{n} \chi_{I_i} (\tau_j^{-1}(x)) |\tau_j'|^{-1} \chi_{\tau_j(I_j)}(x).$$

Since $\tau_j^{-1}$ has range $I_j$, $\chi_{I_i} (\tau_j^{-1}(x))$ will be zero for all $i \neq j$.

Thus

$$P_\tau f(x) = \sum_{i=1}^{n} f_i |\tau_i'|^{-1} \chi_{\tau_i(I_i)}(x). \quad (2.3)$$

This proves that $P_\tau f$ is a piecewise constant function, i.e., $P_\tau f \in F$.

Therefore, there exist constants $d_1, d_2, \ldots, d_n$ such that, for $x \in I_j$

$$P_\tau f(x) = d_j, \quad 1 \leq j \leq n. \quad (2.4)$$

Now we shall compute the constants $d_j$, for $j = 1, 2, \ldots, n$. Let $x \in I_j$

then the right hand side of (2.3) is $f_i |\tau_i'|^{-1}$ if $I_j \subset \tau_i(I_i)$ and zero otherwise. If we define

$$m_{ij} = \begin{cases} |\tau_i'|^{-1}, & \text{if } I_j \subset \tau_i(I_i) \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

then for $x \in I_j$ (2.3) becomes
\[ d_j = \sum_{i=1}^{n} f_i m_{ij}, \quad 1 \leq j \leq n. \quad (2.6) \]

Using matrix representation, we can write (2.6) in the following form

\[ (d_1, d_2, \ldots, d_n) = f M_{\tau} \quad (2.7) \]

where \( f \) is a row vector with coefficients \( f_i, \ 1 \leq i \leq n \), and \( M_{\tau} = (m_{ij}) \) is a \( n \times n \) matrix defined by (2.5).

Therefore using (2.4) and (2.7) we get

\[ p_{\tau} f = f M_{\tau} \quad Q.E.D. \]

As we mentioned in the introduction there exists a class of transformations such that each member of that class admits a unique absolutely continuous invariant measure. The existence of such a class is shown in [3] and is defined as follows:

**Definition 2.1.1:** We say that a point transformation \( \tau: I \to I, \ I = [a,b], \) is in class \( \mathcal{T} \) if it satisfies the following conditions for a fixed partition \( J \):

1. \( \tau \) is piecewise \( C^2 \) with respect to \( J \).
2. \( \inf_{x} |\tau'(x)| > 1. \)
3. \( \tau \) takes partition points into partition points.
4. The partition \( J \) has the communication property under \( \tau \).

In [14] it is proved that the invariant density \( f \) under \( \tau \) is a fixed point of the Frobenius-Perron operator \( P_{\tau} \), i.e., is an invariant function under \( P_{\tau} \), \( P_{\tau} f = f \). Moreover, in [10], it is proved that the invariant density \( f \) under \( P_{\tau} \) is a piecewise constant function on the Markov partition \( J = \{I_i\}_{i=1}^{n} \) of \( I = [0,1] \).
This permits us to consider the function \( f \) as an \( n \)-vector 
\[ f = (f_1, f_2, \ldots, f_n) \]
defined by \( f_i = f(x) \) for each \( x \in I_i, 1 \leq i \leq n \).

Now, using Theorem 2.1.2, the piecewise constant function \( f \), which is
invariant under \( P_\tau \), becomes an invariant vector under \( M_\tau \). Hence
we can replace the problem of finding the invariant density \( f \) under \( \tau \)
using the operator \( P_\tau \) by the problem of finding the invariant
density \( f \) under \( \tau \), using the matrix \( M_\tau \), which is easier because
\( f \) is given as the solution of a system of linear equations. For that
reason Theorem 2.1.2 is essential to the rest of this thesis.

We shall continue with some remarks to Theorem 2.1.2.

**Remark 1:** The matrix \( M_\tau \) defined in the Theorem 2.1.2 is a
non-negative \( n \times n \) square matrix and, for \( 1 \leq j \leq n \) the non-zero
entries in the \( j \)th row are contiguous and equal to \( |\tau_j \cdot |^{-1} \). We
shall refer to \( M_\tau \) as the matrix induced by \( \tau \).

**Remark 2:** Since the transformation \( \tau \) is piecewise linear the
derivative \( \tau_i \), for \( 1 \leq i \leq n \), is given by the following:
\[
\frac{d \tau_i(x)}{dx} \bigg|_{x \in I_i} = \pm \frac{m(\tau_i(I_i))}{m(I_i)},
\]
where \( m \) is the Lebesgue measure on \( I \). Therefore we have
\[ |\tau_i^{-1}| = |\pm m(I_i)/m(\tau_i(I_i))|. \]
Thus we have the following form for the entries of the matrix \( M_\tau \):
\[
m_{ij} = \begin{cases} 
\frac{m(I_i)}{m(\tau_i(I_i))}, & \text{if } I_j \subset \tau_i(I_i) \\
0, & \text{otherwise} \end{cases}
\]
Remark 3: Although $\tau$ determines $M_{\tau}$ uniquely, by Theorem 2.1.2, the converse is not true. Specifically, an $M$ with equal, contiguous, non-negative entries in each row determines an equivalence class of $2^n$ transformations, since $\tau_j$ can be positive or negative without altering $M_{\tau}$. We illustrate that with the following example.

Example 2.1: Let $M$ be a matrix as described above given by

$$M = \begin{bmatrix}
0 & 0 & 0.5 & 0.5 \\
0.25 & 0.25 & 0.25 & 0.25 \\
0 & 0.5 & 0.5 & 0 \\
0.25 & 0.25 & 0.25 & 0.25
\end{bmatrix}$$

the transformations $\tau_1$ and $\tau_2$ show in Figure (2.1)

\[ a_4 = 1 \]
\[ a_3 \]
\[ a_2 \]
\[ a_1 \]
\[ a_0 = 0 \]

\[ I_1 \]
\[ I_2 \]
\[ I_3 \]
\[ I_4 \]

Assume $I_1 = I_2 = I_3 = I_4 = \frac{1}{4}$

\[ \tau_2 \]

Fig. 2.1

are 2 of the 16 transformations in the equivalence class. Obviously, $\tau_i, i = 1, 2$, belong to $\mathcal{T}$, since $\inf_x |\tau_i(x)| > 1$. 

2.2. PROPERTIES OF $M_r$

In this section we shall study some properties of $M_r$. First we shall give another definition for irreducibility of a matrix, and we shall prove that the new definition and the definition 1.1.4 are equivalent. Using Markov chain terminology [20], we give the following definitions.

**Definition 2.2.1:** A sequence $(i_0, i_1, i_2, \ldots, i_{m-1}, j)$, for $m \geq 1$, $i_0 = i$, from the index set $\{1, 2, \ldots, n\}$ of a non-negative matrix $A$, is said to form a chain of length $m$ between the ordered pair $(i, j)$ if

$$a_{i_0 i_1} a_{i_1 i_2} \cdots a_{i_{m-2} i_{m-1}} a_{i_{m-1} j} > 0.$$

A chain for which $i = j$ is called a cycle of length $m$ between $i$ and itself.

We say that $i$ leads to $j$, and write $i \rightarrow j$, if there exists an integer $m \geq 1$ such that $a_{ij}^{(m)} > 0$ or, equivalently, if there is a chain between $i$ and $j$.

We say that $i$ and $j$ communicate if $i \rightarrow j$ and $j \rightarrow i$ and write $i \leftrightarrow j$.

**Definition 2.2.2:** If $i \rightarrow i$, $d(i)$ is the period of the index $i$, if it is the greatest common divisor of those $k$ for which $a_{ii}^{(k)} > 0$. Note that if $a_{ii} > 0$ then $d(i) = 1$.

**Definition 2.2.3:** An $n \times n$ non-negative square matrix $A$ is said to be irreducible (in Markov chain terminology, all states (indices) communicate), if for every pair $i, j$ of its index set, there exists a positive integer $q_0$ (which may depend on $i$ and $j$).
i.e., \( q_0 = q_0^{(i,j)} \) such that \( a_{ij}^{(q_0)} > 0 \). Note that an irreducible matrix \( A \) cannot have a zero row or column.

An irreducible matrix is said to be periodic with period \( d \), if the period of any one of its indices satisfies \( d > 1 \), and is aperiodic if \( d = 1 \).

In [12], the following result was proved.

**Lemma 2.2.1:** If \( A \) is a non-negative and irreducible \( n \times n \) matrix, then:

\[
(I + A)^{n-1} > 0,
\]

where \( I \) is the unit \( n \times n \) square matrix.

**Proof:** It is sufficient to show that for every column vector \( y \geq 0 \) (\( y \neq 0 \)) the inequality

\[
(I + A)^{n-1} y > 0,
\]

holds. (Setting \( y = e_j \), \( 1 \leq j \leq n \), where \( e_j \) is the \( j \)th coordinate vector, we have \((I + A)^{n-1} > 0\).)

This inequality will be established if we can show that the vector \( z = (I + A)y = y + Ay \) always has at least one more positive coordinate than \( y \) does. (For, by repeating this argument \( n-1 \) times for an arbitrary non-zero \( y \geq 0 \), we get \((I + A)^{n-1} y > 0\).) Let us assume the contrary. Then, since \( A \geq 0 \), \( Ay \geq 0 \), \( z \) has at least as many positive coordinates as \( y \). So we assume that \( y \) and \( z \) have the same positive coordinates. Without loss of generality, we may assume that the columns \( y \) and \( z \) have the form

\[
y = \begin{bmatrix} u \\ 0 \end{bmatrix}, \quad z = \begin{bmatrix} v \\ 0 \end{bmatrix}, \quad (u > 0, v > 0)
\]
where the columns \( u \) and \( v \) are of the same dimension. (The columns \( y \) and \( z \) can be brought into this form by suitably renumbering the coordinates.) Setting:

\[
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}
\]

we have

\[
\begin{bmatrix} u \\ 0 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix}
\]

Thus

\[
A_{21} u = 0.
\]

Since \( u > 0 \) and \( A \geq 0 \), it follows that \( A_{21} = 0 \), which contradicts the hypothesis that \( A \) is irreducible.

Q.E.D.

**Corollary:** If \( A \) is a non-negative and irreducible \( n \times n \) square matrix, then for every index pair \( (i,j), 1 \leq i, j \leq n \), there exists a positive integer \( q_0 \) such that,

\[
(a_{ij})^{(q_0)} > 0.
\]

Moreover, \( q_0 \) can always be chosen within the bounds:

\[
q_0 \leq t - 1, \quad \text{if } i \neq j \]  
\[
q_0 \leq t, \quad \text{if } i = j
\]

(2.8)

where \( t \) is the degree of the minimal polynomial \( \psi(\lambda) \) of \( A \).
Proof: Let \( \rho(\lambda) \) denote the remainder on dividing \( (\lambda+1)^{n-1} \) by \( \psi(\lambda) \); i.e.,
\[
(\lambda+1)^{n-1} = \varphi(\lambda) \psi(\lambda) + \rho(\lambda)
\]
Then,
\[
(A + I)^{n-1} = \varphi(A) \psi(A) + \rho(A)
\]
or,
\[
(A + I)^{n-1} = \rho(A) > 0
\]
Since \( \psi(A) = 0 \) and \( (A+I)^{n-1} > 0 \), by virtue of Lemma 2.2.1. Since the degree of \( \rho(\lambda) \) is less than \( t \), it follows from the inequality \( \rho(A) > 0 \) that for arbitrary \( i, j (1 \leq i, j \leq n) \), at least one of the non-negative numbers,
\[
\delta_{ij}, a_{ij}, a_{ij}^{(2)}, \ldots, a_{ij}^{(t-1)}
\]
is not zero. Since \( \delta_{ij} = 0 \) for \( i \neq j \), the first of the relations (2.8) follows. The other relation (for \( i = j \)) is obtained in a similar manner if the inequality \( \rho(A) \) is replaced by \( A \rho(A) > 0 \).
Q.E.D.

Note that the product of an irreducible non-negative matrix and a positive matrix is itself positive.

Therefore, Definition 1.1.4 implies Definition 2.2.3. To show that Definition 2.2.3 implies Definition 1.1.4, we first note that Definition 1.1.4 can be stated as follows.

A \( n \times n \) square matrix \( A = (a_{ij}) \) is called reducible if the index set \( \{1, 2, \ldots, n\} \) can be split into two complementary sets (without common indices) \( \{i_1, i_2, \ldots, i_\mu\}, \{j_1, j_2, \ldots, j_\nu\}, \mu + \nu = n \), such that:
\[ a_{ij} = 0 \quad (\alpha = 1, 2, \ldots, \mu; \beta = 1, 2, \ldots, \nu). \]

**Lemma 2.2.2:** Let \( A \) be a non-negative \( n \times n \) square matrix such that, for each pair \((i, j), 1 \leq i, j \leq n\), there exists a positive integer \( m \), such that \( a_{ij}^{(m)} > 0 \) i.e., that the state \( i \) leads to \( j \) (\( i \rightarrow j \)) or, in other words, there exists a chain of length \( m \) between the ordered pair \((i, j)\):

\[ a_{i1}^{(1)} a_{11}^{(1)} \ldots a_{1m-1}^{(1)} a_{m-1}^{(1)} > 0. \]

Then the matrix \( A \) is irreducible.

**Proof:** Assume that \( A \) is reducible, then we can split the index set \( \{1, 2, \ldots, n\} \) into two complementary sets, without common indices, \( I = \{i_1, i_2, \ldots, i_{\mu}\}, J = \{j_1, j_2, \ldots, j_{\nu}\}, \mu + \nu = n \), such that:

\[ a_{i_j} = 0 \quad (\alpha = 1, 2, \ldots, \mu; \beta = 1, 2, \ldots, \nu). \]

Let us consider the ordered pair \((i, j)\) where \( i \in I \) and \( j \in J \). For that pair, there exist a chain of length \( q \) such that:

\[ a_{i1}^{(1)} a_{11}^{(1)} \ldots a_{q-1}^{(1)} > 0. \]

This is true only if \( a_{i_k k+1}^{(1)} > 0 \), \( k = 0, 1, \ldots, q \), \( i_0 = i, i_{k+1} = j \) since \( i \in I \) and \( a_{i_1} > 0, j_1 \notin J \), i.e., \( i_1 \in I \). Similarly, since \( a_{i_2} > 0, j_2 \notin J \), i.e., \( j_2 \in I \). Continuing in this way, we finally obtain, \( i_{q-1} \in I \) and \( a_{iq-1} > 0 \), which implies that \( j \notin J \). But this contradicts the hypothesis that \( j \in J \). Therefore, we cannot split the index set \( \{1, 2, \ldots, n\} \) into two complementary sets, i.e., \( A \) is irreducible.
Therefore, Definitions 1.1.4 and 2.2.3 are equivalent.

Let us now study some properties of the matrix $M_\tau$.

**Lemma 2.2.3:** The matrix $M_\tau$, for $\tau \in \mathcal{T}$, $\tau$ piecewise linear, is irreducible.

**Proof:** From the definition of $M_\tau$, we have:

$$m_{ij} = \begin{cases} |\tau_i^j|^{-1}, & \text{if } I_j \subset \tau_i(I_i) \\ 0, & \text{otherwise} \end{cases}$$

We claim that, if $I_j \subset \tau_i^k(I_i)$, then

$$m_{ij}^{(k)} = |(\tau_i^k)|^{-1}.$$

We shall prove this by induction. For $k=1$, the result is true. Assume that it is true for $k=r$ and consider the case $k=r+1$. Then

$$|\tau_i^{r+1}|^{-1} = |[\tau_i^1, \tau_i^r]|^{-1}$$

$$= \left[ \frac{d\tau_i^1(x)}{d\tau_i^r(x)} \frac{d\tau_i^r(x)}{dx} \right]^{-1}.$$

Consider now the term $\frac{d\tau_i^r(x)}{d\tau_i^r(x)}$. Putting $\tau_i^r(x) = y$, we get,

$$\frac{d\tau_i^r(x)}{d\tau_i^r(x)} = \frac{d\tau_i^r(y)}{dy}.$$

It is obvious that $\frac{d\tau_i(y)}{dy} \neq 0$ if and only if $y \in I_i$ and in that case,

$$\frac{d\tau_i^r(y)}{dy} = \frac{d\tau_i(y)}{dy}.$$

This result is due to the author.
since $\tau$ is piecewise linear, we have $\frac{d\tau_i(y)}{dy} = \tau'_i$. Therefore, we have
\[
|((\tau_i^{r+1})')^{-1}| = \left| \tau'_i \cdot \frac{d\tau_i^r(x)}{dx} \right|^{-1}
\]
\[
= |\tau'_i|^{-1} |(\tau_i^r)'|^{-1}
\]
\[
= \prod_j m_{ij}^{(r)} = m_{ij}^{(r+1)}.
\]

Now, since $\tau \in \mathcal{T}$, $\tau$ has the communication property, (irreducibility in Markov chain terminology) i.e., for any $I_i, I_j \in J$, there exist integers $p$ and $q$ such that:
\[
I_i = \tau^p(I_j) \text{ and } I_j = \tau^q(I_i).
\]
Thus for any pair $(i,j)$ of the index set there exists an integer $q$
such that $m_{ij}^{(q)} = |(\tau_i^q)'|^{-1} > 0$, since $I_j = \tau^q(I_i)$. Hence $m_{ij}^{(q)} > 0$
and so, $\tau$ is irreducible.

Q.E.D.

Lemma 2.2.4: Let $\tau \in \mathcal{T}$, $\tau$ piecewise linear. Assume that
the subintervals of the partition $J = \{I_i\}_{i=1}^n$ are all of equal
length. Then the matrix $M_\tau$ is row stochastic.

Proof: Let the length of each subinterval be denoted by $\Delta$, i.e.,
\[
\ell(I_i) = m(I_i) = a_i - a_{i-1} = \Delta, \quad i = 1, \ldots, n, \quad a_0 = 0, \quad a_n = 1.
\]
$\ell(I_i) = m(I_i)$, since $m$ is the Lebesgue measure. Suppose the Markov map $\tau$ maps $I_i$ onto $I_{t} \cup I_{t+1} \cup \ldots \cup I_{t+k}$. Then,
\[
\tau^i(I_i) = I_{t} \cup I_{t+1} \cup \ldots \cup I_{t+k}, \quad \text{and} \quad m(\tau^i(I_i)) = a_{t+k} - a_{t-1} = (k+1)\Delta.
\]
Now since
\[ m_{ij} = \begin{cases} \frac{m(I_i)}{m(\tau_j(I_i))} & \text{if } I_j \subseteq \tau_j(I_i) \\ 0 & \text{otherwise} \end{cases} \]
it follows that the \( i \)th row of \( M_\tau \) has entries
\[ m_{ij} = \frac{m(I_i)}{m(\tau_j(I_i))} = \frac{\Delta}{(k+1)\Delta} = \frac{1}{k+1} , \quad j = t, t+1, \ldots, t+k \]
in columns \( t, t+1, \ldots, t+k \) and zero in all \( t \) remaining columns.
Therefore,
\[ \sum_{j=t}^{t+k} m_{ij} = \sum_{j=t}^{t+k} \frac{1}{k+1} = \sum_{i=1}^{k+1} \frac{1}{k+1} = 1 \]
Since \( i \) was arbitrary, the matrix \( M_\tau \) is row stochastic.

Q.E.D.

We shall close this chapter with a Lemma which says under what conditions \( M_\tau \) is primitive. Before that we recall from matrix theory [20] the following.

Theorem 2.2.1: A non-negative square matrix \( A \) is primitive if and only if it is irreducible and has at least one aperiodic index.

