SYSTEM THEORY FOR TWO POINT BOUNDARY VALUE DESCRIPTOR SYSTEMS

by

RAMINE NIKOUKHAH

SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING
AND
COMPUTER SCIENCE
IN PARTIAL FULFILLMENTS OF THE REQUIREMENTS
FOR THE DEGREES OF

ELECTRICAL ENGINEER

and

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1986

© Massachusetts Institute of Technology 1986

Signature of Author _____
Department of Electrical and Computer Science
May 16, 1986

Certified by _____
Alan S. Willsky
Thesis Supervisor

Certified by _____
Bernard C. Levy
Thesis Supervisor

Accepted by _____
Arthur C. Smith
Chairman, Department Committee

SYSTEM THEORY FOR TWO POINT BOUNDARY VALUE DESCRIPTOR SYSTEMS

by

RAMINE NIKOUKHAH

Submitted to the Department of Electrical Engineering and Computer Science on May 16, 1986, in partial fulfillment of the requirements for the Degrees of Electrical Engineer and Master of Science in Electrical Engineering.

ABSTRACT

The two point boundary value descriptor system (TPBVDS)

$$Ex_{k+1} = Ax_k + Bu_k$$
$$V^i x_0 + V^f x_N = v$$
$$y_k = Cx_k$$

where E and A can be singular, is the natural internal model for discrete-time noncausal linear systems. In this thesis, we develop a deterministic and a stochastic system theory for this class of systems. This theory is closely related to the work of Krener for continuous-time, non-descriptor boundary value linear systems, but it must also deal with the possible singularity of E and A. In particular, in the deterministic case, we investigate concepts of reachability, observability, minimality, stationarity, and stability. In the stochastic case, we investigate the concept of stochastic stationarity and relate it to stability by studying the properties of a generalized Lyapunov equation.

THESIS SUPERVISORS:

Professor Alan S. Willsky
Title: Professor of Electrical Engineering

Professor Bernard C. Levy
Title: Associate Professor of Electrical Engineering

Acknowledgements

The research and the preparation of this thesis have been conducted under continuous guidance of my thesis supervisors. I wish to express my thanks to Professor Alan Willsky whose ideas motivated this work and whose insight, guidance and encouragement have been an invaluable resource to me. I would also like to thank Professor Bernard Levy for his accessibility, for his suggestions and for the many technical insights he has provided.

Most of all, I would like to thank my parents who supported me through many years of my education and were always there when I needed them.

This research was carried out at the MIT Laboratory for Information and Decision Systems and was supported, in part, by National Science Foundation grant #ECS-83-12921 and Air Force Office of Scientific Research grant #AFOSR-82-0258.

TABLE OF CONTENTS

CHAPTER I: Introduction..... 9

CHAPTER II: Background..... 14
2.1- Descriptor Dynamics..... 14
2.1.1- Introduction..... 15
2.1.2- Well-Posedness..... 17
2.1.3- Reachability and Observability..... 18
2.2- Boundary Value Linear Systems..... 21

CHAPTER III: Two Point Boundary Value Descriptor Systems: The Deterministic Problem..... 26
3.1- Well-Posedness and Standard Form..... 26
3.2- Strong Reachability/Strong Observability..... 31
3.2.1- Inward and Outward Boundary Value Processes.... 31
3.2.2- Notation..... 33
3.2.3- Strong Reachability and Strong Observability Matrices and Spaces..... 35
3.3- Reachability and observability: General Case..... 42
3.4- Stationary TPBVDS..... 50
3.4.1- Reachability/Observability: Stationary Case.... 53
3.4.2- Minimality for Stationary TPBVDS..... 58
3.4.3- Stability for stationary TPBVDS..... 67

CHAPTER IV: The Stationary TPBVDS: The Stochastic Problem... 75
4.1- Introduction..... 75
4.2- Stochastically Stationary TPBVDS..... 79
4.3- Stability..... 95

CHAPTER V: Conclusion.....109
5.1- Contributions.....109
5.2- Suggestions for Further Research.....111

APPENDIX A: The Inward Boundary Value Process z'.....113
A.1- Introduction.....113
A.2- Recursive Method to Compute z'.....113
A.3- Reachable Space of z'.....116
A.4- Case of Invertible E and A.....122

APPENDIX B: Recursive Solutions for TPBVDS's.....124
B.1- The Two Filter Solution.....124
B.2- The Inward-Outward Solution.....127

I- Introduction

Most systems evolving in time are causal, in the sense that the state at present time is only a function of the inputs and the disturbances of the past. However, systems, where the independent variable is the space and not the time, are not in general causal. Consider, for example, a beam clamped at both ends that supports a distributed load. Clearly, the deflection of the beam at any point is a function of the load on both sides of that point. In this case, by writing out the differentiel or difference equation describing the dynamics of the deflection of the beam and the equation describing the boundary conditions, we obtain a non-causal system. Other examples of non-causal systems are cyclic systems. A system is called cyclic if its initial and final states are identical. For example, the system representing the temperature distribution in a ring is a cyclic system.

In general, non-causal systems are described by a dynamics equation and a boundary equation. In the case of the clamped beam, the boundary equation is of the form which we call separable which means that the boundary condition imposed at one end of the beam is independent of the boundary condition imposed at the other end of the beam. Non-causal systems, however, are not always separable, for

example, cyclic systems where the initial and final conditions are clearly not independent. We study a class of non-causal systems where the dynamics equation is linear and time-invariant first order descriptor difference equation and the boundary equation is a linear equation in the initial and final states. We shall call such systems two point boundary value descriptor systems (TPBVDS).

The two point boundary value descriptor system (TPBVDS)

$$Ex_{k+1}=Ax_k+Bu_k \tag{1.1.1}$$

$$V^i x_0+V^f x_N=v \tag{1.1.2}$$

$$y_k=Cx_k \tag{1.1.3}$$

where E, A, V^i and V^f are nxn (possibly singular) matrices is a generalization of the autonomous boundary value linear system (BVLS) introduced by Krener [1], and more specifically of its discrete-time equivalent. TPBVDS's arise naturally in the study of non-causal systems, for example, they arise in the study of two-dimensional nearest neighbour models [20].