Proof: Let \( A \) be primitive. Then is obviously irreducible and has \( d(1) = 1 \) for every \( i \), since for any fixed \( i \), and \( n \) large enough, we have \( a_{ii}^{(n)} > 0, a_{ii}^{(n+1)} > 0 \) and the greatest common divisor of \( n \) and \( n+1 \) is 1.

Conversely, suppose that \( a_{ii} \) is aperiodic for some \( i \). Since we have:
\[ a_{ii}^{(n+m)} = \sum_{k} a_{kk}^{(m)} a_{ii}^{(n)} , \]
the set \( D = \{ n \geq 1 : a^{(n)}_{ij} > 0 \} \) is closed under addition. This, with the fact that its greatest common divisor is 1, implies that \( D \) contains all large positive integers say all \( n \geq M \).

Now, consider any other index \( j \) and choose \( k \) and \( \ell \) so that \( a_{ij}^{(k)} > 0 \) and \( a_{ji}^{(\ell)} > 0 \). Then for all \( n \geq M \) we have:

\[
a_{jj}^{(\ell + n + k)} \geq a_{ij}^{(\ell)} a_{ii}^{(n)} a_{ij}^{(k)} > 0,
\]

where the first inequality follows from the rule of matrix multiplication and the non-negativity of the elements of \( A \). This means that \( a_{jj}^{(m)} > 0 \) for all large \( m \), implying \( a_{jj} \) is also aperiodic. Thus any element on the diagonal is aperiodic.

Similarly, we have:

\[
a_{ji}^{(n + \ell)} \geq a_{ji}^{(\ell)} a_{ii}^{(n)} > 0 \quad \text{for} \quad n \geq M.
\]

Thus, the off-diagonal entries of \( A^n \) are likewise positive when \( n \) is large enough. Since \( A \) is a finite matrix, it is clear that for large values of \( n \), \( A^n \) is strictly positive, and so, primitive.

Q.E.D.

**Lemma 2.2.5**: Let \( \tau \in \mathcal{T} \) be piecewise linear and let it have a fixed point. Then the matrix \( M_\tau \) is primitive.

**Proof**: By Lemma 2.2.3, \( M_\tau \) is irreducible, and if we show that \( M_\tau \) has an aperiodic index then Theorem 2.2.1 is applicable and so \( M_\tau \) is primitive. Since \( \tau \) has a fixed point, there exists \( x_0 \in I_i \) such that \( \tau_i(x_0) = x_0 \) for some \( i \). This implies that \( I_i \subset \tau_i(I_i) \),
since \( \tau \) is Markov and linear on \( I \). It follows that

\[
m_{11} = |\tau_i|^{-1} > 0,
\]

i.e., the \( i \)th index is aperiodic.

Q.E.D.
CHAPTER III

IN Variant Densities For A Class Of
Pieewise Linear Transformations

3.1. The Perron-Frobenius Theorem For Irreducible Matrices

In this section we shall prove a theorem known as Perron-Frobenius dealing with non-negative irreducible matrices. Since, as we know from Lemma 2.2.3, $M_T$ is a non-negative irreducible matrix, the Perron-Frobenius Theorem is essential to the rest of this thesis, as we shall see in section 3.2. We shall start with some preliminaries.

Let $\mathbb{R}$ be the field of real or complex numbers. Let $A \in \mathbb{R}_{n \times n}$, the space of all $n \times n$ matrices, and $x \in \mathbb{R}_n$, the space of all $n$ column vectors, then the vector $Ax$ is in $\mathbb{R}_n$ and is a member of the range of $A$. Let us consider the vectors $x$, which on multiplication by $A$ are transformed into multiples of themselves: that is, consider those vectors $x \neq 0$ for which there exists a member $\mu_j$ of $\mathbb{R}$, the scalar field, such that:

$$Ax = \mu_j x.$$

Such a non-zero vector $x$ is called a (right) eigenvector of $A$ and $\mu_j$ is the corresponding eigenvalue. The left eigenvector $y$ of $A$ corresponding to the eigenvalue $\mu_j$ is defined by the relation:

$$y^T(\mu_j I - A) = 0^T.$$

It is well known [12 page 54] that if $A \in \mathbb{R}_{n \times n}$, the $\mu_j \in \mathbb{R}$ is an eigenvalue of $A$ if and only if $\mu_j$ is a zero of the characteristic polynomial of $A$:

$$\xi(\mu) = \det(\mu I - A).$$
Let us now state and prove the Perron-Frobenius Theorem [5, 9, 12, 20].

**Theorem 3.1.1.** (Perron-Frobenius) A non-negative irreducible matrix $A = (a_{ij}), 1 \leq i, j \leq n$, always has a positive eigenvalue $r$, which is a simple root of the characteristic equation. The absolute value of all the other eigenvalues do not exceed $r$. To the "maximal" eigenvalue $r$, there corresponds an eigenvector with positive coordinates. Moreover, if $0 \leq B \leq A$ and $\beta$ is an eigenvalue of $B$, then $|\beta| \leq r$. If $|\beta| = r$, then $B = A$.

**Proof:** Consider the real-valued function $r$ defined on the non-zero vectors $x \geq 0$ by

$$r(x) = \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i},$$

where $(Ax)_i$ denotes the $i$th component of the vector $Ax$, i.e.,

$$(Ax)_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \ldots, n.$$ We interpret as $\infty$ the ratio if $x_i = 0$ for some $i$. Clearly,

$$0 \leq r(x) < \infty.$$

Now, since:

$$r(x)x_i \leq (Ax)_i = \sum_{j=1}^{n} a_{ij}x_j \quad \text{for each } i,$$

$$r(x)x \leq Ax,$$

and so,

$$r(x)x \leq IAx.$$
Since $IA \leq IK$, where $K = \max \sum a_{ij}$, it follows that,

$$r(x) \leq \frac{IKx}{lx} = K = \max \sum a_{ij},$$

in other words, $r(x)$ is uniformly bounded above for all such $x$. Thus $r(x)$ is the largest real number $q$ for which,

$$qx \leq Ax.$$

We shall show that the function $r(x)$ assumes a maximum real value $r$ for some real vector $z \geq 0$.

Let $L$ denote the domain of the function $r$, that is the set of all non-zero, non-negative vectors of order $n$, and define

$$r = \sup_{x \in L} r(x). \tag{3.1}$$

From the definition of $r(x)$ we observe that the value of $r(x)$ does not change when we multiply the vector $x \geq 0$ ($x \neq 0$) by a number $\lambda > 0$. Therefore, in the computation of this supremum, we can restrict ourselves to the closed and bounded set $M$ of vectors $x$ for which $x \geq 0$ and $\|x\| = \sum_{i=1}^{n} x_i^2 = 1$. Thus $M \subseteq L$ and,

$$r = \sup_{x \in M} r(x).$$

It is known [19] that if the function $r(x)$ was continuous on $M$ we could equate the sup $r(x)$ to $\max_{x \in M} r(x)$, and since a continuous function takes on its maximum value on its domain, then the existence of a maximum (and so a supremum) would be guaranteed. However, $r(x)$ may have discontinuities at points where elements of $x$ vanish. Therefore, we replace $M$ by $N$ where $N$ consists of all vectors $y$ of the form,

$$y = (I + A)^{n-1} x, \quad x \in M.$$
By Lemma 2.2.1 \( N \) consists of positive vectors only and so \( N \subseteq L \).
Moreover, the set \( N \), like \( M \), is bounded and closed\(^1\), and \( r(y) \) is continuous on \( N \).

If we multiply both sides of the inequality,

\[
r(x)x \leq Ax, \quad x \in M
\]

by \((I+A)^{n-1} > 0\), we obtain:

\[
r(x)y \leq Ay, \quad y = (I+A)^{n-1}x \in N.
\]

Now \( r(y) \) is the greatest number \( q \) such that \( qy \leq Ay \) and hence \( r(x) \leq r(y) \). Thus,

\[
r = \sup_{x \in M} r(x) \leq \max_{y \in N} r(y).
\]

But since \( N \subseteq L \),

\[
\max_{y \in N} r(y) \leq \sup_{x \in L} r(x) = \sup_{x \in M} r(x).
\]

Hence

\[
r = \max_{y \in N} r(y)
\]

and there is a \( z > 0 \) such that \( r = r(z) \).

It is possible to have other vectors in \( L \) for which \( r(x) \) attains the value \( r \). Any such vector is called an extremal vector of \( A \).

It remains to be shown that:

(i) The number \( r \) defined by (3.1) is positive and is an eigenvalue of \( \hat{A} \).

(ii) Every extremal vector \( z \) is positive and is an eigenvector of \( A \) for the eigenvalue \( r \), i.e., \( Az = rz \).

\(^1\)The set \( N \) is bounded and closed because it is a continuous map, \( y = (I+A)^{n-1}x \), of the bounded, closed set of vectors \( x \) for which \( x \geq 0 \) and \( \sum x_i^2 = 1 \).
Consider the n-vector \( u = (1,1,\ldots,1)^T \), then

\[
 r(u) = \min_{k} \sum_{i} a_{ik} > 0. 
\]

Since \( r \geq r(u) \), we deduce that \( r > 0 \).

Let \( z \) be an extremal vector and let \( x = (I + A)^{n-1}z \). Without loss of generality, we may suppose that \( z \in M \). By Lemma 2.2.1, \( x > 0 \), and clearly \( x \in N \). We also assume that \( Az - rz \neq 0 \), then,

\[
(I + A)^{n-1}(Az - rz) > 0.
\]

Hence \( Ax - rx > 0, \quad rx < Ax \), which implies that \( r < r(x) \). But this contradicts the definition (3.2) of \( r \). Therefore we must have \( Az = rz \). Thus any extremal vector \( z \) is a right eigenvector of \( A \) for the eigenvalue \( r \). Now, since \( Az = rz \) we have

\[
0 < x = (I + A)^{n-1}z = (I + r)^{n-1}z,
\]

and since \( r > 0 \) we have \( z > 0 \).

We shall show now that the absolute value of all the other real or complex eigenvalues do not exceed \( r \). Let

\[
Ay = \alpha y, \quad y > 0
\]

since \( A \geq 0 \), we have:

\[
|\alpha y| = |Ay| \leq Ay, \quad (|y| = (|y_1|, |y_2|, \ldots, |y_n|))
\]

Hence \( |\alpha| \leq r(|y|) \leq r \), which we wanted to prove.

Now, let \( y \) be some eigenvector corresponding to \( r \), \( Ay = \alpha y \) \( (y > 0) \). Then, setting \( \alpha = r \) in (3.3) and (3.4), we conclude that

Since \( A \) is irreducible, it can have no columns consisting entirely of zeros.
\[ |y_i| \text{ is an external vector, so that } |y_i| > 0, y_i \neq 0, 1 \leq i \leq n. \text{ This implies that the dimension of the right eigenspace of } \lambda \text{ is } 1, \text{ since, otherwise, we could find two linearly independent right eigenvectors } z_1, z_2 \text{ and then determine numbers } \alpha, \beta \text{ such that } \alpha z_1 + \beta z_2 \text{ has a zero coordinate. From what we have shown, this is impossible. This proves that the geometric multiplicity of } \lambda \text{ is } 1. \]

Let us consider now the adjoint matrix of the characteristic matrix \( \lambda I - A \):

\[ B(\lambda) = [B_{ik}(\lambda)]_n = \Delta(\lambda)(\lambda I - A)^{-1}, \]

where \( \Delta(\lambda) \) is the characteristic polynomial of \( A \) and \( B_{ik}(\lambda) \), the algebraic complement of the element \( \lambda \delta_{ki} - a_{ki} \) in the determinant \( \Delta(\lambda) \). Now, since only one eigenvector \( z = (z_1, z_2, \ldots, z_n)^T, z_i > 0 \) for \( 1 \leq i \leq n \), corresponds to the eigenvalue \( \lambda \), it follows that \( B(\lambda) \neq 0 \), and that in every non-zero column of \( B(\lambda) \) all the elements are different from zero and are of the same sign. The same is true for the rows of \( B(\lambda) \), since \( A \) can be replaced by the transposed matrix \( A^T \).

From these properties of the rows and columns of \( A \), it follows that all the \( B_{ik}(\lambda), 1 \leq i, k \leq n \), are different from zero and of the same sign \( \sigma \). Therefore:

\[ c\Delta'(r) = \sigma \sum_{i=1}^{n} B_{ii}(r) > 0, \]

i.e., \( \Delta'(r) \neq 0 \) (\( \Delta'(r) \) means the derivative of \( \Delta(\lambda) \) at \( r \)) and \( r \) is a simple root of the characteristic equation \( \Delta(\lambda) = 0 \). In other words, the algebraic multiplicity of the eigenvalue \( r \) is 1.

---

\(^1\)By geometric multiplicity of an eigenvalue \( r \), we mean the maximal number of linearly independent right eigenvectors associated with \( r \).

\(^2\)Up to constant multiples.
Finally, let \( y \neq 0 \) be a right eigenvector of \( B \) corresponding to \( \beta \). Then taking absolute value, we have:

\[
|\beta| |y| \leq |B| |y| \leq |A| |y|.
\]  

(3.5)

Now, taking the vector \( z \), which we used above we get:

\[
|\beta| |z| \leq |z| |y| = |z| |y|,
\]

and since \( |z| |y| > 0 \), we get:

\[
|\beta| \leq r.
\]

Suppose now \( |\beta| = r \). Then, from (3.5) we have, \( r |y| \leq |A| |y| \). It follows that \( |A| |y| = r |y| > 0 \), while from (3.5) we get \( r |y| = |B| |y| = |A| |y| \). So, it must follow from \( B \preceq A \) that \( B = A \).

Q.E.D.

Corollary 3.1.1:

\[
\min_{i,j=1}^{n} a_{ij} \leq r \leq \max_{i,j=1}^{n} a_{ij}
\]

(3.6)

with equality on either side implying equality throughout (i.e., \( r \) can only be equal to the maximal or minimal row sum, if all row sums are equal).

Proof: From (3.2) we have that:

\[
r = \max_{x \succ 0} \min_{i} \frac{(Ax)_{i}}{x_{i}}.
\]

(3.7)

Similarly we can show [9, vol. II, p. 64] that:

\[
r = \min_{x \succ 0} \max_{i} \frac{(Ax)_{i}}{x_{i}}.
\]

(3.8)

and a combination of (3.7) and (3.8), together with \( x = (1,1,\ldots,1)^{T} \) gives (3.6).
Now assume that one of the equalities in (3.6) holds, but not all row sums are equal. Then by increasing (or decreasing) the positive elements of $A$ (but keeping them positive), we produce a new irreducible matrix, with all row sums equal and the same $r$, in view of (3.6) which is impossible by hypothesis of the last part of Theorem 3.1.1.

**Remark:** We can state an analogous theorem with the corresponding corollary for primitive matrices. The proof remains the same.

**Lemma 3.1.1:** Let $M$ be a non-negative, row stochastic and irreducible $n \times n$ matrix. Then $M$ has 1 as the eigenvalue of maximum absolute value, and furthermore the algebraic and geometric multiplicities of this eigenvalue are also 1.

**Proof:** By Corollary 3.1.1 we have:

$$\min_{i,j} \sum_{j=1}^{n} m_{ij} \leq r \leq \max_{i,j} \sum_{i=1}^{n} m_{ij},$$

since $M$ is stochastic $\sum_{j=1}^{n} m_{ij} = 1$ for each $i$. Therefore,

$$\min_{i,j} \sum_{j=1}^{n} m_{ij} = 1 = \max_{i,j} \sum_{i=1}^{n} m_{ij}.$$

This implies that $r = 1$. Moreover, Theorem 3.1.1 shows that the geometric and algebraic multiplicities are also 1.

Q.E.D.

We shall close this section with one more result from matrix theory. Before that, we recall [12] that two $n \times n$ matrices $A$ and $B$ are similar if there exists a non-singular $n \times n$ matrix $T$ (non-singular means that $\det T \neq 0$), such that $A = T^{-1}BT$.

**Lemma 3.1.2:** Similar matrices have the same eigenvalues.
Proof: We shall prove that similar matrices have the same characteristic polynomial. Hence, it follows that they have the same eigenvalues.

Suppose that \( A = T^{-1} B T \), i.e., \( A \) and \( B \) are similar, and let \( C(\mu) = \det(\mu I - A) \) be the characteristic polynomial of \( A \). Then:

\[
C(\mu) = \det(\mu I - T^{-1} B T)
\]
\[
= \det(\mu I - T^{-1} T^{-1} B T)
\]
\[
= \det(\mu I - T^{-1} B T)
\]
\[
= \det T^{-1} (\mu I - B) T
\]
\[
= \det(\mu I - B)
\]

since \( \det(T^{-1}) \det(T) = 1 \). Thus, the characteristic polynomials of \( A \) and \( B \) are the same, i.e., the eigenvalues of a matrix are invariant under similar transformation.

Q.E.D.

3.2. THE MATRIX \( M_\tau \) AND ITS EIGENFUNCTIONS

As we show in Chapter II the piecewise linear Markov map \( \tau, \tau \in \cal T \), induces an \( n \times n \) matrix \( M_\tau \) defined as follows:

\[
m_{ij} = \begin{cases} 
|\tau_j|^{-1}, & \text{if } I_j \subset \tau_i(I_i) \\
0, & \text{otherwise}
\end{cases}
\]

Lemma 2.2.3 shows that \( M_\tau \) is irreducible and Lemma 2.2.4 shows that, when the subintervals of the partition \( J \) are all of equal length, the matrix \( M_\tau \) is stochastic. Therefore, by Lemma 3.1.1, the stochastic
matrix $M$ has 1 as the eigenvalue of maximum absolute value, with geometric and algebraic multiplicities also 1.

In [6] it is proved that the same is true for any matrix $M$, $\tau \in T$, $\tau$ piecewise linear, and not necessarily stochastic. This implies that there exists a large class of non-negative, non-stochastic, irreducible matrices whose eigenvalues of maximum absolute value is 1.

Before we present this fact, we remark that pre- or post-multiplication by a non-singular diagonal matrix cannot change any non-zero elements of a matrix $A$ to zero, and therefore such multiplication has no effect on irreducibility.

Theorem 3.2.1: Let $\tau \in T$, be piecewise linear. Then the matrix $M$ induced by $\tau$ has 1 as the eigenvalue of maximum absolute value, the algebraic and geometric multiplicities of this eigenvalue are also 1.

Proof: Let $0 = a_0 < a_1 < \ldots < a_n = 1$ be a partition of $[0,1] = I$, and let $I_j = (a_{j-1}, a_j)$, $1 \leq j \leq n$. Define

$$\delta = \prod_{j=1}^{n} (a_j - a_{j-1})^2,$$

and

$$\delta_i = \frac{\delta}{a_i - a_{i-1}} = \prod_{j=1 \atop j \neq i}^{n} (a_j - a_{j-1})^2, \quad 1 \leq i \leq n.$$ 

Now, let us define the diagonal matrix $D$ to have entries $d_{ii} = \delta_i$, $1 \leq i \leq n$. Then $E = D^{-1}$ is a diagonal matrix with entries $e_{ii} = \delta_i^{-1}$, $1 \leq i \leq n$. 
Suppose now that

$$\tau(I_j) = I_j \cup I_{j+1} \cup \ldots \cup I_{j+k}$$

Then, by definition, the $i$th row of $M_\tau$ has entries equal to

$$m(I_i) = \frac{a_i - a_{i-1}}{m(\tau(I_i))} \text{ in columns } j,j+1,\ldots,j+k, \text{ and zero in every remaining column. Now, let } B = DM_\tau D^{-1}. \text{ Then,}$$

$$b_{rs} = d_{rs} \frac{m_{rs}}{d_{ss}} \delta^{-1}$$

$$= \delta_{rs} m_{rs} \delta^{-1}$$

We claim that $B$ is row stochastic. To show that, let us consider the row sum of the $i$th row of $B$.

$$\sum_{s=1}^{n} b_{is} = \sum_{s=1}^{n} \delta_{is} m_{is} \delta^{-1}$$

$$= \sum_{s=j}^{j+k} \delta_{is} \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}} \delta^{-1}$$

$$= \delta_{ij} \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}} \left[ \delta^{-1} + \delta^{-1} + \ldots + \delta^{-1} \right]$$

$$= \delta_{ij} \left[ \frac{a_i - a_{i-1}}{a_{j+k} - a_{j-1}} \delta^{-1} + \frac{a_{j+1} - a_{j+1} - 1}{\delta} + \ldots + \frac{a_{j+k} - a_{j+k} - 1}{\delta} \right]$$

$$= 1$$

Therefore $B$ is row stochastic, and since $M_\tau$ is irreducible, then so is $B$. Thus the result follows from Lemma 3.1.1 and Lemma 3.1.2.

Q.E.D.

As we mentioned in the introduction, one of the main goals of this thesis is to find the unique invariant densities for Markov maps $\tau$, belonging in $\mathcal{T}$. In the following we shall find the unique invariant
density of a piecewise linear transformation \( \tau \in \mathcal{T} \) when it satisfies two simple conditions.

Let \( I = [0, 1] \), and \( 0 = a_0 < a_1 < \ldots < a_n = 1 \) be any partition of \( I \). Let \( I_i = (a_{i-1}, a_i), 1 \leq i \leq n \). Let us consider the piecewise linear transformation \( \tau : I \to I \), \( \tau \in \mathcal{T} \), defined by the conditions:

\[
(\varepsilon_1) \quad \tau(I_i) = I_{i+1}, \quad 1 \leq i < n
\]

and

\[
(\varepsilon_2) \quad \tau(I_n) = \bigcup_{i=1}^{n} I_i = I.
\]

The transformation \( \tau \) induces a matrix \( M_\tau \) which, according to Theorem 3.2.1, has 1 as the eigenvalue of maximum absolute value, with algebraic and geometric multiplicity both equal to 1. To that eigenvalue there corresponds a unique (left) eigenvector (since the geometric multiplicity is 1) \( \pi \), i.e.,

\[ \pi M_\tau = \pi. \]

In Chapter II, page 13, we show that the vector \( \pi \), when considered as a piecewise constant function is the unique density invariant under \( \tau \).

The next theorem shows how we can construct invariant vectors \( \pi \).

**Theorem 3.2.2:** Let \( \tau \in \mathcal{T} \), piecewise linear, be a transformation on \([0, 1]\), satisfying conditions \((\varepsilon_1)\) and \((\varepsilon_2)\). Then \( \tau \) admits a unique (up to constant multiples) invariant density (eigenfunction) \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \), where:

\[ \pi_i = \pi|_{I_i} = \frac{a_i - a_{i-1}}{a_i - a_0}, \quad a_0 = 0, a_n = 1. \]

**Proof:** By condition \((\varepsilon_1)\), the \( n \times n \) matrix \( M_\tau \) is given by:
\( m_{ij} = \begin{cases} \frac{a_i - a_{i-1}}{a_i + a_{i-2}} , & \text{if } j = i + 1 \quad (\text{or } i = j - 1) \\ 0 , & \text{otherwise} \end{cases} \)

Also,
\[
m_{nj} = \frac{a_n - a_{n-1}}{a_n - a_0} , \quad 1 \leq j \leq n ;
\]
i.e., \( M_\pi \) has nonzero entries on the superdiagonal and on the last row.

We know that \( M_\pi \) has \( 1 \) as the eigenvalue of maximum absolute value. It only remains to show that \( \pi M_\pi = \pi \). We have,
\[
\sum_{r=1}^{n} \pi_r m_{rs} = \begin{cases} \pi_{s-1} m_{s-1,s} + \pi_n m_{ns} , & s \neq 1 \\ \pi_n m_{ns} , & s = 1 \end{cases}
\]

For \( s = 1 \), we have,
\[
\pi_n m_{ns} = \frac{a_n - a_0}{a_n - a_0} = 1
\]
\[
= \frac{a_1 - a_0}{a_1 - a_1-1} = \pi_1
\]

For \( s \neq 1 \), we have,
\[
\pi_{s-1} m_{s-1,s} + \pi_n m_{ns} = \frac{a_{s-1} - a_0}{a_{s-1} - a_{s-2}} + \frac{a_{s-1} - a_s}{a_s - a_s-1} + 1
\]
\[
= \frac{a_{s-1} - a_0}{a_s} + \frac{a_{s-1} a_s}{a_s} = \pi_s
\]
Hence, $\pi M_\tau = \pi$. Since the geometric multiplicity of the eigenvalue 1 is 1, $\pi$ is unique up to constant multiples. 

Q.E.D.

Now we consider the inverse problem. Given any vector $g = (g_1, g_2, \ldots, g_n)$ of length $n$, is it possible to construct a Markov map $\tau$ whose induced matrix $M_\tau$ satisfies $gM_\tau = g$? The answer is affirmative if $g$ satisfies the following two conditions:

(i) $g_1 = c$, where $c > 0$, any constant,

and

(ii) $g_i > c$, $2 \leq i \leq n$.

If (i) and (ii) are satisfied, we can show that for every such $g$, there exists a Markov map $\tau$ satisfying conditions $(e_1)$ and $(e_2)$ above and such that $g$ is the unique eigenfunction associated with the eigenvalue 1 of the matrix $M_\tau$. Without loss of generality, we can assume $c = 1$.

**Theorem 3.2.3:** Let $g = (g_1, g_2, \ldots, g_n)$ be any vector of length $n$, such that $g_1 = 1$ and $g_i > 1$, $2 \leq i \leq n$. Then there exists a Markov map whose induced matrix $M_\tau$ satisfies $gM_\tau = g$.

**Proof:** Let $a_0 = 0$. We choose $a_1, 0 < a_1 < 1$ such that:

$$a_i = a_{i-1} + \frac{a_{i-1} - a_0}{g_{i-1}} \leq 1,$$

for all $2 \leq i \leq n$. It follows that:

$$a_1 - a_0 = a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_{i-1}},$$

and
\[ a_i - a_{i-1} = \frac{a_{i-1} - a_0}{g_i - 1} \]

Consider:

\[ \pi_i = \frac{a_i - a_0}{a_i - a_{i-1}} \]

\[ = \left[ a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_i - 1} \right] \cdot \frac{a_{i-1} - a_0}{g_i - 1} \]

\[ = \left[ a_{i-1} - a_0 + \frac{a_{i-1} - a_0}{g_i - 1} \right] \cdot \frac{g_i - 1}{a_{i-1} - a_0} \]

\[ = g_i - 1 + 1 \]

\[ = g_i \]

Clearly, \( \pi_1 = 1 = g_1 \). Therefore, from Theorem 3.2.2, we have:

\[ gM_T = g. \]

\[ \text{Q.E.D.} \]

We shall now show by means of an example how to construct the unique invariant density for a given Markov map.
Example 3.1: Let \( \tau : [0,1] \to [0,1] \) be the piecewise linear transformation in \( \mathcal{T} \), defined by \( \tau(0) = 0, \tau(\frac{1}{2^n}) = 1, \tau(1) = \frac{1}{2^n}, n \geq 2 \), (Fig. 3.1).

![Diagram](image)

**Fig. 3.1.** \( \tau : [0,1] \to [0,1] \) **Fig. 3.2.** \( \tau : [\frac{1}{2^n}, 1] \to [\frac{1}{2^n}, 1] \).

Let \( n \geq 2 \) be fixed, and let us consider the restriction of \( \tau \) on \( [\frac{1}{2^n}, 1] \) (Fig. 3.2). Let \( J \) be a partition of \( [\frac{1}{2^n}, 1] \) given by \( a_0 = \frac{1}{2^n} < \frac{1}{2^{n-1}} < \ldots < \frac{1}{2} < 1 = a_n \). Obviously \( J \) is a Markov partition, with respect to \( \tau \), of \( [\frac{1}{2^n}, 1] \). Moreover \( \tau(I_i) = I_{i+1} \), \( 1 \leq i \leq n \), and \( \tau(I_n) = \bigcup_{i=1}^{n} I_i \), where \( I_i = (a_{i-1}, a_i) \) and \( a_i = \frac{1}{2^{n-1}}, 0 \leq i \leq n \), i.e., \( \tau \) satisfies conditions (\( \varepsilon_1 \)) and (\( \varepsilon_2 \)).
The transformation $\tau$ induces an $n \times n$ matrix $M_\tau$ given by:

$$m_{ij} = \begin{cases} \frac{m(I_i)}{m(\tau(I_i))} \frac{\ell(I_i)}{\ell(\tau(I_i))}, & \text{if } I_j \subset \tau(I_i) \\ 0, & \text{otherwise} \end{cases}$$

From the definition of $\tau$, the $i$th row of $M_\tau$, for $i < n$, has:

zero everywhere except for the $(i+1)$th entry, which is given by,

$$|\tau_i|^{-1} = \begin{vmatrix} \ell(I_i) \\ \ell(\tau(I_i)) \end{vmatrix} = \begin{vmatrix} \frac{1}{2^{n-1}} - \frac{1}{2^{n-1}+1} \\ \frac{1}{2^{n-1}+1} - \frac{1}{2^{n-1}} \end{vmatrix} = \begin{vmatrix} 2-1 \\ 2^{n-1} \end{vmatrix} = \frac{2^{n-1}}{2^{n-1}+1}$$

$$= \frac{1}{2}.$$

The $n$th row of $M_\tau$ has all entries equal given by,

$$|\tau_n|^{-1} = \begin{vmatrix} -\ell(I_n) \\ \ell(\tau(I_n)) \end{vmatrix} = \begin{vmatrix} a_n - a_{n-1} \\ a_n - a_0 \end{vmatrix} = \frac{1 - \frac{1}{2}}{1 - \frac{1}{2^n}} = \frac{2^{n-1}}{2^{n-1}+1}.$$
Therefore, the \( n \times n \) matrix induced by \( \tau \) is:

\[
M_\tau = \begin{bmatrix}
\frac{1}{2} & & \\
& \frac{1}{2} & \\
& & \frac{1}{2} \\
\frac{2n-1}{2^{n-1}} & \cdots & \frac{2n-1}{2^{n-1}}
\end{bmatrix}
\]

Since \( \tau \) is Markov, \( M_\tau \) is irreducible. Hence \( 1 \) is its eigenvalue of maximum absolute value. Indeed, the unique (up to constant multiples) eigenvector is given by \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \), where

\[
\pi_i = \frac{a_i - a_0}{a_i - a_{i-1}} = \frac{2^{n+1} - 2^{n+1-i}}{2^{n+1} - 2^n} = 2 - 2^{-1-i}
\]

When viewed as a density, \( \pi \) can be expressed as:

\[
f(x) = \sum_{i=1}^{n} \left( \frac{2^{n+1} - 2^{n+1-i}}{2^{n+1} - 2^n} \right) X_{I_i}(x)
\]

\[
= \sum_{i=1}^{n} (2 - 2^{-1-i}) X_{I_i}(x),
\]

where \( X_{I_i} \) is the characteristic function on the set \( I_i, 1 \leq i \leq n \).
The area under $f$, for $n \to \infty$, is 1, since,

\[
\lim_{n \to \infty} \int_{\left[\frac{1}{2^n}, 1\right]} f(x)dx = \lim_{n \to \infty} \sum_{t=1}^{n} \frac{1}{2^n} \int_{\left[\frac{1}{2^n}, 0\right]} \pi_{\frac{1}{2^n}} x dx
\]

\[
= \lim_{n \to \infty} \sum_{t=1}^{n} \pi_{\frac{1}{2^n}} \frac{a_t}{a_{t-1}} \int_{a_{t-1}}^{a_t} dx
\]

\[
= \lim_{n \to \infty} \sum_{t=1}^{n} \frac{a_t - a_0}{a_t - a_{t-1}} (a_t - a_{t-1})
\]

\[
= \lim_{n \to \infty} (a_n - a_0)
\]

\[
= 1.
\]
CHAPTER IV

MATRICES INDUCED BY PIECEWISE LINEAR TRANSFORMATION
AND THEIR IN Variant VECTORS

4.1. PRELIMINARIES

In Chapter II, we showed that a piecewise linear transformation \( \tau \in \mathcal{T} \) induces a uniquely determined square matrix \( M_{\tau} \), which is non-negative, irreducible and row contiguous. We also showed that starting from \( M_{\tau} \) we can determine an equivalent class of \( 2^n \) transformations, and also that the problem of finding the unique invariant densities under the piecewise linear transformations \( \tau \in \mathcal{T} \) is equivalent to the problem of finding the invariant vectors under \( M_{\tau} \).

In Chapter III, we showed that each member from the class of non-negative square, irreducible (or primitive) stochastic matrices possesses a unique\(^1\) invariant vector. That means that every square matrix \( M_{\tau} \) which is non-negative, primitive, stochastic and contiguous possesses a unique invariant vector, which can serve as the unique invariant density for each one of the transformations in the class of \( 2^n \) transformations which are determined by \( M_{\tau} \).

In this Chapter, we shall characterize those classes of \( n \)-vectors which can serve as a unique invariant vector for a matrix \( A \) in a class \( A_n \), defined below.

\(^1\)Up to constant multiples.
Definition 4.1.1: We say that the non-negative, $n \times n$ square matrix $A \in A_n$, if $A$ satisfies the following conditions:

(i) $A$ is primitive.
(ii) $A$ is row stochastic.
(iii) $A$ is row contiguous.

Since we shall use some results from graph theory, to check the primitivity of a matrix $A$, and to construct $A_n$-invariant vectors, we introduce in the following section, some basic definitions from graph theory [2, 4, 11, 78].

4.2. ELEMENTS FROM GRAPH THEORY

Directed Graphs

A directed graph $G$ (or digraph) consists of two things:

(i) A set $V$ whose elements are called vertices, or points.
(ii) A set $E$ of ordered pairs of vertices called arcs (edges).

We shall denote a directed graph by $G = (V,E)$. We can picture directed graphs by diagrams in the plane where each vertex $v$ in $V$ is represented by a dot and each arc $a = <u,v>$ is represented by an arrow from the initial (beginning) point $u$ of $a$ to its terminal (ending) point $v$. If the initial point for an arc is the same vertex as its terminal point, then the arc is called a loop.

Since we are concerned only with directed graphs we shall generally omit the word "directed".

We say that $G = (V,E)$ is finite if its set $V$ of vertices is finite and its set $E$ of arcs is finite.
Let $V'$ be a subset of $V$ and let $E'$ be a subset of $E$ whose endpoints belong to $V$. Then $G' = (V', E')$ is a (directed) graph and is called a subgraph of $G = (V, E)$.

A complete graph on $n$-vertices is a graph with the property that every pair of vertices is an arc. A $k$-clique of a graph $G$ is a complete subgraph of $G$ on $k$ vertices.

The outdegree and indegree of a vertex $v$ are equal, respectively, to the number of arcs beginning and ending at $v$. A vertex with zero outdegree is called a sink and a vertex with zero indegree is called a source.

A walk from $v$ to $w$ in a graph $G$ is a sequence of vertices $v_1, v_2, \ldots, v_k$ such that $v = v_1$ and $w = v_k$ and $<v_{i-1}, v_i> \in E$ for all $1 < i < k$. A semiwalk from $v$ to $w$ is a sequence of vertices $v_1, v_2, \ldots, v_k$ such that $v_1 = v$, $v_k = w$ and either $<v_{i-1}, v_i> \in E$ or $<v_i, v_{i-1}> \in E$ for all $1 < i < k$. A (semi) path is a (semi) walk with all vertices distinct.

A graph is strongly connected if for all $v, w \in V$, $v \leftrightarrow w$ there is a path from $v$ to $w$. It is weakly connected if every pair of distinct points is joined by a semipath. A weak component of a graph is a maximal weakly connected subgraph.

**Rooted Trees**

A cycle is a closed walk with distinct vertices except for the first and the last vertices. A semicycle of a graph $G$ is a semipath from a vertex $v$ to itself. A graph $G$ is said to be acyclic or cycle-free if it contains no cycles.
A rooted tree is a weakly connected graph which does not contain semicycles with a designated vertex \( r \), called the root.

A forest is an acyclic graph in which every weak component is a tree.

An in-tree is a rooted tree in which there is a path from every vertex to the root. It is known [11 pg. 201] that a weakly connected graph is an in-tree if and only if exactly one vertex has outdegree 1 and all others have outdegree 1.

A vertex \( w \) is said to be a child of the vertex \( v \) in a rooted tree if the (unique) semipath from the root \( r \) to \( w \) has \( v \) as its penultimate vertex (i.e., a semipath \( r = v_1, \ldots, v_k, v, w \)).

A vertex is a leaf of a rooted tree if it has no children.

The length of a path from the root \( r \) to \( v \) is called the level or depth of \( v \).

Matrices and Graphs

Consider a finite graph \( G \) and let us label the vertices of the graph \( G \) by assigning subscripts to them so that \( V = \{v_1, v_2, \ldots, v_m\} \). It is known [4] that we can represent that graph by a matrix in many different ways (adjacency matrix, reachability matrix, distance matrix).

For the purpose of this thesis, we shall employ the adjacency matrix, which is defined as follows: \( A = A(G) = (a_{ij}) \), where,

\[
a_{ij} = \begin{cases} 
1 & , \text{ if } \langle v_i, v_j \rangle \in E \\
0 & , \text{ otherwise}
\end{cases}
\]

1 \leq i, j \leq m. Thus a graph can be represented by a 0-1 square matrix and, conversely, a 0-1 square matrix can be considered as an
adjacency matrix for some graph.

We elucidate that with the following example.

Example 4.1: A digraph and its adjacency matrix.

\[
A(G) = 
\begin{bmatrix}
0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

Fig. 4.1

From the above example we note that the row sums of \(A(G)\) give the outdegrees of the vertices of \(G\) and the column sums are the indegrees.

We shall continue with some interesting observations for the adjacency matrix of a digraph \(G\).

Theorem 4.2.1: Let \(A = A(G)\) be an adjacency matrix of \(G\). Then, the \((i,j)\)th entry of the matrix \(A^n\) gives the number of walks of length \(n\) from the vertex \(v_i\) to the vertex \(v_j\).

**Proof:** We prove this by induction on \(n\). Note that a walk of length 1 from \(v_i\) to \(v_j\) is precisely an arc \(<v_i, v_j>\). Thus the theorem holds for \(n = 1\), since the \((i,j)\)th entry of \(A\) gives the number of arcs \(<v_i, v_j>\) (note this number is zero or 1), which is the number of walks of length 1 from \(v_i\) to \(v_j\).
Suppose \( n > 1 \), \( A^{n-1} = (a_{ik}) \) and \( A = (b_{kj}) \). Then \( A^n = (c_{ij}) \) is given by \( c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \). By induction, \( a_{ik} \) equals the number of walks of length \( n-1 \) from \( v_i \) to \( v_j \). Also, \( b_{kj} \) equals the number of walks (arcs) from \( v_k \) to \( v_j \). Thus, \( a_{ik} b_{kj} \) gives the number of walks of length \( n \) from \( v_i \) to \( v_j \), where \( v_k \) is the next-to-last vertex in the walk. Thus, all walks of length \( n \) from \( v_i \) to \( v_j \) can be obtained by summing up the \( a_{ik} b_{kj} \) for all \( k \), that is, \( c_{ij} \) is the number of walks of length \( n \) from \( v_i \) to \( v_j \). Thus, the theorem is proved.