Special Cases

System (1.1) has a very general form. It is useful to describe at this point some of its special cases.

(1) Linear Causal Systems

Linear causal systems are a special case of TPBVDS where $E=V^i=I$ and $V^f=0$. These systems have been extensively

studied in the past (see for example [23]).

(2) Non-Descriptor Two-Point Boundary Systems

In this case, E and A are assumed to be invertible. These systems are straight forward extensions of Krener's BVLS's to the discrete time.

(3) Cyclic Systems (Anticyclic Systems)

Cyclic systems (anticyclic systems) are obtained from TPBVDS by letting $V^i=-V^f=I$ ($V^i=V^f=I$).

(4) Descriptor (Singular) Systems [13-19]

Descriptor systems are a special class of TPBVDS where E is singular and V^f equals 0.

The objective of our work is to develop a system theory for the TPBVDS (1.1). Clearly, this theory should be a generalization of causal system theory and descriptor system theory.

Summary

In Chapter II, we review that part of descriptor system theory that is related to the study of TPBVDS. We also review Krener's study of boundary value linear systems which is used as guidance for some of our study of TPBVDS.

Chapter III contains the main body of our work. In Section 3.1 we introduce the concepts of well-posedness and standard form. We also derive the Green's function solution

for the TPBVDS. In Section 3.2, we present the idea of an inward boundary process and an outward boundary process and we define the concepts of strong reachability and strong observability. Then, using the concept of standard form, we derive simple expressions for the strong reachability and strong observability matrices and spaces. In Section 3.3, we introduce the weaker notions of reachability and observability and show that in general these notions differ from those of strong reachability and strong observability. We also derive reachability and observability matrices and spaces. In Section 3.4, we introduce the concept of stationarity. We show that reachability and observability matrices take simple forms in the stationary case. We also explore the properties of the reachability and the observability matrices. In Section 3.4.2, we explore the question of minimality and obtain a method to reduce any stationary TPBVDS to a minimal TPBVDS. Finally, in Section 3.4.3, we consider the problem of stability for stationary TPBVDS.

Chapter IV is devoted to the study of stationary TPBVDS driven by white noise. In Section 4.2, we introduce the concept of stochasic stationarity and derive necessary and sufficient conditions for a TPBVDS to be stochastically stationary. Later, we show how the covariance of the state of a TPBVDS can be computed using a generalized Lyapunov

equation and a simple recursion. Finally, we investigate the relationships between the notion of stability and the generalized Lyapunov equation.

Chapter V is the concluding chapter. There, we first present a list of contributions of our work, and then, suggestions for future research.

In Appendix A, we derive a recursive method for computing the inward boundary process introduced in Chapter III. We also present proofs for some of the results discussed in Chapter III. Finally, in Appendix B, we present algorithms to solve TPBVDS.

II-Background

Even though no system theory has ever been developed for two point boundary value descriptor systems (TPBVDS) there exist results from other system theories that are very helpful in our study. In fact, the development of our theory, in many instances, parallels that of the (continuous-time, non descriptor) boundary value linear system (BVLS) theory studied extensively by Krener [1,2] and by Gohberg and Kaashoek [3-5]. The reason a direct analogy between TPBVDS and BVLS does not exist is the singular nature of the dynamics of the TPBVDS. As we will see later in the case of the BVLS the state transition matrix is always non-singular. In the TPBVDS case, however, the state transition matrix, in general, is singular. This, of course, complicates our study. Some work has been carried out, however, by Luenberger, Lewis, Cobb, Yip and Sincovec, and others [6-14] on a special class of TPBVDS's which corresponds to what we will call separable systems. In Section 2.1, we review those results from existing descriptor system theory that are of use to us. In this development we draw on material from [6,7,12,22]. Section 2.2 is devoted to a review of the BVLS theory.

2.1-Descriptor Dynamics

In this section, we consider the discrete descriptor system

$$Ex_{k+1}=Ax_k+Bu_k \qquad k=0,1,2,\ldots,N-1 \qquad (2.1.1)$$

$$y_k=Cx_k \qquad k=0,1,2,\ldots,N \qquad (2.1.2)$$

where E and A are arbitrary nxn matrices, and B and C are nxm and pxn matrices, respectively.

2.1.1-Introduction

Solvability of system (2.1) has been examined by Gantmacher [22] and Luenberger [6]. Solvability is the property that guarantees that system (2.1) has a solution x_k for any input sequence u_k. The system (2.1) is solvable if and only if det(sE-A) does not vanish identically. Matrices E and A satisfying this condition constitute, by definition, a regular pencil. In our study, we will assume that {E,A} always comprise a regular pencil.

Equation (2.1) can also be written as follows

$$S\begin{bmatrix} x_0 \\ \vdots \\ x_N \end{bmatrix}=\begin{bmatrix} Bu_0 \\ \vdots \\ Bu_{N-1} \end{bmatrix} \qquad (2.2.1)$$

where

$$S=\begin{bmatrix} -A & E & & \\ & -A & E & \\ & & \ddots & \ddots \\ & & -A & E \end{bmatrix} \qquad (2.2.2)$$

(missing elements of the matrix S in equation (2.2.2) are zero). Solvability of system (2.1) implies that S has full

row rank, however, since S is not square, the solution x_k is not unique and thus additional conditions are needed to make the system well-posed. Before considering these additional conditions we present the following useful properties of regular pencils {E,A}.(for detail see [9,11]).

a) There exist invertible matrices V and W such that

$$VEW = \begin{bmatrix} I & \\ & \mathcal{N} \end{bmatrix} \quad \text{and} \quad VAW = \begin{bmatrix} J & \\ & I \end{bmatrix} \tag{2.3}$$

where $\mathcal{N}$ is nilpotent. Since we can multiply (2.1) on the left by V without affecting the system, then by a simple change of coordinate $z=W^{-1}x$, we can decouple the system into two subsystems

$$z^1_{k+1} = Jz^1_k + B_1u_k \tag{2.4.1}$$

$$z^2_{k-1} = \mathcal{N}z^2_k - B_2u_{k-1} \tag{2.4.2}$$

$$y_k = C_1z^1_k + C_2z^2_k \ . \tag{2.4.3}$$

Note that z^1 is recursive in the forward direction and z^2 in the backward direction.

b) In a deleted neighborhood of zero (i.e. away from s=0), the following Laurent expansion exists

$$(sE-A)^{-1} = s^{-1}\sum_{k=-\mu}^{\infty} \phi_k s^{-k} \tag{2.5}$$

where μ is the index of nilpotency of the pencil {E,A} (i.e. $\mathcal{N}^{\mu-1}\neq 0$, $\mathcal{N}^{\mu}=0$). The sequence of matrices ϕ_k is called the sequence of relative fundamental matrices.

c) Relative Cayley-Hamilton theorem

Let

$$\Delta(s) = \det(sE-A) = p_0 s^n - p_1 s^{n-1} \ldots - p_n \qquad (2.6.1)$$

then

$$\Delta(\phi_k) = p_0\phi_k - p_1\phi_{k-1} \ldots - p_n\phi_{k-n} = 0, \; k \geq n \text{ and } k < -1. \qquad (2.6.2)$$

2.1.2-Well-Posedness

As seen previously, we need additional conditions to make the descriptor system well-posed. In the literature, the issue od boundary conditions has not been dealt with thoroughly and in some cases not at all. For example, in his work Lewis does not consider any additional conditions and only studies the descriptor dynamics.

Yip and Sincovec, and Bender consider a very special type of additional condition, namely, an initial condition. Their system is an extension of the continuous-time descriptor system (see [13,17]). In essence, they decompose the system as in (2.4) and consider an initial condition for system (2.4.1). They also consider the system to be defined over all non-negative integers so that the final condition required for system (2.4.2) to make (2.4) well-posed becomes irrelevent, since $\mathcal{N}$ is nilpotent.

Luenberger has considered a slightly more general case.

He has shown that a complete set of additional conditions can be specified in terms of pure initial and pure final conditions. We call these systems separable.

A separable system is a special case of a TPBVDS. The boundary condition of our more general class of TPBVDS's is a linear function of the initial and the final state. All the above authors except Lewis, explicitly consider separable systems. Lewis's work. also, implicitly considers separable systems. Many of our results reduce to the results obtained by these authors in the separable case, and in fact our analysis provides new and useful results and insights in this special case.

2.1.3-Reachability and Observability

There are a variety of definitions for reachability and observability for the descriptor system (2.1) with the special separable boundary conditions described previously given in the literature [12-14,17,18]. In this section, we present Lewis's definitions for reachability and observability. Lewis does not consider any boundary conditions for his descriptor dynamics; however, in his definition of reachability, he implicitly assumes causality. Therefore, it is not surprising that his defintion is essentially equivalent to that proposed by Yip and Sincovec, and Bender for their systems.

Reachability for causal systems is defined as the ability to drive the system from any initial state to any desired state by proper choice of the inputs. Lewis's definition (given below) is very similar.

<u>Definition 2.1</u>

System (2.1) is reachable if for any z_1 and $z_2 \in \mathbb{R}^n$, there exist controls u_j where $j \in [0,N-1]$ for some $N>0$, such that x_k where $k \in [0,N]$ is a solution to (2.1.1) and $(x_0, x_N) = (z_1, z_2)$.

He has shown that the following statements are equivalent.

a) System (2.1) is reachable.

b) The matrices $[sE-A \mid B]$ and $[E \mid B]$ have full column rank for all s.

c) Rank $[\Phi_{-\mu}B \mid \ldots \mid \Phi_{-1}B \mid \Phi_0 B \mid \ldots \mid \Phi_{n-1}B] = n$.

d) The matrices $[B_1 \mid JB_1 \mid \ldots \mid J^{n-1}B_1]$ and $[B_2 \mid NB_2 \mid \ldots \mid N^{\mu-1}B_2]$ have full column rank.

The problem of observability is more complex. Yip and Sincovec's defintion of observability differs from that of Lewis and Cobb. Lewis and Cobb's concept for observability is the dual of the reachability concept and it turns out to be the one most useful in our study. At this point, it is

very difficult to explain Lewis's definition given below and the only expanation seems to be that it leads to a concept dual to the concept of reachability. In Chapter III, we explore the significance of this concept.

<u>Definition 2.2</u>

System (2.1) is observable if for $u_k=0$ and some $N>0$, knowledge of the output y_k where $k\epsilon[0,N]$ is sufficient to uniquely determine Ax_0 and Ex_{N+1}.(The choice of the final time (i.e. N+1 instead of N) results in duality with the definition of reachability).

The following statements are equivalent.

a) System (2.1) is observable.

b) The matrices $\begin{bmatrix} sE-A \\ C \end{bmatrix}$ and $\begin{bmatrix} E \\ C \end{bmatrix}$ have full rank for all s.

c) Rank $\begin{bmatrix} \phi_{-\mu}C \\ \vdots \\ \phi_n C \end{bmatrix} = n.$

d) The matrices $\begin{bmatrix} C_1 \\ \vdots \\ C_1 J^{n-1} \end{bmatrix}$ and $\begin{bmatrix} C_2 \\ \vdots \\ C_2 N^{\mu-1} \end{bmatrix}$ have full rank.

We will see in Chapter III that it is possible to interpret these reachability and observability conditions for the more general TPBVDS without imposing a restrictive separability assumption on the boundary conditions. Not only

do these interpretations shed more light on reachability and observability for descriptor systems, but they also lead to reachability and observability matrices expressed directly in terms of the original matrices A, E, B, and C. In addition, in the more general context of TPBVDS-.i.e. where we explicityly account for the nature of the boundary conditions- we show that, much as in the work of Krener [2] to be discussed in the next subsection, it is actually necessary to define <u>two</u> distinct notions each for reachability and observability.