Q.E.D.

Now, suppose that \( G \) has \( m \) vertices. Then any path of \( G \) cannot contain more than \( m \) vertices and hence must have length \( m-1 \) or less. Thus, \( G \) is strongly connected if for any vertices \( u \) and \( v \), there exists a path from \( u \) to \( v \) and one from \( v \) to \( u \), each of length \( m-1 \) or less.

**Corollary 4.2.1:** (A criterion for connectedness). Let \( A = A(G) \) be an adjacency matrix of \( G \). Let:

\[
C = A + A^2 + \ldots + A^{m-1}
\]

Then \( G \) is strongly connected if and only if \( C \) has no zero entries off the main diagonal.

**Proof:** The \((i,j)\)th entry of \( C \) is zero if and only if the \((i,j)\)th entries of \( A, A^2, \ldots, A^{m-1} \) are all zero. This means there is no walk from \( v_i \) to \( v_j \) of length \( m-1 \) or less, and so there is no path from \( v_i \) to \( v_j \). Thus \( G \) is strongly connected if and
only if the \((i,j)\)th entries of \(C, i \neq j\), are non-zero. 

Q.E.D.

It is well known [4 page 171] that a matrix \(A = A(G)\) is irreducible if and only if \(G\) is strongly connected. It is also known (Theorem 2.2.1) that a non-negative square matrix \(A\) is primitive if and only if \(A\) is irreducible and has at least one aperiodic index. Thus, using Corollary 4.2.1 we get the following theorem.

**Theorem 4.2.2**: (A test of primitivity) A non-negative \((m \times m)\) matrix \(A\) with \(0-1\) entries is primitive if and only if the matrix

\[
C = A + A^2 + \ldots + A^{m-1}
\]

has no zero entries off the main diagonal and \(a_{ij} > 0\) for some \(i, 1 \leq i \leq m\).

**Proof**: By Corollary 4.2.1, \(G\) is strongly connected and so \(A\) is irreducible, and since \(A\) has an aperiodic index (the index 1), Theorem 2.2.1 confirms that \(A\) is primitive. Q.E.D.

4.3. **PERMUTATION INVARIANT MATRICES**

**Induced 0-1 Matrices**

Let \(A\) be a square matrix with non-negative entries. We can associate with \(A\) a 0-1 square matrix by setting all non-zero entries to 1. We denote this matrix by \(g(A)\). We remark that \(g(A)\) is an adjacency matrix for some graph. Let us recall that \(A_n\) is the class of square \(n \times n\) matrices which are stochastic, primitive and each

\[^1\text{This result is due to the author.}\]
row of which has its non-zero entries contiguous with equal value.

In the following we shall prove three lemmas which characterize the map \( g: A_n \rightarrow g(A_n) \), where the map \( g \) is defined [8] as follows, for \( A' = (a'_{ij}) \in A_n \):

\[
g(a'_{ij}) = \begin{cases} 
0 & \text{if } a'_{ij} = 0 \\
1 & \text{if } a'_{ij} \neq 0 
\end{cases}
\]

\( 1 \leq i, j \leq n \). Obviously \( g(A_n) \) is a subspace of the space of square 0-1 matrices.

**Lemma 4.3.1:** The mapping \( g: A_n \rightarrow g(A_n) \) is 1-1.

**Proof:** Let \( A, A' \in A_n \), where \( A \neq A' \). Let the \( i \)th row of \( A \) contain \( k_i \) non-negative entries and let the \( i \)th row of \( A' \) contain \( k_i' \) non-negative entries. Without loss of generality, we assume in view of the fact that \( A \neq A' \), that \( k_i > k_i' \).

Now assume that \( g(A) = g(A') \). The \( i \)th row of \( g(A) \) must contain a contiguous block of \( k_i \) 1's and so must \( g(A') \), since \( g(A) = g(A') \). Now, since \( A' \) is stochastic and contiguous, its \( i \)th row must contain a block of \( \frac{1}{k_i} \) 1's in the same columns as \( g(A') \), which means that \( A' \) must contain \( k_i' < k_i \) non-negative entries in the \( i \)th row. This contradicts the assumption that \( A' \) contains \( k_i' < k_i \) non-negative entries in the \( i \)th row.

Therefore, if \( A \neq A' \), we have \( g(A) \neq g(A') \), i.e., \( g \) is 1-1.

Q.E.D.

**Lemma 4.3.2:** Let \( A \) be an \( n \times n \) square matrix with non-negative entries and let \( g(A) \) be the corresponding 0-1 matrix.
Then \( g(g(A)^k) = g(A^k) \) for any \( k > 0 \).

**Proof:** Let \( A \) and \( B \) be two, non-negative \( n \times n \) matrices. We know that \( g(A) \) and \( A \) both have non-zero entries in the same positions. The same applies for \( g(B) \) and \( B \). Let us consider the two products \( g(A)g(B) \) and \( AB \), and let us denote by \( \alpha_{ij} \) and \( \beta_{ij} \), \( i,j = 1,\ldots,n \), the entries of \( AB \) and \( g(A)g(B) \), respectively. Then,

\[
\beta_{ij} = \sum_{k=1}^{n} g(a)_{ik} g(b)_{kj},
\]

is not zero if and only if, for at least one \( k, 1 \leq k \leq n \), both \( g(a)_{ik} \) and \( g(b)_{kj} \) are non-zero. But then for the same \( k \)'s, we have both \( a_{ik} \) and \( b_{kj} \) are non-zero, and so \( a_{ik} b_{kj} \) is non-zero. Thus

\[
\sum_{k=1}^{n} a_{ik} b_{kj} = \alpha_{ij}
\]

is non-zero and therefore \( \beta_{ij} \) is non-zero if and only if \( \alpha_{ij} \) is, as well.

Now we shall show, by induction on \( k \), that \( g(g(A)^k) = g(A^k) \).

For \( k = 1 \), it is obvious that \( g(g(A)) = g(A) \). Assume that the statement is true for \( k = n-1 \), i.e., \( g(g(A)^{n-1}) = g(A^{n-1}) \). We shall show that it is true for \( k = n \). From the equality \( g(g(A)^{n-1}) = g(A^{n-1}) \), we see that \( g(g(A)^{n-1}) \) and \( g(A^{n-1}) \) have non-zero entries in the same positions. But the same is true for the matrices \( g(A)^{n-1} \) and \( A^{n-1} \). From the definition of \( g \), it is clear that \( g \) does not change any non-zero entries to zero. Then, by the above argument, \( g(A)^{n-1} \), \( g(A) \) and \( A^{n-1} \) have non-zero entries in the same positions, i.e., \( g(A)^n \) and \( A^n \) have non-zero entries in the same positions and so, by the definition of \( g \), \( g(g(A)^n) = g(A^n) \). Thus
the equality is true and for \( k = n \), and the induction is complete.

Q.E.D.

**Lemma 4.3.3:** Let \( A \in \mathbb{A}_n \). Then \( g(A) \) is primitive.

**Proof:** Since \( A \) is primitive, there exists an integer \( k > 0 \) such that \( A^k > 0 \), i.e.,

\[
\begin{align*}
A^k & \text{ has non-zero entries} \\
g(A^k) & \text{ has non-zero entries} \\
g(g(A^k)) & \text{ has non-zero entries} \\
g(A)^k & \text{ has non-zero entries} \\
g(A) & \text{ is primitive.}
\end{align*}
\]

Q.E.D.

Lemma 4.3.3 tells us that instead of checking the primitivity of a non-negative matrix \( A \), it is equivalent to check the primitivity of \( g(A) \), which is much easier from a computational point of view.

As we mentioned at the beginning of this section, the \( 0 \)-1 matrix \( g(A) \) can serve as the adjacency matrix for some graph \( G \). Since the rows and columns of \( g(A) \) correspond to an arbitrary labelling of the vertices of \( G \), it follows that if \( g(A) \) is an adjacency matrix of \( G \), then so is \( Pg(A)P^T \) where \( P \) is a \( n \times n \) permutation matrix. Moreover, if \( g(A) \) and \( g(B) \) are any two adjacency matrices for \( G \), there exists a permutation matrix \( P \) such that \( g(B) = Pg(A)P^T \) [4]. Therefore it is clear that we shall be interested primarily in those properties of \( g(A) \) which are invariant under permutations of the rows and columns. Since \( Pg(A)P^T = g(PAP^T) \) for \( A \in \mathbb{A}_n \), we shall define a class of matrices \( P_n \subset \mathbb{A}_n \), which is closed under pre- and post-multiplication by permutation matrices. \( P_n \) well be the main object
of our investigation in the sequel.

**Definition 4.3.1**: Let $P_n \subseteq A_n$ be the class of matrices $A \in A_n$ with the property that $PAP^T \in A_n$ for any permutation matrix $P$.

The following results characterize those vectors $\pi$ for which there exists a matrix $A \in P_n$ such that $mA = \pi$. We refer to such vectors as $P_n$-invariant vectors [8].

**Lemma 4.3.4** [8]: Let $A \in P_n$ and let $G = (V,E)$ be the digraph associated with $g(A)$. Then,

i) $G$ contains a $k$-clique, $K$, for some $k \geq 1$, such that there exists an arc from each vertex of the $k$-clique to each vertex of $G$;

ii) if $v$ is a vertex of $G$ not in $K$, $v$ has outdegree 1;

iii) the subgraph of $G$, $\overline{G}$, formed by deleting all arcs $<v,w>$, where $v$ is a vertex of $K$, is a forest of $k$ in-trees with roots in $K$.

Note: part (ii) simply says that the $i$th row of $g(A)$, $i > k$ (for $k \geq 1$) contains only one non-zero entry.

**Proof**: We claim that each row of $g(A)$ consists entirely of 1's or of a single 1. If there was a row, say the $i$th row, which contains a contiguous block of $k_i$ 1's ($k_i \geq 1$, $n$), we could find a permutation matrix $P$ such that one of the rows of $Pg(A)P^T$, say the $j$th, has non-contiguous entries of 1's. This implies that the $j$th row of $g(PAP^T)$ is non-contiguous and similarly the $j$th row of $PAP^T$ is non-contiguous. This implies that $A \notin P_n$, which contradicts the hypothesis.
We now show that \( k \neq 0 \). If \( k = 0 \), then every row in \( g(A) \) has only one non-zero entry. Since \( g(A) \) is strongly connected, it must be a permutation matrix. Then \( g(A)^m \) cannot be primitive for any \( m \). Hence \( A \notin A_n \), contradicting the hypothesis that \( A \in P_n \subset A_n \).

We can, therefore, choose a permutation matrix \( P \) such that \( Pg(A) \) has all its rows of \( 1 \)'s in the first \( k \) positions. Then, since \( A \in A_n \), it follows that \( PAP^T \in A_n \), and so \( g(PAP^T) \) has all its rows of \( 1 \)'s in the first \( k \) positions (note that \( g(PAP^T) = Pg(A)P^T \)).

Let \( K \) be the \( k \times k \) matrix of \( 1 \)'s in the upper left corner of \( g(PAP^T) \), so (i) and (ii) of the lemma are established.

Now, for part (iii) we claim that each weak component (maximal weakly connected subgraph) of the graph \( \bar{G} \) corresponding to \( g(PAP^T) \) with the first \( k \) rows replaced by zeros is an in-tree rooted at a vertex of \( K \). It suffices to show that each weakly connected component has exactly one vertex of outdegree 0, since each vertex of \( \bar{G} \) has outdegree at most 1 and the only vertices of outdegree 0 in \( \bar{G} \) are those of \( K \).

If each vertex has outdegree 1, the component would clearly contain a cycle. Moreover, there would be no arc from the cycle to any vertex of \( K \). This contradicts the fact that \( G \) is strongly connected. Therefore, each weakly connected component has at least one vertex of outdegree 0.

Suppose now that \( v \) and \( w \) are two vertices of outdegree 0 in the same weak component. Let \( v = v_1, v_2, \ldots, v_n = w \) be a semiwalk in \( \bar{G} \).
Since \( v \) has outdegree 0, \( <v_2, v_1> \in E \). On the other hand, since \( w \) has outdegree 0, \( <v_{n-1}, v_n> \in E \). Let \( i \) be the smallest index such that \( <v_i, v_{i+1}> \in E \). Clearly \( 1 < i < n \). Since \( <v_{i-1}, v_i> \not\in E, <v_i, v_{i-1}> \in E \) and \( <v_i, v_{i+1}> \in E \), and hence \( v_i \) has outdegree 2. Therefore \( r = w \).

Now, since there are \( k \) vertices of outdegree 0 in \( G \), there are \( k \) such in-trees.

Q.E.D.

The converse of Lemma 4.3.4 is also true. Before proving that, we note that the inverse mapping \( g^{-1}: g(A_n) \rightarrow A_n \) is well-defined. This is true since, by Lemma 4.3.1, the map \( g \) is 1-1. Moreover, if \( A \in g(P_n) \), then \( g^{-1}(A) \) has rows consisting entirely of \( \frac{1}{n} \)'s or containing exactly one 1, and, to those rows from \( g^{-1}(A) \) which consist entirely of \( \frac{1}{n} \)'s there correspond rows of \( A \) which consist entirely of 1's. This is obvious because \( g^{-1}(A) \) is stochastic and contiguous.

Thus, if each row of \( A \) contains exactly one 1, then it follows that \( g^{-1}(A) = A \).

**Lemma 4.3.5:** Let \( G \) be a graph with properties (i), (ii), (iii) of Lemma 4.3.4, and let \( A \) be an adjacency matrix for \( G \). Then \( g^{-1}(A) \in P_n \).

**Proof:** From (i) and (ii), it is clear that \( A \) has rows consisting entirely of 1's or containing exactly one 1; i.e., \( g^{-1}(A) \) is stochastic and contiguous. Therefore, it suffices to show that \( g^{-1}(A) \in A_n \). This will be true if \( g^{-1}(A) \) is primitive, i.e., if \( G \) is strongly connected and there exists an aperiodic vertex (i.e., a loop).
Let \( v \) and \( w \) be any two vertices of \( G \). If \( v \) is a vertex of \( K \), we know from (i) that there is a path (namely an arc) from \( v \) to \( w \). Otherwise, from (iii) we see that there exists a path to some vertex \( u \) of \( K \), and this path can be extended to \( w \) since there is an arc \( <u,w> \) by (i). Hence, \( G \) is strongly connected.

Since \( k \geq 1 \), \( A \) has at least one row with no zero entries. Thus there exists a vertex \( i \) such that \( a_{ii} > 0 \), and so, \( A \) is primitive. Q.E.D.

A Special Labelling Procedure

There are many ways of labelling a graph. We show the following labelling procedure [8].

Consider a graph \( G \) satisfying Lemma 4.3.4, and let \( \mathcal{F} \) be the forest composed of \( k \) in-trees, numbered \( T_1, T_2, \ldots, T_k \) where \( T_i \) has \( n_i \) vertices.

We label the root of \( T_1 \) as 1, and we let \( 2, 3, \ldots, m_1 \) label the vertices of the first level of \( T_1 \), starting from the left. Then \( m_1+1, m_1+2, \ldots, m_2 \) label the vertices of level 2 and so on, until all the \( n_1 \) vertices of \( T_1 \) are labelled. We label \( T_2 \) as \( n_1+1 \) and use the same procedure to label all the vertices of \( T_2 \). In general, the root of \( T_i \) is labelled as \( b_{i}+1 \) where \( b_{i} = \sum_{j=1}^{i-1} n_j \), and its vertices are labelled, as above, consecutively, adding 1 to the label, as we proceed from left to right, along each level of \( T_i \) and then down to the next level, as shown in Figure 4.2.
Fig. 4.2

The graph of the forest $\bar{G}$.

With this labelling scheme, the adjacency matrix for $\bar{G}$ is a block diagonal matrix of the form:

$$\bar{A} = \begin{bmatrix}
A_1 & 0 \\
0 & A_2 \\
0 & \ddots \\
0 & \ddots & 0 \\
& & & & A_k
\end{bmatrix}$$
where \( A_i = (a_{rs}^i) \) is the adjacency matrix for \( T_i \), and is the form

\[
\begin{bmatrix}
0 \\
\vdots \\
1 \\
\vdots \\
1 \\
\ddots \\
1 \\
0
\end{bmatrix}
\]

The first row of \(-A_i\) has zero entries only, since the first row of \( A_i \) corresponds to the root of the in-tree \( T_i \), which has zero outdegree.
Each other row of \(-A_i\) has exactly one 1, since each other row of \( A_i \) corresponds to a vertex of the in-tree \( T_i \), different from the root, and we know that each such vertex has outdegree 1. We can also see that if a vertex labelled \( \beta \) in \( T_i \) has indegree equal to \( t \), and \( \beta + p + 1, \ldots, \beta + p + t, p < n_i \) are the labelled vertices in \( T_i \), from which there starts an arc ending at the vertex \( \beta \), then \( a_{rs}^i = 1 \) for \( r = \beta + p + 1, \ldots, \beta + p + t \).

Now let \( B = \{1, 1+n_1, 1+n_1+n_2, \ldots, 1+n_1+n_2+\ldots+n_{k-1}\} = \{1 + \sum_{j=1}^{i-1} n_j\}^{k-1} \)

i.e., \( B \) contains the integers which are the numbers corresponding to the roots of the above forest \( G \) according to the above labelling procedure.
Let $U$ be the $n \times n$ matrix, $n = n_1 + n_2 + \ldots + n_k$, in which the $j$th row, $j \in B$, consists entirely of 1's and all other rows are zero, i.e.:

$$
U = \begin{bmatrix}
1 & 1 & 1 & \ldots & 1 \\
0 & & & & \\
1 & 1 & 1 & \ldots & 1 \\
0 & & & & \\
1 & 1 & \ldots & & 1 \\
0 & & & & \\
\end{bmatrix}
$$

1st row

$(l+n_1)$th row

$(l+n_1+\ldots+n_{k-1})$th row
Now, let $\mathbf{P} = \mathbf{A} + \mathbf{U}$, i.e.:

\[
\begin{bmatrix}
1 & 1 & \cdots & 1 & 1 \\
1 & 1 & \cdots & 1 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
1 & 1 & \cdots & 1 & 1 \\
1 & 0 & \cdots & 0 & 1 \\
0 & 1 & \cdots & 1 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
1 & 1 & \cdots & 1 & 1 \\
0 & 0 & \cdots & 0 & 1 \\
1 & 1 & \cdots & 1 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1 \\
1 & 1 & \cdots & 1 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\end{bmatrix}
\]

- 1st row
- $(1 + n_1)$th row
- $(1 + n_1 + n_2)$th row
- $(1 + n_1 + \cdots + n_{k-1})$th row
- nth row
From the labelling procedure of $G$ and the construction of $\overline{P}$, it follows that $\overline{P}$ is an adjacency matrix of $G$.

Consider the matrix $g^{-1}(\overline{P})$, where,

$$
\begin{bmatrix}
\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\
\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\
\vdots & & & \vdots \\
\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}
\end{bmatrix}_{1 \text{ th row}}
\begin{bmatrix}
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
\vdots & & & \vdots \\
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n}
\end{bmatrix}_{(1+n_1) \text{ th row}}
\begin{bmatrix}
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
\vdots & & & \vdots \\
\frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n}
\end{bmatrix}_{(1+n_1+\ldots+n_{k-1}) \text{ th row}}
\begin{bmatrix}
0 & 0 & \cdots & 0 & 1 \\
0 & 0 & \cdots & 0 & 1 \\
\vdots & & & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{bmatrix}
$$

By Lemma 4.3.5, $g^{-1}(\overline{P}) \in P_n$.