2.2-Boundary Value Linear Systems

In this section, we review some of the results obtained by Krener [2] for linear boundary value problems specified by

$$\dot{x}(t)=Ax(t)+Bu(t) \tag{2.7.1}$$

$$V^0x(t_0)+V^1x(t_1)=v \tag{2.7.2}$$

$$y(t)=Cx(t) \tag{2.7.3}$$

where A, B and C are real-analytic functions of t. We will call the system autonomous if A,B and C are constant.

System (2.7) is well-posed (i.e. it has a unique solution x for any arbitrary input) if and only if

$$F=V^0+V^1\phi(t_0,t_1) \tag{2.8}$$

where ϕ is the state transition matrix. Thus, if (2.7) is well-posed, without loss of generality we may assume that $F=I$ (that is because we can premultiply (2.7.2) by F^{-1}). If this condition is satisfied, the boundary condition (and the system) is said to be in standard form. The solution to (2.7) is given by

$$x(t)=\phi(t,t_0)v+\int_{t_0}^{t_1}G(t,s)B(s)u(s)ds \tag{2.9}$$

where

$$G(t,s)=\begin{cases}\phi(t,t_0)V^0\phi(t_0,s) & t>s\\ -\phi(t_1,t)V^1\phi(t_1,s) & t<s\ .\end{cases} \tag{2.10}$$

Clearly, the Green's function is not causal and in fact $x(t)$ is a function of all inputs $u(t)$ over the interval $[0,T]$.

Krener has proposed a notion of causality for the system (2.7). He has shown that there exists an inward boundary value process $k(\tau_0,\tau_1)$ such that the system

$$\dot{x}=Ax+Bu \tag{2.11.1}$$

$$K^0x(\tau_0)+K^1x(\tau_1)=k(\tau_0,\tau_1) \tag{2.11.2}$$

has the same solution $x(t)$ for $\tau_0<t<\tau_1$ as system (2.7). $k(\tau_0,\tau_1)$ is only a function of v and of the inputs $u(t)$ for t off the interval $[\tau_0,\tau_1]$, i.e. for $t\epsilon[t_0,t_1]/[\tau_0,\tau_1]$. In fact

$$k(\tau_0,\tau_1)=\int_{t_0}^{\tau_0}+\int_{\tau_1}^{t_1} G(\tau_0,s)B(s)u(s)ds \tag{2.12}$$

$$K^0=\phi(\tau_0,t_0)V^0\phi(t_0,\tau_0) \tag{2.13.1}$$

$$K^1=\phi(\tau_0,t_0)V^1\phi(t_1,\tau_1) \tag{2.13.2}$$

So, if we think of t off $[\tau_0,\tau_1]$ as the past and t on $[\tau_0,\tau_1]$ as the future, $k(\tau_0,\tau_1)$ is a causal process. In the same way it is possible to define an anticausal process

$$j(\tau_0,\tau_1)=\int_{\tau_0}^{\tau_1} \phi(\tau_1,s)B(s)u(s)ds \tag{2.14.1}$$

which depends only on u(t) on the interval $[\tau_0,\tau_1]$. The process j is called the outward boundary process or the jump process and can be shown to be equal to

$$j(\tau_0,\tau_1)=x(\tau_1)-\phi(\tau_1,\tau_0)x(\tau_0). \tag{2.14.2}$$

It can be shown that $x(\tau_0)$ and $x(\tau_1)$ can be uniquely recovered from $j(\tau_0,\tau_1)$ and $k(\tau_0,\tau_1)$. In fact

$$\begin{bmatrix} x(\tau_0) \\ x(\tau_1) \end{bmatrix} = \begin{bmatrix} I & -\phi(\tau_0,t_0)V^1\phi(t_1,\tau_1) \\ \phi(\tau_1,\tau_0) & \phi(\tau_1,t_0)V^0\phi(t_0,\tau_1) \end{bmatrix} \begin{bmatrix} k(\tau_0,\tau_1) \\ j(\tau_0,\tau_1) \end{bmatrix}. \tag{2.15}$$

There are two useful definitions of both controllability and observability for system (2.7): one associated with the process j and the other associated with the process k. The system is defined to be <u>controllable off</u> if the process k is controllable (i.e. by proper choice of the input function outside any interval $[\tau_0,\tau_1]$, $k(\tau_0,\tau_1)$ can be made arbitrary), and it is defined to be <u>observable</u>

<u>off</u> if the process j is observable (i.e. by observing the output function off any interval $[\tau_0,\tau_1]$, $j(\tau_0,\tau_1)$ can be uniquely determined). The system is defined to be <u>controllable on</u> if j is controllable and finally the system is defined to be <u>observable on</u> if k is observable. Notice that controllabilty on and observability on are simply causal controllability and observability.

It can be shown that controllability on and observability on are stronger conditions that controllabilty off and observability off. Thus, any system that is controllable on (observable on) is also controllable off (observable off).

Krener has shown that system (2.7) is minimal if and only if it is controllable and observable off and any state which is unobservable on is controllable on.

In our work, we are interested in the class of autonomous systems. It turns out that the class of autonomous BVLS's considered by Krener, in general, does not contain a minimal realization. However, this <u>is</u> true if we further restrict ourselves to <u>stationary</u> systems. Specifically, a system is defined to be stationary if it is autonomous and its Green's function G(t,s) is only a function of t-s. It can be shown that (2.7) is stationary if the boundary matrices V^0 and V^1 commute with A.

We will see in Chapter III that two point boundary

value descriptor systems (TPBVDS) can be studied in a way similar to the study of boundary value linear systems (BVLS) presented here. In the case of TPBVDS's, however, the singular nature of the dynamics of the system makes for some important differences and for some more delicate analysis to derive and prove results.

III- TWO-POINT BOUNDARY VALUE DESCRIPTOR SYSTEM (TPBVDS): The Deterministic Problem

In this chapter we consider the deterministic two-point boundary value descriptor system (TPBVDS)

$$Ex_{k+1} = Ax_k + Bu_k \tag{3.1.1}$$

$$V^i x_0 + V^f x_N = v \tag{3.1.2}$$

$$y_k = Cx_k \tag{3.1.3}$$

where E, A, V^i, and V^f are constant nxn matrices; and B and C are constant nxm and pxn matrices, respectively. The sequence u_k is the input and y_k the output of system (3.1). We also assume that N>2n. It should be clear that the TPBVDS (3.1) is the natural extension, to the discrete case, of the autonomous BVLS described in the previous section, since the system is as noncausal as possible. Note indeed that the dynamics (3.1.1) are noncausal and that the boundary condition (3.1.2) involves both the initial and the final value of the state, so that no time direction is preferred.