Characterization of $P_n$-Invariant Vectors

For matrices of the form of $g^{-1}(\overline{P})$, it is not difficult to characterize the vectors $\pi$ such that $\pi g^{-1}(\overline{P}) = \pi$. This leads to the following characterization of $P_n$-invariant vectors:
Theorem 4.3.1: [8] Let $\overline{G} = (V,E)$ be a forest of $k$ rooted in-trees. Let $c$ be a constant and let $w: V \rightarrow \mathbb{R}$ be a real valued function which assigns weights to the vertices of $\overline{G}$ as follows:

(i) if $v \in V$ has indegree 0, then $w(v) = c$.

(ii) if $v \in V$ has indegree $m \neq 0$ and $v_1, v_2, \ldots, v_m$ are the vertices from which there starts an arc ending at $v$, then

$$w(v) = c + \sum_{i=1}^{m} w(v_i).$$

Suppose $v_1, v_2, \ldots, v_n$ is a labelling of the elements of $V$, then the vector $(w(v_1), w(v_2), \ldots, w(v_n))$ is $P_n$-invariant. Moreover, for any $P_n$-invariant vector $\pi$, there is a forest $\overline{G}$ and a labelling of the vertices such that $\pi = (w(v_1), w(v_2), \ldots, w(v_n))$.

Proof: Let us label the vertices of $\overline{G}$ following the above procedure, and let this be $v_1, v_2, \ldots, v_n$. Let $w(v)$ be the weight assigned to the vertex $v$, as defined in (ii) of the statement of the theorem, let $A$ and $\overline{P}$ be the adjacency matrix for $G$ and $\overline{G}$, respectively, as constructed above. We claim that $\pi = (w(v_1), w(v_2), \ldots, w(v_n))$ satisfies $\pi \cdot g^{-1}(\overline{P}) = \pi$. To see this we assume the following:

a) Let $v$ be a root of the $\text{i}^{\text{th}}$ in-tree of $\overline{G}$ possessing $n_i$ vertices $1 \leq i \leq k$. Assume that with the above labelling procedure this in-tree has $q_i$ levels, and let the $j^{\text{th}}$ level has $h_j$ vertices, $j = 1, 2, \ldots, q_i$. Then it is clear that $\sum_{j=1}^{q_i} h_j = n_i - 1$ and, by (ii), for the root $v$, we have
\[
\begin{align*}
    w(v) &= c + \sum_{i=1}^{h_1} w(v_i), \\
    \text{and} & \\
    w(v_i) &= c + \sum_{r=1}^{t_i} w(v_r), \\
\end{align*}
\]
where \( t_i \) represents the number of vertices in level 2, from which there originates an arc ending at \( v_i \). It is clear that \( \sum_{i=1}^{h_1} t_i = h_2 \).

Thus we have the following:

\[
\begin{align*}
    w(v) &= c + \sum_{i=1}^{h_1} \left( c + \sum_{r=1}^{t_i} w(v_r) \right) \\
    &= c + \sum_{i=1}^{h_1} c + \sum_{r=1}^{t_i} w(v_r) \\
    &= c + h_1 c + \sum_{r=1}^{h_2} w(v_r), \\
\end{align*}
\]

where \( v_r \) is a new rearrangement of the elements of level 2.

Continuing in this way, we get:

\[
\begin{align*}
    w(v) &= c + h_1 c + h_2 c + \ldots + h_{q-1} c + \sum_{e=1}^{h_q} w(v_e). \\
\end{align*}
\]

But since each vertex on the \( q \) level has indegree 0, by (i) we have \( w(v_e) = c \). Therefore:

\[
\begin{align*}
    w(v) &= c + h_1 c + h_2 c + \ldots + h_q c \\
    &= c + \left( \sum_{j=1}^{q} h_j \right) c \\
    &= c + (n_1 - 1) c \quad \checkmark \\
    &= n_1 c \quad \checkmark
\end{align*}
\]
Now, let $B = g^{-1}(\overline{F})$ and $D = g^{-1}(U)$, if $A_1$ is an $n_i \times n_i$ matrix as defined above, since each row of $A_1$ contains exactly one 1, then $g^{-1}(A_1) = A_1$. Therefore we have

$$B = g^{-1}(\overline{F}) = g^{-1}(\overline{A} + U) = g^{-1}(\overline{A}) + g^{-1}(U) = \overline{A} + D.$$

Now, let $\pi = (\pi(1), \pi(2), \ldots, \pi(k))$, where $\pi(i) = (w(v_{b_i + 1}), w(v_{b_i + 2}), \ldots, w(v_{b_i + n_i}))$, $i = 1, 2, \ldots, k$. We note that the vector $\pi$ has $k$ components, i.e., a number equal to the number of roots (or in-trees) in the forest $\overline{G}$, and that each component $\pi^{(i)}$ has length $n_i$, i.e., equal to the number of vertices in the in-tree $T_i$. Therefore, it is clear that $\pi$ has length $n = n_1 + n_2 + \ldots + n_k$. Then,

$$\pi B = \pi(\overline{A} + D) = \pi \overline{A} + \pi D = (\pi(1), \pi(2), \ldots, \pi(k)) \overline{A} + \pi D = (\pi(1)A_1, \pi(2)A_2, \ldots, \pi(k)A_k) + \pi D = (\pi(1)A_1 + \pi D, \pi(2)A_2 + \pi D, \ldots, \pi(k)A_k + \pi D),$$

but,
\[ \pi_D = (n^{(1)}, n^{(2)}, \ldots, n^{(k)}) g^{-1}(U) \]
\[ = (w(v_{b_1+1}), w(v_{b_1+2}), \ldots, w(v_{b_1+n_1}), w(v_{b_2+1}), \ldots, w(v_{b_k+n_k})) g^{-1}(U) \]
\[ = (w(v_1), w(v_2), \ldots, w(v_{n_1}), w(v_{n_1+1}), \ldots, w(v_n)) g^{-1}(U) \]
\[ \]
Thus, we obtain,

$$\Pi B = (\pi(1)A_1+c, \pi(2)A_2+c, \ldots, \pi(k)A_k+c). \quad (4.1)$$

b) Let \( v_r \) be a vertex of \( T_i \) with indegree 0. Then

$$\pi_r = w(v_r) = c,$$

and

$$c = (\Pi B)_r = c + (\pi(1)A_1)_{r-b_1} \quad \text{from (4.1)},$$

$$= c + \sum_{j=1}^{n_i} \pi_j a_j a_{j,r-b_1}.$$

Since \( v_r \) has indegree zero and so \( a_{j,r-b_1} = 0 \), \( j, j=1,2,\ldots,n_i \).

Hence,

$$c = (\Pi B)_r = \pi_r.$$

c) Let \( v_r \) be a vertex of \( T_i \), with indegree \( m \), and let

$$v_{i+1}, \ldots, v_{i+m}$$

be the vertices from which there starts an arc ending at \( v_r \). Then,

$$\pi_r = w(v_r)$$

$$= c + \sum_{\ell=1}^{m} w(v_{i+\ell}) \quad \text{by (4.1)}.$$
\[(πB)_r = c + (π^1 A_i)_{r-b_i} \]
\[= c + \sum_{j=1}^{n_i} \delta^j a_{j,r-b_i} \]
\[= c + \sum_{\ell=1}^{m} \delta^{i+\ell-b_i} \]
\[= c + \sum_{\ell=1}^{m} w(v_{i+\ell}) \]
\[= \pi_r \]

Hence,
\[(πB)_r = \pi_r \]

Therefore, the claim is established.

Now let \( \pi' \) be a vector of weights corresponding to a different labelling of the vertices. Then it is obvious that the elements of \( \pi' \) are a permutation of those of \( \pi \), so there exists a permutation matrix \( P \), such that,
\[\pi' = \pi P \]

Thus,
\[\pi' = \pi P \]
\[= \pi B P \]
\[= \pi P P^T B P \]
\[= \pi (P^T B P)^T \]

and, since \( B \in P_n \) (note \( B = g^{-1}(P) \), \( P^T B P \in P_n \). Therefore from \( \pi' P^T B P = \pi' \), we deduce that \( \pi' \) is \( P_n \)-invariant.
Conversely, let \( \pi' \) be a \( P_n \)-invariant vector and let \( B' \in P_n \) be a matrix such that \( \pi'B' = \pi' \). We know that there exists a permutation matrix \( P \) such that \( PB'P^T = B \) is of the form \( g^{-1}(P) \).

Let \( \pi = \pi'P^T \) and \( c = \sum_{i=1}^k \pi_{b_i+1} \). Accordingly if \( v_r \) is a vertex of \( \overline{G} \) with indegree \( m \) and \( v_{i+1}, \ldots, v_{i+m} \) vertices of \( \overline{G} \) from which there starts an arc ending at \( v_r \), then

\[
\pi_r = c + \sum_{\ell=1}^{m} w(v_{i+\ell})
\]

\[
= c + \sum_{\ell=1}^{m} \pi_{i+\ell}
\]

Therefore, \( \pi \) consists of weights as described in (ii) and \( \pi = \pi P^T \) is a permutation of those weights.

Q.E.D.

**Examples**

**Example 4.2:**

![Diagram](image)

Fig. 4.3.

Fig. 4.3 shown the weight assignment for a forest consisting of two in-trees.
Example 4.3:

Fig. 4.4 shows the labelling of the vertices of an in-tree and the weight assignment of each vertex.
\[ B = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

\[ \pi = (13, 4, 4, 4, 1, 1, 1, 1, 1, 1, 1) \]

It can be easily shown that \( \pi B = \pi \), i.e., that the vector \( \pi \) is \( P_n \)-invariant.

**Example 4.4:**

![Diagram](Fig. 4.5)
\[ B = \begin{bmatrix}
    1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
    0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
    1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix} \]

\( \pi = (5, 3, 1, 1, 1, 4, 1, 2, 1, 2, 1) \), and \( \pi B = \pi \)

**Example 4.5:**

\[ v_1 : nc \]
\[ v_2 : (n-1)c \]
\[ \cdots \]
\[ v_{n-1} : 2c \]
\[ v_n : c \]

\[ \pi = (n, n-1, \ldots, 2, 1) \] and \( \pi B = \pi \)
Example 4.6: Since the vector $\pi = (n, n-1, \ldots, 2, 1)$ is $P_n$-invariant, $\pi' = P\pi$ is also $P_n$-invariant, where $P$ is a permutation matrix. Therefore, the vector $\pi = (1, 2, \ldots, n-1, n)$ is $P_n$-invariant, with the corresponding forest $\mathcal{G}$ and the matrix $B'$ given by,

$$
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
$$

Fig. 4.7

4.4. EXTENSION OF $P_n$-INVARIANT VECTORS

It is known (example 4.6) that the vector $\pi = (1, 2, \ldots, n)$ is $P_n$-invariant, and therefore it is the unique invariant density of a transformation $\tau$ which can be constructed. Suppose now, that we need a vector $\pi$ with the $j$th term equal to $mj$ rather than $j$, where $m$ is a positive integer. We shall see that it is possible to construct a new $\tau$ for which this is the case.

Theorem 4.4.1: Let $n, m$ and $j$ be fixed positive integers satisfying $1 < m < n$, $m < j+1 < n$. Then the $n$-vector

$$
\pi = (1, 2, \ldots, m-1, j \cdot (m-1), 2j - (m-2), \ldots, (m-1) \cdot j - 1, mj, j+1, \ldots, n)
$$
is $A_n$-invariant, i.e., there exists a matrix $A \in A_n$ such that

$$mA = n.$$ 

Proof: Let us construct the matrix $A = (a_{ij})$ as follows:

$$a_{i,j+1} = 1, \quad i \neq j$$

$$a_{j,k} = \begin{cases} 
\frac{1}{m}, & j-(m-2) \leq k \leq j+1 = j-(m-(m-1)) \\
0, & \text{otherwise}
\end{cases}$$

$$a_{n,s} = \frac{1}{n}, \quad s = 1, 2, \ldots, n$$

i.e.,

$$A = \begin{bmatrix}
0 & 1 & & & \\
& & 0 & & \\
& & & 1 & \\
& & & & \frac{1}{m} & \ldots & \frac{1}{m} \\
& & & & \frac{1}{n} & & \frac{1}{n}
\end{bmatrix}$$
First, we shall show that \( A \in A_n \):

(i). By construction, \( A \) is contiguous.

(ii) \( A \) is stochastic, since, for \( k = n \),

\[
\sum_{i=1}^{n} a_{n,i} = \sum_{i=1}^{n} \frac{1}{n} = 1.
\]

for \( k \neq j \),

\[
\sum_{i=1}^{n} a_{k,i} = a_{k,k+1} = 1.
\]

and for \( k = j \),

\[
\sum_{i=1}^{n} a_{j,i} = \sum_{t=1}^{m} \frac{1}{m} = \sum_{t=1}^{m} \frac{1}{m} = 1.
\]

(iii) To show that \( A \) is primitive, we consider the matrix \( B' \),

where \( B' \) is the matrix associated with \( \pi' \) as constructed in Example 4.6. It is known that \( B' \) is primitive.

Clearly, \( A \) has a non-zero entry corresponding to each non-zero entry of \( B' \) (plus others in the \( j \)th row). Hence, \( g(A) \) is strongly connected and possesses an aperiodic vertex. Thus, \( A \) is primitive.

"We claim now that \( \pi A = \pi \). Let \( m < j \). Then, for \( i = 1 \),

\[
\pi_{r} a_{r+1} = \pi_{r+1} = \frac{1}{n} = 1 = \pi_{1}.
\]

For \( 1 < i \leq j-(m-1) \),

\[
\pi_{r} a_{i+1} = \pi_{i+1} a_{i+1} = \pi_{i+1} a_{i+1} + \pi_{i} = \pi_{i+1} + 1 = \pi_{i+1} + 1 = \pi_{i+1}.
\]

For \( i = j-(m-1) \),

\[
\pi_{r} a_{j} = \pi_{j} = \pi_{j}.
\]
For $1 = j - (m-k), 2 \leq k \leq m$, i.e., for $j-(m-2) \leq i \leq j-(m-m) = j$,

$$
\sum_{r=1}^{n} \pi_r a_{r,j-(m-k)} = \pi_{j-(m-(k-1))} a_{j-(m-(k-1)),j-(m-k)} + \pi_{j-(m-k)} + \pi_{n, n} \frac{1}{n}
$$

$$
= [(k-1)j-(m-(k-1))m] + mj \frac{1}{m} + 1
$$

$$
= kj-(m-k)
$$

$$
= \pi_{j-(m-k)}
$$

For $j+1 \leq i \leq n$,

$$
\sum_{r=1}^{n} \pi_r a_{r,i} = \pi_{i-1} a_{i-1,i} + \pi_{n, n} \frac{1}{n}
$$

$$
= \pi_{i-1} + 1
$$

$$
= i-1 + 1 = i = \pi_{i}
$$

Thus, we have shown that $\pi A = \pi$.

Q.E.D.

**Corollary 4.4.1:** Let $n,m$ be fixed positive integers and let $j = m-1$. Then,

$$
\pi = (m,2m,3m,\ldots,(m-1)m,m,m+1,\ldots,n)
$$

is $A_n$-invariant.

**Proof:** Let $m \neq j+1$ in Theorem 4.4.1. Then the vector

$$
\pi = (1,2,\ldots,j-(m-1),2j-(m-2),\ldots,mj,j+1,\ldots,n)
$$

becomes

$$
\pi = (m,2m,\ldots,(m-1)m,m,m+1,\ldots,n).
$$

Q.E.D.
Example 4.7: Let \( n = 6, m = 3, j = 4 \). Then the vector \( \pi = (1, 2, 7, 12, 5, 6) \) is \( A_n \)-invariant, where

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 3 & 3 & 3 & 0 \\
1 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}
\]

Example 4.8: Let \( n = 6, m = 4, j = 4 \). Then the vector \( \pi = (4, 8, 12, 4, 5, 6) \) is \( A_n \)-invariant, where

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 \\
4 & 4 & 4 & 4 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

Example 4.9: Letting \( m = 2 \), Theorem 4.4.1 proves that the vector \( \pi = (1, 2, \ldots, j-1, 2j, j+1, \ldots, n) \) is \( A_n \)-invariant where,
Corollary 4.4.2: Let \( \pi = (\pi(1), \pi(2), \ldots, \pi(k)) \), where \( \pi(1) \)

is a vector of length \( n_1 \) and is obtained from the vector

\( (k_1, k_1+1, \ldots, k_1+n_1-1) \), where \( k_1 = \sum_{j=1}^{i-1} n_j \), by multiplying \( k_1+j_1 \) by

\( m_i \), where \( 1 \leq m_i < n_i \), \( m_i \leq j_i < n_i \). Then the vector \( \pi \) is \( A_n \)-invariant.

Proof: The construction of Theorem 4.4.1 can be applied to each row \( k_1+j_1 \) of \( A \). Since none of the modified entries overlap, the proof goes through exactly as in Theorem 4.4.1.

Q.E.D.
4.5. **A SPECIAL CONSTRUCTION**

In this section, we discuss an interesting and useful subclass of $A_n$-invariant vectors.

**Theorem 4.5.1**: Let $m, n, k \in \mathbb{N}$ such that $m = \frac{n}{k}$. Then the vector $\mathbf{v} = (m, m+1, \ldots, n)$ is $A_n$-invariant.

**Proof**: Let us construct a square matrix $A$ with $(k-1)m+1 = n-(m-1)$ rows, as follows:

For $0 \leq p \leq k-2$,

$$a_{pm+1,j} = \frac{1}{(p+1)n}, \quad k-p \leq j \leq k + (p+1)(m-1)$$

$$a_{pm+1,j} = 0, \quad otherwise$$

For $p = k-1$,

$$a_{pm+1,j} = \frac{1}{(k-1)m+1}, j = \frac{1}{k}, \quad i \leq j \leq k$$

i.e.,
A is clearly contiguous. It is also stochastic, since for $t = pm + 1$,

$$0 \leq p \leq k - 2,$$

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = \sum_{j=1}^{(k-1)m+1} \frac{1}{(p+1)m} = \sum_{j=k-p}^{k-1} \frac{1}{(p+1)m} = \sum_{j=1}^{(p+1)m} \frac{1}{(p+1)m} = 1,$$

for $t = (k-1)m+1$,

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = \sum_{j=1}^{(k-1)m+1} \frac{1}{k} = 1.$$
for $i = pm+r$, where $0 \leq p \leq k - 2$ and $2 \leq r \leq m$, 

$$\sum_{j=1}^{(k-1)m+1} a_{ij} = a_{pm+r, k+p(m-1)+r-1} = 1.$$ 

Thus $A$ is stochastic. We shall now show that $A$ is primitive. To do this we shall prove that the graph $G$, which corresponds to $g(A)$ is strongly connected and has an aperiodic index.

a) We claim that for all $i \neq \ell$, there exists a path from vertex $v_i$ to the vertex $v_\ell$, where $\ell = (k-1)m+1$. It is clear that for each $i \neq \ell$, there exists $j > i$ such that $g(a)_{ij} = \bar{P}_{ij} = 1$, since for each $i$, the $i$th row of $A$ contains at least one non-zero entry.