3.1- Well-Posedness and Standard Form:

Definition 3.1:

The system (3.1) is well-posed if it has a unique

solution x_k for any input sequence u_k.

By writing together in matrix form all the equations (3.1.1) (i.e. for k=0,...,N-1) and (3.1.2), we obtain

$$S\begin{bmatrix} x_0 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} Bu_0 \\ \vdots \\ Bu_{N-1} \\ v \end{bmatrix} \tag{3.2.1}$$

where

$$S = \begin{bmatrix} -A & E & & & \\ & -A & E & & \\ & & \cdot & \cdot & \\ & & & -A & E \\ V^i & & & & V^f \end{bmatrix}. \tag{3.2.2}$$

It is easy to see a necessary and sufficient condition for well-posedness is that the matrix S be invertible.

A necessary condition for well-posedness is that {E,A} comprise a regular pencil (see section 2.1). To obtain a sufficient condition we would like to be able to perform simple row cancellations on the matrix S. But in order to accomplish this, E and A must commute. Fortunately this can be guaranteed. Specifically, assuming that {E,A} is a regular pencil, there exist α and β such that $\alpha E+\beta A$ is invertible. Then we simply multiply equation (3.1.1) by $(\alpha E+\beta A)^{-1}$ on the left. This does not change the system or the state variable x but guarantees that the resulting new E and A matrices commute.

Note indeed that since

$$\tilde{E} = (\alpha E+\beta A)^{-1}E \quad , \quad \tilde{A} = (\alpha E+\beta A)^{-1}A$$

are the new E and A matrices, we have

$$\alpha\tilde{E}+\beta\tilde{A} = I \tag{3.3.1}$$

and

$$\tilde{A}(\alpha\tilde{E}+\beta\tilde{A}) = (\alpha\tilde{E}+\beta\tilde{A})\tilde{A} \tag{3.3.2}$$

so that if $\alpha\neq 0$, we must have $\tilde{E}\tilde{A}=\tilde{A}\tilde{E}$, i.e. $\tilde{E}$ and $\tilde{A}$ commute. In the case when $\alpha=0$ we can take $\beta=1$ and $\tilde{A}=I$, so that $\tilde{E}$ and $\tilde{A}$ necessarily commute.

For this new form of the system, we have

$$\alpha E + \beta A = I \ . \tag{3.4}$$

Now since E and A commute we can examine the invertibility of the matrix S by simple row elimination, and thus we obtain the following result.

<u>Theorem 3.1:</u>

Let $\alpha E + \beta A = I$ for some α and β , then (3.1) is well-posed if and only if

$$V^i E^N + V^f A^N$$

is invertible.

This result is obtained by applying row elimination to solve for x_0 and x_N in (3.2.1) (also see Appendix B).

<u>Definition 3.2:</u>

The TPBVDS (3.1) is in Standard Form if and only if

(1) $\exists\ \alpha,\beta \in R$ such that $\alpha E+\beta A = I$

(2) $V^iE^N+V^fA^N = I$

The concept of standard form is motivated by Theorem (3.1). It is clear that any well-posed TPBVDS can be transformed into Standard Form without any change of coordinates. That is we only use left multiplications on the system equations (3.1.1) and (3.1.2) first to ensure that (1) is satisfied by premultiplying (3.1.1) by $(\alpha E+\beta A)^{-1}$, and then to obtain (2) by multiplying (3.1.2) by $(V^iE^N+V^fA^N)^{-1}$. On the other hand, any system in Standard Form is well-posed. From this point on, we shall assume that system (3.1) is in Standard Form.

The solution to (3.1) is obtained by writing the matrix S as the following product

$$S = \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ (V^iA^N+\omega V^fE^N)\Gamma^{-1} & (V^iEA^{N-1}+\omega V^fE^2A^{N-2})\Gamma^{-1} & \cdots & (V^iE^N+V^fA^N)\Gamma^{-1} \end{bmatrix} \cdot \begin{bmatrix} -A & E & & & \\ & -A & E & & \\ & & \ddots & \ddots & \\ & & & -A & E \\ \omega E & & & & -A \end{bmatrix} \tag{3.5}$$

where

$$\Gamma = \omega E^{N+1} - A^{N+1} \tag{3.6}$$

and where ω is any scalar for which Γ is invertible. Note that ω can always be chosen so that Γ is invertible because if α and β are both non-zero then since $\alpha E+\beta A=I$, E and A have the same Jordan structure, and thus since $\{E,A\}$ comprise a regular pencil, E and A cannot both have a zero eigenvalue associated to a common eigenvector which implies that $\{E^k,A^k\}$ comprise a regular pencil for all k. When either α or β is zero clearly either E or A must equal I in which case $\{E^k,A^k\}$ comprise again a regular pencil.

Inverting each of the two matrices in (3.5) separately gives us S^{-1}. The entries of the S^{-1} matrix form the Green's function solution of system (3.1) because

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix} = S^{-1} \begin{bmatrix} u_0 \\ \vdots \\ u_{N-1} \\ v \end{bmatrix}. \tag{3.7}$$

We obtain

$$x_k = A^k E^{N-k} v + \sum_{i=0}^{N-1} G(k,i).B.u_i \tag{3.8}$$

where

$$G(i,j) = \begin{cases} A^i(A-E^{N-i}(V^iA+\omega V^fE)E^i)E^{j-i}A^{N-j-1}\Gamma^{-1} & j\geq i \\ E^{N-i}(\omega E-A^i(V^iA+\omega V^fE)A^{N-i})E^jA^{i-j-1}\Gamma^{-1} & i>j. \end{cases} \tag{3.9}$$

G is called the Green's function of system (3.1.1)-(3.1.2)

and the input-output weighting pattern is given by

$$W(i,j) = CG(i,j)B \quad . \tag{3.10}$$

Note that the Green's function is the analog of the state-transition matrix for deterministic causal systems, and that the weighting pattern is the analog of the system's impulse response.