Let $f(i)$ be the smallest such $j$. Similarly for that $f(i)$ (assuming that $f(i) \neq \ell$, since where $f(i) = \ell$ the same result is true), there exists $j > f(i)$ such that $\bar{P}_{f(i), j} = 1$. Let $f^2(i)$ be the smallest such $j$. Continuing the process, we have $i, f(i), f^2(i), \ldots, f^k(i)$, which defines a path, starting at $i$ and visits vertices of strictly increasing number. Since there are a finite number of vertices, the sequence must terminate. Moreover, it is clear that the sequence terminates when $f^k(i) = \ell$.

b) Since there are arcs from vertex number $\ell$ to vertices $1, 2, \ldots, k$, the set of vertices $v^(-1) = \{1, \ldots, k, \ell\}$ is strongly connected. From (a), there exists a path from all other vertices into this set (specifically $\ell$). Therefore, it suffices to show that there are paths from this set to the remaining vertices.

Let $v(s) = \{k+s(m-1)+1, k+s(m-1)+2, \ldots, k+s(m-1)+m-1\}$, where $0 \leq s \leq k-2$. Now, there exists an arc from vertex $1$ to each vertex in
\[ v^{(0)} = \{k+1, k+2, \ldots, k+m-1\} \] (which is obvious from the matrix \( A \)).

Therefore, \( v^{(-1)} \cup v^{(0)} \) is strongly connected. We proceed by induction. Suppose \( v^{(-1)} \cup v^{(0)} \cup \ldots \cup v^{(i-1)} \), \( i \leq k-3 \), is strongly connected. We claim that \( \cup_{j=-1}^{i} v^{(j)} \) is strongly connected. To prove this, we note that there are arcs from the vertex \( im+1 \) to all vertices in \( v^{(1)} \). Therefore, it suffices to show that

\[ \cup_{j=-1}^{i-1} \left( im+1 \cup \bigcup_{j=-1}^{i-1} v^{(j)} \right) \] is strongly connected. By the construction, this is true if

\[ im+1 \leq k+i(m-1) \] (i.e., if \( i \leq k-1 \)), which is certainly true.

Now, we must show that \( P = g(A) \) has an aperiodic vertex. For \( k \geq 2 \), \( P \) can take on the value \( k-2 \). Hence, from

\[ P_{pm+1,j} = \frac{1}{(p+1)m} \left( k-p \leq j \leq k+(p+1)(m-1) \right) \]

we obtain

\[ P_{(k-2)m+1,j} = \frac{1}{(k-1)m} \quad 2 \leq j \leq k+(k-1)(m-1) \]

Therefore, the \( [(k-2)m+1] \)th row of \( P \) intersects the diagonal since \( P_{k+(k-2)(m-1)} \geq (k-2)m+1 \) for \( k \geq 2 \). Hence, there exists an \( i \) such that \( P_{ii} > 0 \). Consequently, \( P \) is primitive, and so is \( A \).

To complete the proof, we have to show that \( \pi A = \pi \), for \( j = 1 \)

\[ \sum_{i=1}^{(k-1)m+1} \pi_i a_{ij} = \pi_{(k-1)m+1} \]

Therefore, \( \frac{1}{k} = m = \pi_1 \).
for $2 \leq j \leq k$,

$$
\sum_{i=1}^{(k-1)m+1} \pi_i a_{i,j} = \sum_{p=0}^{k-2} \pi_{pm+1} a_{pm+1,j} + \sum_{p=k-j}^{k-2} \pi_{pm+1} \frac{1}{(p+1)m} + m \quad \text{(since $k-pz-j$ or $p=k-j$)}
$$

$$
= \sum_{p=k-j}^{k-2} \frac{(p+1)m}{(p+1)m} + m
$$

$$
= j - 1 + m
$$

$$
= \pi_j
$$

for $(k+1) + s(m-1) \leq j \leq k + (s+1)(m-1)$, where $0 \leq s \leq k-2$,

$$
\sum_{i=1}^{(k-1)m+1} \pi_i a_{i,j} = \sum_{p=s}^{k-2} \pi_{pm+1} a_{pm+1,j} + \pi_{j-k+s+1} a_{j-k+s+1,j}
$$

$$
= \sum_{p=1}^{k-s-1} \pi_{j-k+s+1} a_{j-k+s+1,j}
$$

$$
= (k-s-1) + (j-k+s+m)
$$

$$
= m + j - 1
$$

$$
= \pi_j
$$

Hence, $\pi A = \pi$. This completes the proof that $\pi$ is $A_n$-invariant.

Q.E.D.

Remark 1: For $k=2$, the matrix $A$ becomes an $(m+1) \times (m+1)$ matrix of the form.
where the 1's are on the superdiagonal. It is clear that the $(m+1)$th entries on the diagonals of $A^2$ and $A^3$ are greater than 0. Thus, the set \( \{ n \geq 1 : a_{m+1,m+1}^{(n)} > 0 \} = \{2,3\} \), and therefore, \( \text{g.c.d} \{ n \geq 1 : a_{m+1,m+1}^{(n)} > 0 \} \) is equal to 1. Hence, the $(m+1)$th vertex is aperiodic. The graph of $g(A)$ is strongly connected. Therefore $P$ is primitive and so is $A$.

Remark 2: For $k = n$, the matrix $A$ becomes a skew-triangular matrix,

\[
A = \begin{bmatrix}
1 \\
\frac{1}{2} & \frac{1}{2} \\
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{n} & \cdots & \cdots & \cdots & \frac{1}{n}
\end{bmatrix}
\]

Clearly, $A \notin P_n$. On the other hand, the matrix obtained for the same vector using the construction of section 4.3 is
\[
A' = \begin{bmatrix}
1 & 0 & & \\
0 & & & \\
\frac{1}{n} & & & 1 \\
\frac{1}{n} & \cdots & \frac{1}{n}
\end{bmatrix}
\]

where the 1's are on the superdiagonal, i.e., \(A' \in P_n\).

Example 4.10: Let \(m=4\), \(n=12\). Then we have \(k=3\) and

\[n \neq \{4,5,6,7,8,9,10,11,12\}\]. The vector \(\pi\) is \(A_n\)-invariant, where

the associated matrix \(A\) is given by.

\[
0 \quad 0 \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad 0 \quad 0 \quad 0 \\
0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \\
0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \\
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \\
0 \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \\
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \\
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \\
0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \\
\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \\
\]

We can generalize Theorem 4.5.1. This generalization is given

as a corollary.
Corollary 4.5.1: Let \( m, n, j \in \mathbb{N} \) be fixed, such that \( n = m + \ell_j \), \( j \) divides \( m \) (i.e., \((k-1)m = \ell_j\)). Then the vector \( \pi = (m, m+j, m+2j, \ldots, n) \) is \( \Lambda_n \)-invariant.

Proof: We construct the \( [(k-1)m_j + 1] \times [(k-1)m_j + 1] \) matrix \( A \) in a directly analogous manner to that in Theorem 4.5.1. Namely, as follows.

For \( 0 \leq p \leq k-2 \):

\[
\begin{align*}
a_{p, \frac{m}{j} + 1, s} &= \begin{cases} 
\frac{1}{(p+1)m_j}, & k-p \leq s \leq (p+1)(\frac{m}{j} - 1) + k \\
0, & \text{otherwise}
\end{cases} \\
a_{p, \frac{m}{j} + r, s} &= \begin{cases} 
1, & \text{if } j = k+p(\frac{m}{j} - 1) + \pi - 1 \\
0, & \text{otherwise}
\end{cases}
\end{align*}
\]

where \( 2 \leq r \leq \frac{m}{j} \).

For \( p = k+1 \):

\[
\begin{align*}
a_{(k-1)m_j + 1, s} &= \begin{cases} 
\frac{1}{k}, & 1 \leq s \leq k \\
0, & \text{otherwise}
\end{cases}
\end{align*}
\]

The proof that \( A \in \Lambda_n \) is analogous to the proof of this fact in Theorem 4.5.1.

Q.E.D.
Theorem 4.6.2: Let $\pi$ and $\sigma$ be $A_n$-invariant vectors of lengths $n_1$ and $n_2$ respectively, such that $\pi A^{(1)} = \pi$ and $\sigma A^{(2)} = \sigma$.

Suppose that the following two conditions hold, for some $i$ and $j$:

(i) $a_{i,n_1}^{(1)} = \frac{1}{x}$, $a_{j,1}^{(2)} = \frac{1}{y}$, $x, y \in N$, $x \leq n_1$, $y \leq n_2$.

(ii) $\pi_i = kx$, $\sigma_j = ky$, where $k \in N$.

Then $(\pi, \sigma)$ is $A_n$-invariant.

Proof: First we note that since $A^{(1)}$ and $A^{(2)}$ belong to $A_n$, $A^{(1)}$ and $A^{(2)}$ are both stochastic and contiguous. Then from (i) we have,

$$a_{i,n_1}^{(1)} = \frac{1}{x} \text{ implies that } a_{i,s}^{(1)} = \frac{1}{x} \text{ for } n_1 - x + 1 \leq s \leq n_1,$$

and

$$a_{j,1}^{(2)} = \frac{1}{y} \text{ implies that } a_{j,s}^{(2)} = \frac{1}{y} \text{ for } 1 \leq s \leq y.$$

Let us define the matrix $A$ as follows:

$$A = \begin{bmatrix} A^{(1)} & D^{(1)} \\ D^{(2)} & A^{(2)} \end{bmatrix}$$
where,

\[ a_{i,s}^{(1)} = \begin{cases} \frac{1}{x+y}, & 1 \leq s \leq n_1 \cap x \leq i \leq s \leq n_2, \\ 0, & \text{otherwise} \end{cases} \]

\[ a_{r,s}^{(1)} = \frac{1}{x+y}, \quad \forall \, r \neq i, \quad a_{i,s}^{(1)} = \begin{cases} \frac{1}{x+y}, & 1 \leq s \leq y \\ 0, & \text{otherwise} \end{cases} \]

\[ a_{r,s}^{(2)} = a_{r,s}, \quad \forall \, r \neq j, \quad a_{j,s}^{(2)} = \begin{cases} \frac{1}{x+y}, & 1 \leq s \leq y \\ 0, & \text{otherwise} \end{cases} \]

\[ d_{r,s}^{(1)} = 0, \quad \forall \, r \neq i, \quad d_{i,s}^{(1)} = \begin{cases} \frac{1}{x+y}, & 1 \leq s \leq y \\ 0, & \text{otherwise} \end{cases} \]

\[ d_{r,s}^{(2)} = 0, \quad \forall \, r \neq j, \quad d_{j,s}^{(2)} = \begin{cases} \frac{1}{x+y}, & 1 \leq s \leq y \\ 0, & \text{otherwise} \end{cases} \]

Let \((\pi, \sigma)A = (\rho, \mu)\), where \(\rho\) and \(\mu\) are of lengths \(n_1\) and \(n_2\) respectively. Then,

\[ \rho = \sum_{r=1}^{n_1} \pi_{rs} a_{rs}^{(1)} + \sum_{r=1}^{n_2} \sigma_{rs} d_{rs}^{(2)} \]

\[ \pi_{rs} a_{rs}^{(1)} = \pi_s, \quad \text{if} \quad s \notin [n_1-x+1, n_1] \]

\[ \pi_{rs} a_{rs}^{(1)} + \pi_{js} a_{js}^{(1)} + \sigma_{js} d_{js}^{(2)}, \quad \text{if} \quad s \in [n_1-x+1, n_1] \]
\[
\begin{align*}
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} + \pi_j a^{(1)}_{js} + \frac{n_1}{x+y} \sum_{r=1}^{n_1} \frac{\sigma_j}{x+y}, & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} + \frac{ky}{x+y}, & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} + k, & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} + \frac{n_1}{x}, & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} + \pi_j a^{(1)}_{is}, & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\left\{ \begin{array}{ll}
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \notin [n_1-x+1, n_1] \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s & \text{if } s \in [n_1-x+1, n_1]
\end{array} \right. \\
\sum_{r=1}^{n_1} \pi_r a^{(1)}_{rs} = \pi_s, & \text{if } s \in [n_1-x+1, n_1]
\end{align*}
\]
\[ = \pi_s \text{ for each } s, 1 \leq s \leq n_1. \]

Similarly, we can prove that \( \mu_s = \sigma_s \) for each \( s, 1 \leq s \leq n_2. \)

Therefore, we have \((\rho, \mu) = (\pi, \sigma), \) i.e.,

\[(\pi, \sigma)A = (\pi, \sigma).\]

Now, from the definition of \( A, \) we can see that \( A \) is stochastic and contiguous. To see that \( A \) is primitive, we consider the matrix \( \overline{A} = g(A) \) and the corresponding graph \( G = g(A). \) We label the vertices of \( g(A) \) by \( a_1, \ldots, a_{n_1} \) and \( b_1, \ldots, b_{n_2} \) as in the proof of Theorem 4.6.1.

Since \( g(\overline{A}^{(1)}) = g(A^{(1)}) \) and \( g(\overline{A}^{(2)}) = g(A^{(2)}) \), in order to show that \( g(A) \) is strongly connected, it suffices to show that for \( r, 1 \leq r \leq n_1 \) and for \( s, 1 \leq s \leq n_2 \), there are paths \( a_r \to b_s \) and \( b_s \to a_r \). We omit this, since it is analogous to the proof of Theorem 4.6.1.

Finally, let \( Q^{(n)} \) represent the upper left \( n_1 \times n_1 \) block of \( A^n. \) Since \( g(\overline{A}^{(1)}) = g(A^{(1)}) \), it is clear that \( Q^{(n)} \geq (A^{(1)})^n \) and hence \( \overline{A} \) has one aperiodic state. Thus, \( A \) is primitive. Therefore, \( (\pi, \sigma) \) is \( A_n \)-invariant.

Q.E.D.

**Example 4.13:** Let \( \pi = (1, 2, 3, 4, 5) \) and \( \sigma = (1, 2, 3, 4) \) be \( A_n \)-invariant. Then

\[
A^{(1)} = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5}
\end{bmatrix}, \text{ and } A^{(2)} = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}
\end{bmatrix}.\]
since \( a_{55} = \frac{1}{5} \) and \( a_{41} = \frac{1}{4} \), we have \( i = 5, j = 4, x = 5, y = 4 \).

The condition (ii) is satisfied for \( k = 1 \). Then the vector

\((\pi, \sigma) = (1, 2, 3, 4, 5, 1, 2, 3, 4)\) is \( A_n \)-invariant with \( A \) given by

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\
\frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

We can extend this example to any two vectors \( \pi^{(1)} \) and \( \pi^{(2)} \),

where the \( n_i \)th entry of \( \pi^{(i)} \) is \( n_i \) and the \( n_i \)th row of \( A^{(i)} \)

consists of equal entries \( \frac{1}{n_i} \), \( i = 1, 2 \). Therefore, this example

applies to all vectors of section 4.4.
4.7. PERMUTATION RESULTS

Many of the matrices constructed in the preceding sections are not in $P_n$ (for example, see section 4.5), i.e. they are not permutation invariant. It is of interest, however, to investigate the extent to which these vectors can be permuted.

Let $A \in A_n$ and let $P$ be a permutation matrix. It is known that post-multiplication by $P$ (i.e., $PA$) rearranges the rows of $A$ and so post-multiplication does not have any effect on the contiguity. On the other hand, pre-multiplication by $P^\top$, rearranges the columns of $A$ and this may destroy the contiguity. We also know that, given any $A \in A_n$, $PAP^\top$ is the matrix which corresponds to a relabelling of the vertices of $g(A)$.

The following two results show that certain permutations are admissible.

**Theorem 4.7.1:** $\pi = (\pi_1, \ldots, \pi_n)$ is $A_n$-invariant if and only if $\pi' = (\pi_n, \pi_{n-1}, \ldots, \pi_1)$ is, as well.

**Proof:** Consider the skew-diagonal matrix:

$$P = \begin{bmatrix}
0 & 1 \\
1 & 0 \\
1 & 1
\end{bmatrix}$$

and assume that $\pi$ is $A_n$-invariant with $A$ the matrix induced by $\pi$, i.e., $\pi A = \pi$. Then $(PA)P^\top$ just reverses the order of the,
columns of $PA$. Thus the contiguity of $PA$ still stands and therefore, $PAP^T \in A_n$. Note that $\pi' = \pi P^T$ and so, we have,

$$\pi' PAP^T = \pi P^T PAP^T = \pi AP^T = \pi P^T = \pi'. $$

Hence, $\pi'$ is $A_n$-invariant. Similarly, we can show that $\pi$ is $A_n$-invariant if $\pi'$ is.

Q.E.D.

**Theorem 4.7.2**: Suppose that $nA = \pi'$, where $A \in A_n$ is an $n \times n$ matrix which can be written in block form as $(A_1, A_2, A_3)$ and $A_1$ is an $n \times p_1$ matrix with $0 \leq p_1 \leq n$,

$$\sum_{i=1}^{3} p_i = n.$$ Let $A_2$ have the property that all its rows contain either 1 or $p_2$ non-zero elements. Then,

$$\pi' = (\pi_1, \ldots, \pi_{p_1}, \sigma_1, \ldots, \sigma_{p_2}, \pi_{p_1+p_2+1}, \ldots, \pi_n)$$

is $A_n$-invariant, where $(\sigma_1, \ldots, \sigma_{p_2})$ is a permutation of $(\pi_{p_1+1}, \ldots, \pi_{p_2})$.

**Proof**: Let $P_2$ be a $p_1 \times p_2$ permutation matrix such that,

$$(\pi_{p_1+1}, \ldots, \pi_{p_2})P_2^T = (\sigma_1, \ldots, \sigma_{p_2}).$$
Let us define

\[ p = \begin{bmatrix} I_1 & 0 \\ p_2 & I_3 \end{bmatrix} \]

where \( I_i \) is the \( p_i \times p_i \) identity matrix \( i = 1, 3 \). Then,

\[ PAP^T \in A_n, \quad \pi P^T = \pi' \quad \text{and} \]

\[ \pi' PAP^T = \pi P^T PAP^T = \pi AP^T \pi P^T = \pi', \]

i.e., \( \pi' \) is \( A_n \)-invariant.

Q.E.D.
CHAPTER V

MARKOV MAPS AND IRREDUCIBLE
(OR PRIMITIVE) MATRICES.

5.1. PRELIMINARIES

As we have seen, the irreducibility or the primitiveness of a non-negative matrix has been essential to our work. Therefore, it is interesting to know if a given matrix is irreducible (primitive) or not. Direct verification for large matrices (i.e., using the definition of an irreducible (primitive) matrix) can be very time-consuming, even on fast computers. For that reason, we would like to relate the irreducibility (primitiveness) of large matrices to the irreducibility (primitiveness) of smaller ones. In this chapter we shall analyze this idea using Markov maps.

Until now, we are familiar with Markov maps from [0,1] into itself, which are piecewise linear and belong to the class $T$. We showed (Theorem 2.1.2) that each $\tau \in T$, $\tau$ piecewise linear, determines uniquely a non-negative square matrix, which by Lemma 2.2.3 is irreducible. In this chapter we consider the following class of Markov maps:

**Definition 5.1.1**: Let $J = \{I_i\}_{i=1}^n$ be a partition of $I$ (I any interval), and let $\tau : I \to I$ be a Markov map with respect to $J$. We say that $\tau$ is in the class $W$ if:

(i) $\tau$ is piecewise $C^2$ with respect to $J$

(ii) $\inf_{x} |\tau'(x)| > 0$.

It follows that $\tau$ is piecewise monotonic with respect to $J$. 

(Lemma 2.1.2) and nonsingular. Obviously $\mathcal{T} \subseteq W$.

In section 5.2 we define the transition matrix, $B_\tau$, for each Markov map and we study the relations between $M_\tau$ and $B_\tau$ for $\tau \in \mathcal{T}$, $\tau$ piecewise linear.

Finally in section 5.3, we consider the following problem:

Let $J_1 = \{I_1\}_{i=1}^{n_1}$ be a Markov partition with respect to $\tau$ of the interval $I = [a, b]$, let $J_2 = \{I_2\}_{i=1}^{n_2}$ be any other partition of $I$ (assume $n_1 < n_2$). Let $g(B_\tau^{(1)})$ and $g(B_\tau^{(2)})$ be the 0-1 matrices (where $B_\tau^{(1)}$ and $B_\tau^{(2)}$ are the transition matrices and $g$ the function defined in 4.3) induced by $J_1$ and $J_2$, respectively.

Suppose that $g(B_\tau^{(1)})$ is irreducible (primitive). Under what conditions on $\tau$ will $g(B_\tau^{(2)})$ be irreducible (primitive)? We shall show that if $\tau$ satisfies a certain easily verifiable condition, then the above question is answered in the affirmative [7], in other words, that the irreducibility (primitivity) of $g(B_\tau^{(1)})$ implies the irreducibility (primitivity) of $g(B_\tau^{(2)})$. 
5.2. **Matrix Representation of a Markov Map**

Let us consider an interval \( I = [a, b] \) of the real line, and let us take a finite partition \( J = \{I_i\}_{i=1}^n \) of \( I \). Let \( \tau : I \to I \) be a Markov map with respect to the partition \( J \). We define the transition matrix \( B = B_\tau \) for the Markov map \( \tau \) as follows:

\[
   b_{ij} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)} \quad i, j = 1, 2, \ldots, n,
\]

where \( m \) is the Lebesgue measure on \( I \). Clearly \( b_{ij} \) is the proportion of the interval \( I_i \) which is mapped onto the interval \( I_j \).

It is known (Theorem 2.1.2) that when \( \tau \) is piecewise linear and belongs to the class \( \mathcal{T} \), then \( \tau \) induces a matrix \( M = M_\tau \), which is defined as follows:

\[
   m_{ij} = \begin{cases} 
   \frac{m(I_i)}{m(\tau^{-1}(I_i))} & \text{if } I_j \subset \tau^{-1}(I_i) \\
   0 & \text{otherwise}
   \end{cases}
\]

Thus, for a piecewise linear transformation \( \tau \in \mathcal{T} \) we can construct two matrices, namely \( B = B_\tau \) and \( M = M_\tau \). In the following we shall study the relations between \( B_\tau \) and \( M_\tau \).

In general \( B_\tau \) and \( M_\tau \) are not equal. We can see that noticing that all the non-zero entries of \( M_\tau \) on the \( i \)th row are contiguous and equal to \( \frac{m(I_i)}{m(\tau^{-1}(I_i))} \), but the non-zero entries of \( B_\tau \) on the same \( i \)th row are not necessarily equal since \( I_i \cap \tau^{-1}(I_j) \) varies, and so \( m(I_i \cap \tau^{-1}(I_j)) \), as \( I_j \) varies. We can see that easily from the following example:
Example 5.1: Let \( \tau: [0,1] \rightarrow [0,1] \) be a piecewise linear transformation defined by \( \tau(0) = 0, \tau\left(\frac{1}{4}\right)^- = \frac{1}{2}, \tau\left(\frac{1}{4}\right)^+ = 0, \tau\left(\frac{1}{2}\right)^- = 1, \tau\left(\frac{1}{2}\right)^+ = 0, \tau(1) = 1 \), and shown in Fig. 5.1.

\[ \begin{align*}
\text{Fig. 5.1}
\end{align*} \]

It is easy to see that

\[
M_\tau = \begin{bmatrix}
0.5 & 0.5 & 0 \\
0.25 & 0.25 & 0.25 \\
0.5 & 0.5 & 0.5
\end{bmatrix}
\]

and

\[
B_\tau = \begin{bmatrix}
0.5 & 0.5 & 0 \\
0.25 & 0.25 & 0.5 \\
0.25 & 0.25 & 0.5
\end{bmatrix}
\]

Obviously \( M_\tau \neq B_\tau \).

With the following Lemma we shall show that under certain conditions \( M_\tau = B_\tau \).
Lemma 5.2.1: Let $\tau \in \mathcal{T}$ be a piecewise linear transformation of $I$ onto $I$, $I = [a, b]$. The transition matrix $B = B_\tau$ and the induced matrix $M = M_\tau$ are precisely equal when the partition $J$ of $I$ has all subintervals of equal length.

Proof: We have from definitions of $B_\tau$ and $M_\tau$ that

$$b_{ij} = \frac{m(I_i \cap \tau_1^{-1}(I_j))}{m(I_i)}, \quad i, j = 1, 2, ..., n$$

and

$$m_{ij} = \begin{cases} \frac{m(I_i)}{m(\tau_1^{-1}(I_i))}, & \text{if } I_j \subseteq \tau_1(I_i) \\ 0, & \text{otherwise} \end{cases}$$

We consider two cases.

Case 1: Let $m_{ij} = 0$. Then $I_j \not\subseteq \tau_1(I_i)$, which implies that $I_j \cap \tau_1^{-1}(I_i) = \emptyset$ or $\tau_1^{-1}(I_j) \cap I_i = \emptyset$. Thus, $m(\tau_1^{-1}(I_j) \cap I_i) = m(\emptyset) = 0$. Therefore, $b_{ij} = 0$. Conversely, let $b_{ij} = 0$. Then $m(I_i \cap \tau_1^{-1}(I_j)) = 0$. This implies that $I_i \cap \tau_1^{-1}(I_j) = \emptyset$, since the intersection of two open intervals of the real line is $\emptyset$ or an open interval, and the measure of an open interval cannot be zero; equivalently, $\tau(I_i) \cap I_j = \emptyset$. Thus, $I_j \not\subseteq \tau(I_i)$, and so $m_{ij} = 0$. Therefore, the matrices $B_\tau$ and $M_\tau$, both have the same zero entries.

\footnote{All the lemmas in that section are due to the author.}
Case 2: Let $i$ be arbitrary, and $m_{ij} > 0$ for some $j, j = 1, 2, \ldots, n$. Let the length of each subinterval be denoted by $\Delta$, i.e., $m(I_i) = \ell(I_i) = a_i - a_{i-1} = \Delta$. \((m(I_i) = \ell(I_i))\) since $m$ is the Lebesgue measure. Suppose the Markov map $\tau$ maps $I_i$ onto $I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$. Then:

$$\tau_i(I_i) = I_t \cup I_{t+1} \cup \ldots \cup I_{t+k},$$

and $m(\tau_i(I_i)) = (k+1)\Delta$. Thus we get that

$$m_{ij} = \frac{m(I_i)}{m(\tau_i(I_i))} = \frac{\Delta}{(k+1)\Delta} = \frac{1}{k+1}, \quad j = t, t+1, \ldots, t+k. \quad (5.1)$$

Now from $\tau_i(I_i) = I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}$, we get:

$$I_i = \tau^{-1}(I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}).$$

Let $j = t + \ell$, $0 \leq \ell \leq k$. Then $\tau^{-1}(I_j) = \tau^{-1}(I_{t+\ell})$, and

$$I_i \cap \tau^{-1}(I_j) = \tau^{-1}(I_t \cup I_{t+1} \cup \ldots \cup I_{t+k}) \cap \tau^{-1}(I_{t+\ell}) = \tau^{-1}(I_{t+\ell}).$$

Thus,

$$m(I_i \cap \tau^{-1}(I_j)) = m(\tau^{-1}(I_{t+\ell})) = \frac{m(I_{t+\ell})}{|\tau_i|} = \frac{\Delta}{k+1}.$$

Therefore,
\[ b_{ij} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_j)} = \frac{\Delta}{k+1} = \frac{1}{k+1}, \quad j=t, t+1, \ldots, t+k, \quad (5.2) \]

and from (5.1) and (5.2) we get that \( m_{ij} = b_{ij} \) for \( i \) arbitrary and \( j=t, t+1, \ldots, t+k \). This completes the proof.

Q.E.D.

Example 5.2: Let us consider the Example 2.1. We have

\[
M_\tau = \begin{bmatrix}
0 & 0 & .5 & .5 \\
.25 & .25 & .25 & .25 \\
0 & .5 & .5 & 0 \\
.25 & .25 & .25 & .25
\end{bmatrix}
\]

Using the Definition of \( B_\tau \), we find that

\[
B_\tau = \begin{bmatrix}
0 & 0 & .5 & .5 \\
.25 & .25 & .25 & .25 \\
0 & .5 & .5 & 0 \\
.25 & .25 & .25 & .25
\end{bmatrix}
\]

Therefore \( M_\tau = B_\tau \).

Remark: From the proof of Lemma 5.2.1 it is clear that \( M_\tau \) and \( B_\tau \) both have the same zero entries not only in the case where we consider partition \( J \) with all subintervals of equal length, but also in the case where \( J \) is any arbitrary Markov partition. Thus \( M_\tau \) and \( B_\tau \) both have the same non-zero entries. Therefore, we can state the following lemma.
Lemma 5.2.2: Let \( \tau \in \mathcal{T} \), \( \tau \) piecewise linear, and let \( g(B) \) and \( g(M) \) be the 0-1 matrices induced (see section 1.3) by \( B_{\tau} \) and \( M_{\tau} \), respectively. Then \( g(B) = g(M) \).

**Proof:** We must show that \( g(B) \) and \( g(M) \) both have the same non-zero entries. Let us denote by \( C \) and \( T \) the matrices \( g(B) \) and \( g(M) \) respectively.

Let \( c_{ij} = 1 \), for some \( i,j, 1 \leq i,j \leq n \), then \( b_{ij} \neq 0 \) or \( m(I_j \land \tau^{-1}(I_j)) \neq 0 \), which implies that \( I_j \land \tau^{-1}(I_j) \neq \emptyset \) or \( I_j \land \tau(I_j) \neq \emptyset \). This implies that \( I_j \subset \tau(I_j) \) or \( I_j \supset \tau^{-1}(I_j) \), since any other relation between \( I_j \) and \( \tau(I_j) \) contradicts the fact that \( \tau \) takes partition points to partition points, i.e., \( I_j \subset \tau^{-1}(I_j) \). Thus \( t_{ij} = 1 \). Therefore, \( C = T \) or \( g(B) = g(M) \).

Q.E.D.

It is clear that the matrix \( B \) is irreducible (primitive) if and only if the matrix \( g(B) \) is irreducible (primitive). Thus we can prove the following result.

**Lemma 5.2.3:** The transition matrix \( B_{\tau} \), of a piecewise linear transformation \( \tau, \tau \in \mathcal{T} \), is irreducible matrix.

**Proof:** To show that \( B_{\tau} \) is irreducible we shall show that \( g(B_{\tau}) \) is. From Lemma 5.2.2 we have that \( g(B_{\tau}) = g(M_{\tau}) \). Therefore, if we show that \( g(M_{\tau}) \) is irreducible the result is proved. But \( g(M_{\tau}) \) is irreducible since \( M_{\tau} \) is in view the Lemma 2.2.3. Hence \( B_{\tau} \) is irreducible.

Q.E.D.
5.3. IRREDUCIBILITY AND PRIMITIVITY OF TRANSITION MATRICES

Let us consider a Markov map $\tau \in W$, $\tau : I \to I$ with respect to
the partition $J_1 = \{I_i\}_{i=1}^n$, and let $J_2 = \{K_j\}_{j=1}^m$ be any other
partition of $I$ (not necessarily Markov with respect to $\tau$). We say
that the condition (A) is satisfied if for each $K_j \in J_2$ there
exist $m_j$ and some $i, 1 \leq i \leq n$ such that:

$$\tau^{m_j}(K_j) \supset I_i \tag{A}$$

Theorem 5.3.1: Let $J_1$ be a Markov partition of $I$ with respect
to $\tau$, $\tau \in W$, and let $J_2$ be any other partition of $I$. Let $g(B^{(1)}_\tau)$
and $g(B^{(2)}_\tau)$ be the 0-1 matrices induced by $J_1$ and $J_2$,
respectively. If condition (A) is satisfied, then,

(i) $g(B^{(1)}_\tau)$ irreducible implies that $g(B^{(2)}_\tau)$ is irreducible,

and

(ii) $g(B^{(1)}_\tau)$ primitive implies that $g(B^{(2)}_\tau)$ is primitive.

Proof: Let $E = g(B^{(1)}_\tau)$ and $D = g(B^{(2)}_\tau)$.

(i) Let $E$ be irreducible. Then, by definition (2.1.4) for
each pair $(i,j)$ there exists $t$ such that $e_{ij}^t > 0$. We can see that
$e_{ij} > 0$, if and only if $m(I_i \cap \tau^{-1}(I_j)) > 0$, or if and only if
$I_i \cap \tau^{-1}(I_j)$ contains an open interval, or if and only if
$\tau(I_i) \cap I_j$ contains an open interval. Similarly, $e_{ij}^t > 0$ if and only if
$m(I_i \cap \tau^{-t}(I_j)) > 0$ and thus, if and only if $\tau^t(I_i) \cap I_j$ contains an
open interval.

So it suffices to show that for all $i,j$ there exists an $m$ such that
$\tau^m(K_i) \cap K_j$ contains an open interval. By condition (A), there exists an $m_i$
and $p$, $1 \leq p \leq n$ such that $\tau^{m_i}(K_i) \supset I_p$. Now, $K_j \cap I_p \neq \emptyset$ for some $\ell$. By the
irreducibility of \( E \), there is an \( m' \) such that \( \tau^{m'}(I_p) \supset I_\infty \). Then
\[
\tau^{m'+m} (K_i) \supset \tau^{m'}(I_p) \supset I_\infty \supset K_j \cap I_\infty \neq \emptyset.
\]
Therefore, \( \tau^{m'+m}(K_i) \cap K_j = \tau^{m}(K_i) \cap K_j \neq \emptyset \), i.e., contains an open interval.

\[(if)\] Let \( E \) be primitive. Then exists a \( k \) such that \( e_{i,j}^k > 0 \) for all \( i,j \), which means that \( m(I_i \cap \tau^{-k}(I_j)) > 0 \), or that \( I_i \cap \tau^{-k}(I_j) \) contains an open interval for each \( i \) and \( j \), which means that \( \tau^k(I_i) = I \) for each \( i \), the same is true for \( k+1 \), i.e., \( \tau^{k+1}(I_i) = I \) for each \( i \). Let \( m' = m + m_j \). Then by condition (A), we have,
\[
\tau^{m_j}(K_j) \supset I_i,
\]
or
\[
\tau^{m}(\tau^{m_j}(K_j)) \supset \tau^{m}(I_i)
\]
or
\[
\tau^{m}(K_j) \supset I = \tau^{m}(I_i), \quad (5.3)
\]
and since,
\[
\tau^{m}(K_j) \subset I \quad (5.4)
\]
for each \( K_j, j = 1, 2, \ldots, n \),
\[
\tau^{m'}(K_j) = I
\]
Moreover,
\[
\tau^{m'+1}(K_j) = I
\]
It follows that \( D^{m'} \) and \( D^{m'+1} \) both have non-zero entries in the \( j \)th position on the diagonal. Hence \( D \) is primitive.

Q.E.D.
Let \( J_1 = \{ I_q \}_{q=1}^n \) be a fixed partition of \( I = [0,1] \) and let \( \tau \in \mathcal{W} \). We define

\[
\mu = \min_{1 \leq q \leq n} m(I_q) \quad \text{and} \quad \sigma = \max_{1 \leq q \leq n} m(I_q)
\]

Let \( c = \mu / \sigma \). Clearly \( 0 < c \leq 1 \). When the partition consists of equal intervals, \( c = 1 \).

Let \( \tau_i = \tau| I_i \) and let \( \ell_i \) be the straight line joining the endpoints of \( \tau_i(x) \). Define,

\[
\beta_i = \inf_{x \in I_i} |\tau_i(x)|
\]

and

\[
\delta_i = \frac{\beta_i}{\ell_i}
\]

Obviously, \( 0 < \delta_i \leq 1 \). If \( \tau_i \) is linear, then \( \delta_i = 1 \).

**Lemma 5.3.1:** Let \( K \) be an interval contained in \( I_p \), where

\[ \tau(I_p) = \bigcup_{j=1}^q I_{p'+j} \]

for some \( p' \) and \( q \). Then

\[
1 \geq \frac{m(\tau(K))}{\sum_{j=1}^q m(I_{p'+j})} \geq \delta_p \eta(K),
\]

where,

\[
\eta(K) = \frac{m(K)}{m(I_p)}
\]

**Proof:** The first inequality is obvious, since \( \tau(K) \subset \tau(I_p) \)

implies that \( m(\tau(K)) \leq m(\tau(I_p)) = m(\bigcup_{j=1}^q I_{p'+j}) = \sum_{j=1}^q m(I_{p'+j}) \).
To show the second inequality, let us consider the straight line \( \ell_p \) joining the endpoints of \( \tau(I_p) \). Then,

\[
\ell_p' = \frac{m(\tau(I_p))}{m(I_p)} = \frac{m(\tau(K))}{m(K)}.
\]

Since \( \tau \) is monotone, we have

\[
\ell_p' = \beta_k = \inf_{x \in K} |\tau'(x)| \geq \beta_p = \inf_{x \in I_p} |\tau'(x)|.
\]

Then

\[
m(\tau(K)) = \ell_p' m(K) \geq \beta_p m(K).
\]

(5.5)

Now, since

\[
\delta_p' = \beta_p = \frac{\beta_p}{\ell_p'} = \beta_p \frac{m(\tau(I_p))}{m(I_p)} = \frac{\beta_p m(I_p)}{m(\tau(I_p))},
\]

or

\[
\delta_p = \frac{\beta_p m(I_p)}{m(I_p)} = \frac{\beta_p m(I_p)}{m(\cup_{j=1}^{q} I_{p,j})} = \frac{\beta_p m(I_p)}{\sum_{j=1}^{q} m(I_{p,j})},
\]

we obtain

\[
\beta_p = \frac{\delta_p}{m(I_p)} \sum_{j=1}^{q} m(I_{p,j}).
\]

Thus, using (5.5), we get

\[
m(\tau(K)) \geq \frac{\delta_p}{m(I_p)} \sum_{j=1}^{q} m(I_{p,j}) m(K),
\]

and

\[
\frac{m(\tau(K))}{\sum_{j=1}^{q} m(I_{p,j})} \geq \frac{\delta_p}{m(I_p)} m(K) = \delta_p \eta(K).
\]

Q.E.D.
Remark: Since $K \subseteq I_p$, $m(K) \leq m(I_p)$. This means that $n(K) \leq 1$.

Corollary 5.3.1: If $\tau(I_p) = I_p$, then

$$n(\tau(K)) \geq \delta_p n(K).$$

Proof: $\tau(I_p) = I_p$ implies that $q = 0$. Then $m(\tau(I_p)) = m(I_p)$, and so

$$n(\tau(K)) = \frac{m(\tau(K))}{m(\tau(I_p))} = \frac{m(\tau(K))}{m(I_p)} \geq \delta_p n(K) \sum_{j=1}^{q} \frac{m(I_p') / m(I_p)}{m(I_p)} = \delta_p n(K).$$

Q.E.D.

Definition 5.3.1: A Markov partition $J$ is said to be a strong Markov partition if the transition matrix $B$, is irreducible and not a permutation matrix.