In the rest of this chapter, we assume that Γ is invertible for $\omega=1$ and use the expression (3.9) for G with ω set equal to 1. This, of course, means that no $(N+1)^{st}$ root of unity is an eigenmode of the system. We do this only for simplicity in what follows. All of the results in this chapter have obvious extensions to the case of an arbitrary value of ω (essentially we must simply carry ω along in the various expressions).

3.2- Strong Reachability / Strong Observability:

3.2.1-Inward and Outward Boundary Value Processes

Similarly to Krener's outward process j (see Equation (2.15)), we define the outward boundary value process

$$z(i,j)=-A^{j-i}x_i+E^{j-i}x_j=$$
$$A^{j-i-1}Bu_i+EA^{j-i-2}Bu_{i+1}+\ldots+E^{j-i-1}Bu_{j-1} \qquad j>i+1 \tag{3.11.1}$$

and,

$$z(i,i+1)=-Ax_i+Ex_j=Bu_i. \tag{3.11.2}$$

Note, however, that since E and A could be singular, z(i,j)

can only be propagated outward whereas, in the continuous case, the outward process j could be propagated inward as well as outward.

The above expression for z(i,j) is obtained by cancelling x_{i+1} through x_{j-1} in Equations (3.2.1) by simple row manipulations on matrix S (see Appendix A). It is easy to show that the four point boundary system

$$Ex_{k+1} = Ax_k + Bu_k \tag{3.12.1}$$

$$V^i x_0 + V^f x_N = v \tag{3.12.2}$$

$$-A^{j-i} x_i + E^{j-i} x_j = z(i,j) \tag{3.12.3}$$

has the same solution as system (3.1) for $k \epsilon [0,N] \backslash [i+1,j-1]$ (i.e. system (3.1) has the same solution as the system (3.11) over [0,i] and [j,N]). So, z is analogous to the anticausal process j defined by Krener. It is also possible to define a process z'(k,l) analogous to Krener's k process. Unfortunately, in the general case, z'(k,l) is a very complex function of the boundary matrices and of the controls u_j for $j \epsilon [0,N-1] \backslash [k,l]$, which we write as

$$z'(k,l) = W^i_{kl} x_k + W^f_{kl} x_l =$$

$$F_{kl}(u_0, u_1, \ldots, u_{k-1}, u_l, u_{l+1}, \ldots, u_{N-1}, v] \tag{3.13.1}$$

where l>k and

$$z'(k,k) = x_k = F_{kk}(u_0, \ldots, u_{N-1}, v). \tag{3.13.2}$$

The above expression for z' and the matrices W^i and W^f

can be computed recursively as described in Appendix A. There we show that when E and A are both invertible, W^i_{kl} and W^f_{kl} have ranks independent of k and l and equal to the ranks of V^i and V^f respectively. We also show that when E and A are both invertible W^i_{kl} and W^f_{kl} have a simple form, similar to that of the inward boundary matrices for the continuous case (Equations (2.13.1) and (2.13.2)); moreover, the Green's function solution is simplified and looks very much like the Green's function solution in the continuous case (Equation (2.10)). In this case, as in the continuous case, we can propagate the inward process outward as well as inward, a property which is not true in general.

It can be shown that the TPBVDS

$$Ex_{k+1} = Ax_k + Bu_k \tag{3.14.1}$$

$$W^i_{jl}x_j+W^f_{jl}x_l=z'(j,l) \tag{3.14.2}$$

has the same solution as system (3.1) for $k\epsilon[j,l]$. We will see later in this chapter that, for stationary systems, the function F takes a simple form.

3.2.2-Notation

Several concepts of reachabilty and observability for TPBVDS will be discussed in Chapter III. In each case, we have to introduce a new symbol for the reachability and the observability matrices and the reachable and the

unobservable spaces. To prevent confusion, we introduce in this section the notation that will be used throughout Chapter III.

1) $R(i,j)$ $\{=R(j-i)\}$: The strong reachability matrix on $[i,j]$ (reachability matrix of $z(i,j)$),

$$z(i,j)=R(i,j)\begin{bmatrix} u_i \\ \vdots \\ u_{j-1} \end{bmatrix}.$$

2) $\mathcal{R}$: The strongly reachable space,

$$\mathcal{R}=\mathrm{Im}(R(n)).$$

3) $O(i,j)$ $\{=O(j-i)\}$: The strong observability matrix on $[i,j]$ (observability matrix of $z'(i,j)$),

$$\begin{bmatrix} y_i \\ \vdots \\ y_j \end{bmatrix}=O(i,j)z'(i,j).$$

4) $\mathcal{O}$: The strongly unobservable space,

$$\mathcal{O}=\mathrm{Ker}(O(n)).$$

5) $R'(i,j)$: The reachability matrix off $[i,j]$ (reachability matrix of $z'(i,j)$),

$$z'(i,j)=R'(i,j)\begin{bmatrix} u_0 \\ \vdots \\ u_{i-1} \\ u_j \\ \vdots \\ u_{N-1} \end{bmatrix}.$$

6) $\mathcal{R}'(i,j)$: The reachable space off $[i,j]$ (the reachable space of $z'(i,j)$),

$$\mathcal{R}'(i,j)=\mathrm{Im}(R'(i,j)).$$

7) $O'(i,j)$: The observability matrix off $[i,j]$ (observability matrix of $z(i,j)$),

$$\begin{bmatrix} y_0 \\ \vdots \\ y_i \\ y_j \\ \vdots \\ y_N \end{bmatrix} = O'(i,j)z(i,j).$$