Theorem 5.3.2: Let $J_1 = (I_i)_{i=1}^n$ be a strong Markov partition with respect to $\tau, \tau \in \mathcal{W}$, and let $J_2 = (K_j)_{j=1}^m$ be any other partition of $I$. Then condition (A) is satisfied if

$$\delta_1 \delta_2 \ldots \delta_n (1+c) > 1.$$
Proof: We first define a sequence $K(1), K(2), \ldots, K(j), \ldots$ of intervals contained in the elements of $J_1$ such that,

$$K(i+1) = \tau(K(i)),$$

i.e.,

$$K(j) = \tau(K(j-1)) = \tau^2(K(j-2)) = \ldots = \tau^{j-1}(K(1)).$$

We shall show that the corresponding sequence $\{n(K(i))\}$ attains the value 1 in a finite number of steps, because if $n(K(i)) = 1$ for some $i$, then $m(K(i)) = m(I_i)$ or $K(i) = I_i$. And since $K(i) = \tau^{i-1}(K(1))$, we have

$$I_i = \tau^{i-1}(K(1)).$$

Therefore, condition (A) is satisfied.

Let $K \in J_2$ and let $I \in J_1$ be such that $K \cap I = \emptyset$. Let $K(1) = K \cap I$. Assuming that $K(1) \not\subset I_p$ (if $K(1) = I_p$ we are through), we define $K(i+1)$ as follows.

**Case 1:** $\tau: I_p \rightarrow I_p'$. Then $K(i+1) = \tau(K(i)) \subset I_p'$. Note that $\tau(K(i)) \not\subset \tau(I_p)$, since $|\tau(I_p)| > 0$. So $K(i+1) \not\subset I_p'$.

**Case 2:** $\tau: I_p \rightarrow \bigcup_{j=1}^{q'} I_p' + j$. There are three possibilities.

1. $\tau(K(1)) = I_p' + \ell$ for some $1 \leq \ell \leq q'$. Then $K(i+1) = \tau(K(i)) \subset I_p' + \ell$.

In this case, either $K(i+1) = I_p' + \ell$ and we terminate the sequence with $n(K(i+1)) = 1$ or $K(i+1) \not\subset I_p' + \ell$. 
(2) \( \tau(K^{(1)}_i) \subseteq I_{p_i}^t + \ell \cup I_{p_i}^t + \ell + 1 \) and \( \tau(K^{(1)}_i) \) intersects both intervals.

Let \( \Delta(k) = \tau(K^{(1)}_i) \cap I_{p_i}^t + \ell + k \), \( k = 0 \) or \( 1 \). If for \( k = 0 \) or \( 1 \), 
\[ \Delta(k) = I_{p_i}^t + \ell + k \] let \( k(i+1) = \Delta(k) \), and we terminate the 
sequence with \( n(k^{(i+1)}) = 1 \). (If this condition holds for both 
\( \Delta(0) \) and \( \Delta(1) \), we choose any one for \( k^{(i+1)} \).) Otherwise, let 
\( k(i+1) = \Delta(k) \) where \( n(\Delta(k)) = n(\Delta(k')) = k' = k \).

(3) \( \tau(K^{(1)}_i) \subseteq \bigcup_{s=1}^{t} I_{p_i}^t + \ell + s \) for some \( 1 < \ell < q \), \( 1 \leq t \leq q - \ell \). Then 
\( I_{p_i}^t + \ell \subseteq \tau(K^{(1)}_i) \), since \( \tau \) is piecewise continuous. Let 
\( k(i+1) = I_{p_i}^t + \ell \); then the sequence is terminated with 
\( n(k^{(i+1)}) = 1 \).

We now obtain, from Case 1 and 2 (1), (2), (3) a set of 
inequalities.

Case 1: By the Corollary of Lemma 5.3.1, we have 
\[ n(\tau(K^{(1)}_i)) = n(k^{(i+1)}) \geq \delta_p n(K^{(1)}_i) \] (5.6)
Case 2: (1) By Lemma 5.3.1, we have,

\[
\delta_p n(k^{(i)}) = \frac{m(\tau(k^{(i)}))}{\sum_{j=1}^q m(I_{p^i+\ell})^j} = \frac{m(k^{(i+1)})}{m(I_{p^i+\ell}) + \mu} \leq \frac{m(k^{(i+1)})}{m(I_{p^i+\ell})} \leq \frac{n(k^{(i+1)})}{1+c}.
\]

Thus,

\[
\delta_p (1+c) n(k^{(i)}) \leq n(k^{(i+1)}) \leq 1,
\]

where the last inequality is obtained from the remark of Lemma 5.3.1.

Case 2: (2) Since \( \tau(k^{(i)}) \subset \Delta(0) \cup \Delta(1) \), we have from Lemma 5.3.1,

\[
\delta_p n(k^{(i)}) = \frac{m(\Delta(0))}{\sum_{j=1}^q m(I_{p^i+\ell})} + \frac{m(\Delta(1))}{\sum_{j=1}^q m(I_{p^i+\ell})} \leq n(\Delta(0)) + n(\Delta(1)) \leq 2n(k^{(i+1)}).
\]

Thus,

\[
\frac{\delta_p n(k^{(i)})}{2} \leq n(k^{(i+1)}) \leq 1.
\]

It is clear from this construction that the sequence \( \{k^{(i)}\} \)

terminates only if there exists \( p \) such that \( K^p \in J_1 \). Then,
condition (A) is satisfied.

Now we shall show that the sequence $\{K(i)\}$ must terminate. First, we shall see that if the sequence is infinite, then Case 2 (2) cannot occur more than once. Suppose that it occurs twice, i.e., $K(i)$ and $K(j)$ are both derived by application of the Case 2 (2) construction. Then,

$$\tau(K(i-1)) = (a, X_p, b) \cup (X_p, +\xi, b),$$

where $X_p, +\xi$ is a partition point of $J_1$. Then, by the definition of $K(i)$, it is equal to $\Delta(p)$ or $\Delta(1)$, so that $X_p, +\xi$ is an end point of $K(i)$. Now, $\tau$ maps partition points to partition points, and since $\tau$ is monotonic on each subinterval it is easy to see that $K(j)$ must be an interval with one of its endpoints a partition point of $J_1$, say $X_p$. Then, by our assumption,

$$\tau(K(j-1)) = (a', x_p) \cup (x_p, b'),$$

where $x_p$ is mapped into either $a'$ or $b'$. Since $\tau$ is Markov, $a'$ or $b'$ is a partition point. Thus $K(j) \in J_1$ and the sequence terminates. Contradiction.

Now, let $s$ be the index such that $K(s)$ is obtained from $K(s-1)$ by applying the construction of Case 2 (2), (if this case is impossible we put $s = 1$) and let us consider the sequence of numbers,

$$\eta(K(s)), \eta(K(s+1)), \ldots.$$ We note that if this sequence did not terminate, we would have $\eta(K(s+i)) < 1$ for all $i \geq 0$.

Now it is clear that for all $p \geq s$, $K(p+1)$ can only be derived from the construction of Case 1 or Case 2 (1). For any such $p$, let $k$ be the smallest index such that $K(p+k)$ is derived by applying
Case 2 (1). Let $k^{(p+j)} = I_{i,j} \in J_1$ for all $0 \leq j \leq k-1$. There can be no repetitions among the sequence $\{I_{i,j}\}_{j=0}^{k-1}$ since the matrix $g(B_1^{(1)})$ is irreducible and hence, the elements of $J_1$ are aperiodic. Therefore, we must have $k \leq n$. If $k = n$, then $B_1^{(1)}$ would be a permutation matrix, contradicting the assumption. Thus $k \neq n$.

Therefore, by inequality (5.6), we have,

$$n(k^{(p+k-1)}) \geq \delta_{i_k} \delta_{i_{k-1}} \cdots \delta_{i_0} n(k^{(p+1)}) .$$

Since $n(k^{(p+1)}) \geq \delta_{i_k} (1+c) n(k^{(p)})$ (by 2(1)) we obtain,

$$n(k^{(p+k-1)}) \geq \delta_{i_k} \delta_{i_{k-2}} \cdots \delta_{i_0} \delta_{i_1} n(k^{(p+1)}) .$$

and, since, $n(k^{(p+k)}) \geq \delta_{i_k} n(k^{(p+k-1)})$ (by (5.6), we finally get).

$$n(k^{(p+k)}) \geq \delta_{i_k} \delta_{i_{k-1}} \delta_{i_0} \delta_{i_1} (1+c) n(k^{(p)}) ,$$

or

$$n(k^{(p+k)}) \geq n_k (1+c) n(k^{(p)}) ,$$

since the $\delta_{i_j}$ are all distinct and $\delta_{i_j} \leq 1$.

Consider now the subsequence $(n(k^{(s)}), n(k^{(s+k_1)}), n(k^{(s+k_1+k_2)}), \ldots, n(k^{(s+\frac{g}{k} k_1)}), \ldots, n(k^{(s+k_1+k_2)}), \ldots, n(k^{(s+\frac{g}{k} k_1)})$, where $k_g$ is the index defined above, corresponding to $p = s + \sum_{i=1}^{g} k_i$. Then,
\[ n^k_{(s+\sum_{i=1}^{g} k_i)} \geq \delta_1 \delta_2 \cdots \delta_{n(1+c)} n(k(s)) > 1, \]

for \( g \) sufficiently large, since \( \delta_1 \delta_2 \cdots \delta_{n(1+c)} > 1 \) by hypothesis.

This contradicts the statement that \( n(k(s+i)) < 1 \), for all \( i \geq 0 \).

Q.E.D.

**Remark 1:** We had assumed that the transition matrix \( B^{(1)}_t \) is not a permutation matrix. This is a necessary condition as we shall see from the following example.

**Example 5.3:** Let \( \tau : [0,3] \rightarrow [0,3] \) be the piecewise linear map defined by \( \tau(0) = 1, \tau(2^-) = 3, \tau(2^+) = 0, \tau(3) = 1 \); and shown in Fig. 5.2.

![Fig. 5.2](image-url)
Then \( J^1 = \{(0,1), (1,2), (2',3)\} \), and

\[
g(B_1^{(1)}) = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 
\end{bmatrix}
\]

\( B_1^{(1)} \) is a permutation matrix and irreducible.

Consider now the partition

\( J_2 = \{(0.5), (\frac{1}{2},1), (1,1.5), (1.5,2), (2,2.5), (2.5,3)\} \).

The 0-1 matrix induced by \( J_2 \) i.e., the \( g(B_1^{(2)}) \) is,

\[
g(B_1^{(2)}) = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 
\end{bmatrix}
\]

But \( g(B_1^{(2)}) \) is not an irreducible matrix as we can easily see from the associated directed graph. It consists of two disjoint cycles

\( 1 \rightarrow 3 \rightarrow 5 \rightarrow 1 \) and \( 2 \rightarrow 4 \rightarrow 6 \rightarrow 2 \), which is not strongly connected.

**Remark 2:** If \( \tau \) is piecewise linear, then \( \delta_i = 1, \ i = 1, \ldots, n \), and in that case, the sufficiency condition in Theorem 5.3.2 reduces to \( (1+c) > 1 \), which is always true. Therefore, Theorem 5.3.2 is true when \( \tau \) is piecewise linear.

If we combine Theorems 5.3.1 and 5.3.2, we obtain the following main result.
Theorem 5.3.3: Let $J_1$ be a strong Markov partition of $I$
with respect to $\tau, \tau \in W$, satisfying the following condition,

$$\delta_1 \delta_2 \cdots \delta_{n(1+c)} > 1.$$ 

Furthermore, let $J_2$ be any other partition of $I$. Let $g(B^{(1)}_\tau)$ and
$g(B^{(2)}_\tau)$ be the 0-1 matrices induced by $J_1$ and $J_2$, respectively.
Then $g(B^{(2)}_\tau)$ is irreducible. If $g(B^{(1)}_\tau)$ is also primitive, then
so is $g(B^{(2)}_\tau)$.

Proof: This follows from Theorems 5.3.1 and 5.3.2. Q.E.D.

Theorem 5.3.3 may not hold if $J_1$ is not Markov. To prove this,
consider the following example.

Example 5.4: Consider the piecewise linear map $\tau: [0,1] \rightarrow [0,1]$
defined by the conditions.

$$\tau(0) = 0, \tau(.5^-) = 1, \tau(.5^+) = .9 \text{ and } \tau(1) = .4$$

and, shown in Fig. 5.3,

![Fig. 5.3](image-url)
Then \( J_1 = \{(0, 0.5), (0.5, 1)\} \), and,
\[
g(B_\tau^{(1)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\]

Clearly, \( J_1 \) is not a Markov partition, but the induced matrix \( g(B_\tau^{(1)}) \) is, obviously, primitive.

Now, consider the partition \( J_2 = \{(0, 0.4), (0.4, 0.5), (0.5, 0.9), (0.9, 1)\} \).
The matrix \( g(B_\tau^{(2)}) \), induced by \( J_2 \) is
\[
g(B_\tau^{(2)}) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}
\]

Now, we consider the matrix \( C = g(B_\tau^{(2)}) + g(B_\tau^{(2)})^2 + g(B_\tau^{(2)})^3 \).
\[
C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}
+ \begin{bmatrix} 1 & 1 & 1 & 2 \\
0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix}
+ \begin{bmatrix} 1 & 1 & 1 & 3 \\
0 & 0 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 2 & 3 & 1 \end{bmatrix}
= \begin{bmatrix} 3 & 3 & 3 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 6 & 5 \\ 0 & 3 & 5 & 3 \end{bmatrix}
\]

Since the matrix \( C \) contains zero entries off the main diagonal, according to Corollary 4.2.1 the corresponding graph to \( g(B_\tau^{(2)}) \) is not strongly connected, and so \( B_\tau^{(2)} \) is not irreducible.
Example 5.5: Consider the continuous map $\tau: [0,1] \to [0,1]$ defined by

$$
\tau(x) = \begin{cases} 
2x + \frac{1}{2}, & 0 \leq x \leq \frac{1}{4} \\
-2x + \frac{3}{2}, & \frac{1}{4} \leq x \leq \frac{3}{4} \\
2x - \frac{3}{2}, & \frac{3}{4} \leq x \leq 1
\end{cases}
$$

and shown in Fig. 5.4.

Then $J_1 = \{(0, \frac{1}{4}), (\frac{1}{4}, \frac{3}{4}), (\frac{3}{4}, 1)\}$ and

$$
g(B_\tau^{(1)}) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}
$$

It is clear that $g(B_\tau^{(1)})$ is primitive $(g(B_\tau^{(1)})^2 > 0)$. Now, let $J_2 = \{(0, \frac{1}{4}), (\frac{1}{4}, \frac{1}{2}), (\frac{1}{2}, \frac{3}{4}), (\frac{3}{4}, 1)\}$ be a second partition, which induces the matrix
\[
g(B_\tau^{(2)}) = \begin{bmatrix}
0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
\end{bmatrix}
\]

\(g(B_\tau^{(2)})\) is not primitive (much less, irreducible). Clearly, \(J_1\) is not a Markov partition.

**Example 5.6:** Let \(\tau: [1,5] \rightarrow [1,5]\) be the piecewise linear map defined by,

\[
\tau(1) = 3, \tau(2) = 5, \tau(3) = 4, \tau(4) = 2 \text{ and } \tau(5) = 1,
\]

and shown in Fig. 5.5.
Then, \( J_1 = \{(1,2), (2,3), (3,4), (4,5)\} \). Let \( J_2 \) be another partition of \([1,5]\), given by

\[
J_2 = \{(1,1,4), (1,4,2), (2,2,4), (2,4,3), (3,3,8), (3,8,4), (4,4,6), (4,6,5)\}.
\]

\( J_1 \) is obviously a Markov partition, inducing the matrix

\[
B^{(1)} = \begin{bmatrix}
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & 0 & 0 & 0
\end{bmatrix}
\]

The matrix \( g(B^{(1)}) \) is primitive, since by Theorem 4.2.2 we have

\[
C = g(B^{(1)}) + g(B^{(1)})^2 + g(B^{(1)})^3 = \begin{bmatrix}
1 & 2 & 4 & 3 \\
2 & 0 & 1 & 2 \\
1 & 3 & 3 & 2 \\
2 & 1 & 2 & 2
\end{bmatrix}
\]

and \( g(B^{(1)})_{33} = 1 > 0 \). Hence the matrix

\[
g(B^{(2)}) = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
induced by \( J_2 \), is primitive, by Remark 2 in section 5.3.

**Example 5.7:** Let \( \tau \) and \( J_1 \) be defined as they were in Example 5.6. Let \( n \) be divisible by 8, and let \( J_2 \) be an equal length partition of \([1,5]\) consisting of \( n \) subintervals. Then the matrix induced by \( J_2 \) is

\[
g(B^{(2)}_\tau) = \begin{bmatrix}
0 & 0 & A_n & B_n \\
0 & 0 & 0 & I_n \\
0 & C_n & D_n & 0 \\
I_n & 0 & 0 & 0
\end{bmatrix}
\]

where each block is an \( \frac{n}{4} \times \frac{n}{4} \) matrix and the non-zero blocks are defined as follows.

Let \( H_n \) and \( G_n \) be \( \frac{n}{8} \times \frac{n}{8} \) band matrices given by

\[
H_n = \begin{bmatrix}
1 & 1 & 0 \\
1 & 1 & 0 \\
0 & 1 & 1
\end{bmatrix}, \quad G_n = \begin{bmatrix}
0 & 1 & 1 \\
0 & 1 & 1
\end{bmatrix}
\]

Define, \( A_n = \left( \begin{array}{c} H_n \\ 0 \end{array} \right) \), \( B_n = \left( \begin{array}{c} 0 \\ R_n \end{array} \right) \), \( C_n = \left( \begin{array}{c} G_n \\ 0 \end{array} \right) \) and \( D_n = \left( \begin{array}{c} 0 \\ 0 \end{array} \right) \). \( I_n \) is the \( \frac{n}{4} \times \frac{n}{4} \) skew diagonal matrix. Since \( g(B^{(1)}_\tau) \) is primitive, so is \( g(B^{(2)}_\tau) \).
Example 5.8: Consider the map \( \tau: [0,1] \rightarrow [0,1] \), defined as follows:

\[
\tau(x) = \begin{cases} 
2x & 0 \leq x \leq \frac{1}{2} \\
-2x + 2 & \frac{1}{2} \leq x \leq 1 
\end{cases}
\]

Let \( J_1 = \{(0, \frac{1}{2}), (\frac{1}{2}, 1)\} \). Obviously, \( J_1 \) is a Markov partition which induces the primitive matrix

\[
g(B_{\tau}^{(1)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\]

Now, let us consider the partition

\[
J_2 = \left\{ \left( \frac{1}{n+1}, \frac{1}{n} \right), \left( \frac{1}{3}, \frac{1}{2} \right), \left( \frac{1}{2}, \frac{2}{3} \right), \left( \frac{2}{3}, \frac{3}{4} \right), \ldots, \left( 1 - \frac{1}{n}, 1 - \frac{1}{n+1} \right) \right\},
\]

where \( n+1 \) is even. \( J_2 \) is not a Markov partition (because \( \frac{2}{7} \) is not a partition point). The \( 2n \times 2n \) matrix induced by \( J_2 \) is

\[
g(B_{\tau}^{(2)}) = \begin{bmatrix} A_n & B_n \\ \overline{A}_n & \overline{B}_n \end{bmatrix}
\]

where \( A_n, B_n, \overline{A}_n, \overline{B}_n \) are \( n \times n \) matrices, defined as follows:
where, in $A_n$, the first row consists of $\frac{n+3}{2}$ 1's followed by 0's.

Also, $\overline{A}_n$, $\overline{B}_n$ are obtained from $A_n$, $B_n$, respectively, by interchanging the rows, i.e., the first row becomes the last, the second row, the second of last; and so on. By Theorem 5.3.3, $g(B_t^{(2)})$ is primitive.

Remark: As simple as the transformation of this example is, it can generate by appropriate partition choices, large classes of matrices. All these matrices are primitive.
REFERENCES