8) $\mathcal{O}'(i,j)$: The unobservable space off $[i,j]$ (the unobservable space of $z(i,j)$),

$$\mathcal{O}'(i,j)=\mathrm{Ker}(O'(i,j)).$$

9) R_i': The reachability matrix of x_i $(=z'(i,i))$,

$$R_i'=R'(i,i).$$

10) O_i': The observability matrix of Bu_i $(=z(i,i+1))$,

$$O_i'=O'(i,i+1).$$

11) $\mathcal{R}_i'$: The reachable space of x_i,

$$\mathcal{R}_i'=\mathrm{Im}(R_i').$$

12) $\mathcal{O}_i'$: The unobservable space of Bu_i,

$$\mathcal{O}_i'=\mathrm{Ker}(O_i').$$

13) $\mathcal{R}'$: The reachable space in the stationary case (see section 3.4),

$$\mathcal{R}'=\mathcal{R}_i' \qquad i\epsilon[n,N-n].$$

14) $\mathcal{O}'$: The unobservable space in the stationary case

$$\mathcal{O}'=\mathcal{O}_i' \qquad i\epsilon[n,N-n].$$

3.2.3- Strong Reachability and Strong Observabilty Matrices and Spaces

Definition 3.3.1:

The system (3.1) is reachable on $[i,j]$ if the map

$$\{u_k : k\epsilon[i,j-1]\} \rightarrow z(i,j) \text{ is onto.}$$

System (3.1) is called strongly reachable or reachable on if it is reachable on $[i,j]$ for all i and j such that $j-i \geq n$.

By writing the system in forward-backward form, as was done in the previous chapter, we can show that for the special case of descriptor systems where the backward model is nilpotent and for the more general case where the backward model is not necessarily nilpotent, reachability on is identical to the concept of reachability presented in Chapter II even though they have been defined in terms of different variables (reachability on is defined in terms of z and reachability in terms of x). This will become clear when in Theorem (3.3) (statements (c) and (d)), we derive necessary and sufficient conditions for reachability on and compare them to the necessary and sufficient conditions for reachability presented in in Chapter II. For reasons that will become apparent later, we use the expression strong reachability instead of reachability on.

From (3.11.1), it is clear that the reachability matrix

for z is

$$R(i,j) = [A^{j-i-1}B \vdots EA^{j-i-2}B \vdots \ldots \vdots E^{j-i-1}B] = R(j-i) \qquad (3.15)$$

where we have abused notation, since $R(i,j)$ depends only on $j-i$. Thus, $R(j-i)$ is the strong reachability matrix on $[i,j]$ associated to the system (3.1). It is clear that

a) $\mathrm{Im}\{R(k+1)\} = \mathrm{Im}\{ER(k)\} + \mathrm{Im}\{AR(k)\}$

b) $\mathrm{Im}\{R(k)\} \subseteq \mathrm{Im}\{R(k+1)\}$.

Using the above two properties of $R(.)$ we obtain the following results,

$$\forall\, k, \qquad \mathrm{Im}(R(k)) \subseteq \mathrm{Im}(R(n))$$

$$\text{and} \qquad \mathrm{Im}(R(k)) = \mathrm{Im}(R(n)) \qquad \text{for } k \geq n.$$

So, natural choices of the strong reachability matrix R and strong reachability space $\mathfrak{R}$ are

$$R = R(n)$$

$$\mathfrak{R} = \mathrm{Im}(R).$$

An alternative derivation of $\mathfrak{R}$ is possible using the Generalized Cayley-Hamilton theorem given below.

<u>Theorem 3.2 {Generalized Cayley-Hamilton Theorem}</u>

Let $\{E,A\}$ be a regular pencil in standard form. Then for all $K,L \geq 0$, $\exists\, \alpha_0,\ldots,\alpha_{n-1} \in R$ such that

$$E^K A^L = \sum_{i=0}^{n-1} \alpha_i A^{n-i-1} E^i \quad . \qquad (3.16)$$

This theorem basically states that the space of matrices $E^K A^L$ is spanned by

$$\{A^{n-1},\ EA^{n-2}, \ldots,\ E^{n-1}\}=\{E^i A^j,\ i,j\geq 0,\ i+j=n-1\}.$$

The proof of this theorem relies on the fact that for some α and β, $\alpha E+\beta A = I$. Suppose that $\alpha\neq 0$, then we can express E as a function of A in $E^K A^L$. Apply the usual Cayley-Hamilton theorem to all the powers of A higher than n-1. Then multiply all A^k 's with $(\alpha E+\beta A)^{n-k-1}$'s which are equal to the identity matrix and finally expand the resulting expression.

So, we have the following theorem.

<u>Theorem 3.3:</u>

The following statements are equivalent

a) System (3.1) is strongly reachable.

b) The strong reachability matrix R has full rank.

c) The matrix $[sE-tA \vdots B]$ has full rank for all $(s,t)\neq(0,0)$.

d) The state x_i where $i\epsilon[n,N-n]$ can be made arbitrary by proper choice of the inputs u_j: $j\epsilon[i-n,i+n-1]$ with all other inputs and the boundary value v set to zero, and for all pair of matrices V^i and V^f in standard form. See Fig. 1.

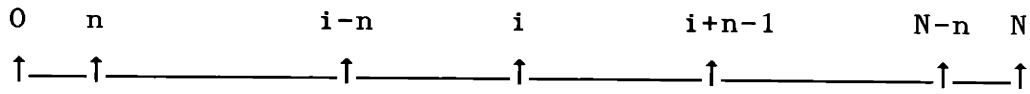

Fig. 1

Note that in statement d), we require that x_i can be made arbitrary by applying proper controls over the 2n point symmetric neighbourhood of i. In fact, we only need an n point neighbourhood of i, but then this interval is not necessarily symmetric and its position depends on the matrices E, A and B. The union of all possible such n point neighbourhoods, however, is the 2n point symmetric neighbourhood of i. Later, we introduce the concept of reachability where we require that x_i can be made arbitrary by proper choice of inputs u over the <u>whole</u> interval [0,N]. Clearly, reachability is a weaker condition than the notion of strong reachability presented here.

<u>Proof</u>

Statement b) is obtained by noting that

$$z(i,j) = R(j-i)\begin{bmatrix} u_i \\ u_{i+1} \\ \vdots \\ u_{j-1} \end{bmatrix}. \tag{3.17}$$

Statement c) is proven as follows. First, assume that $\alpha \neq 0$. In this case

$$\mathfrak{R} = \mathrm{Im}[B \vdots AB \vdots \ldots \vdots A^{n-1}B].$$

This is a direct consequence of the Generalized Cayley-Hamilton theorem. Now, since $\alpha E+\beta A=I$,

$$[sE-tA\,|\,B]=[(s/\alpha)I-(t-s\beta/\alpha)A\,|\,B]=[uI-vA\,|\,B]. \qquad (3.18)$$

It is clear that $(u,v)=(0,0)$ if and only if $(s,t)=(0,0)$. When $v\neq 0$, the problem is reduced to the well known causal case (see for example [23]) because we can write Equation (3.18) as $[(u/v)I-A\,|\,B]$. If v equals zero, u must be nonzero and thus $[uI\,|\,B]$ has full rank. For the case where $\alpha=0$, we have $\beta\neq 0$ and thus we can argue similarly by replacing A by E. Note that if $\alpha\neq 0$ and $\beta\neq 0$ then

$$\mathfrak{R}=\mathrm{Im}[B\,|\,AB\,|\,\ldots\,|\,A^{n-1}B]=\mathrm{Im}[B\,|\,EB\,|\,\ldots\,|\,E^{n-1}B].$$

Statement d) is obtained by writing x_i in terms of $z(i-n,i)$ and $z(i,i+n)$, using expression (3.9), as follows:

$$\begin{aligned} x_i = {} & A^i(A-E^{N-i}(V^iA+V^fE)E^i\ )\ \Gamma^{-1}z(i,i+n) \\ & + E^{N-i}(\ E-A^i(V^iA+V^fE)A^{N-i})\Gamma^{-1}z(i-n,i). \end{aligned} \qquad (3.19)$$

Now, let $z(i,i+n)=-A^{N-i}\xi$ and $z(i-n,i)=E^i\xi$, where ξ is an arbitrary vector. This can be done because R has full rank. By replacing the expressions for $z(i-n,i)$ and $z(i,i+n)$ in (3.19) we obtain

$$x_i = \Gamma\Gamma^{-1}\xi = \xi. \qquad (3.20)$$

This of course means that x_i can be made arbitrary.

<u>Definition 3.3.2:</u>

The system (3.1) is <u>observable on</u> $[i,j]$ if the map

$z'(i,j) \rightarrow \{y_k : k \in [i,j]\}$ is one to one.

System (3.1) is called <u>strongly observable</u> or <u>observable on</u> if it is observable on $[i,j]$ for all i and j such that $j-i \geq n$.

Following a reasoning similar to the one used for the reachability case, we can show that the observability matrix is given by

$$O = \begin{bmatrix} CA^{n-1} \\ CEA^{n-2} \\ \vdots \\ CE^{n-1} \end{bmatrix}. \tag{3.21}$$

The strongly unobservable space is

$$\mathcal{O} = \mathrm{Ker}(O).$$

<u>Theorem 3.4</u>

The following statements are equivalent

a) System (3.1) is strongly observable.

b) The strong observability matrix O has full rank.

b) The matrix $\begin{bmatrix} sE-tA \\ C \end{bmatrix}$ has full rank for all $(s,t) \neq (0,0)$.

d) The state x_i where $i \in [n, N-n]$ can be uniquely determined from the outputs y_j: $j \in [i-n, i+n-1]$ for all V^i and V^f in standard-form.

Later, we shall introduce the concept of observability where x_i can be uniquely determined from y_j's over the <u>whole</u> interval [0,N]. Clearly, observability is a weaker condition than strong observability.

These strong reachability and observability properties are properties of the dynamics of the system given by equation (3.1.1) and do not depend on the boundary conditions. In the next section we present alternative definitions for reachability and observability which depend on the boundary conditions, and later we show how all of these notions are needed to analyze system (3.1).

In the case of continuous-time boundary value linear systems Krener has shown that the controllability and observability on are simply the causal controllability and observability. In our case, however, this is true only if E is invertible. This is one difference that arises in the descriptor context. Others will become clear as we proceed.

3.3- Reachability and Observability: General Case

In Chapter II, we saw that for boundary value linear systems it is useful to define the concept of reachability and observability off. Reachability off (reachability) and observability off (observability) can be easily defined for the case of TPBVDS as well.

Definition 3.4:

The system (3.1) is reachable off $[i,j]$ if the map

$$\{u(k): k\epsilon[0,i-1]\cup[j,N-1]\} \rightarrow z'(i,j) \text{ is onto.}$$

System (3.1) is called reachable if it is reachable off $[i,j]$ for all i and j such that $n\leq i\leq j\leq N-n$.

Definition 3.5:

The system (3.1) is observable off $[i,j]$ if the map

$$z(i,j)\rightarrow \{y_k: k\epsilon[0,i]\cup[j,N] \text{ is one to one.}$$

System (3.1) is called observable if it is observable off $[i,j]$ for all i and j such that $n\leq i\leq j\leq N-n$.

Theorem 3.5

If the reachable space $\mathcal{R}'(i,j)$ of the inward process $z'(i,j)$ has dimension m for some $i,j\epsilon[n,N-n]$, then it has dimension m for all $i,j \epsilon [n,N-n]$.

Corollary

If the system (3.1) is reachable off some $i,j\epsilon[n,N-n]$, then it is reachable off all $i,j\epsilon[n,N-n]$.

Proof

See Appendix A.

Thus, if the reachable space of $z'(i,j)$ has dimension n for some i,j then it has dimension n for all $i,j \in [n,N-n]$. So, to test reachability, we can test the reachability space of $z'(i,i)=x_i$ for example. Similarly in the case of observability, testing, for example, whether $z(i,i+1)=Bu_i$ can be uniquely determined is enough to show if the system is observable or not. This of course means that reachability is equivalent to being able to make any x_i far enough from the boundaries arbitrary by proper choice of the inputs u and that observability is equivalent to being able to determine Bu_i from the outputs y.

The next step is to obtain the reachability and observability spaces. In many cases of interest such as the causal case and the cyclic case ($V^i=I$, $V^f=-I$) the two definitions of reachability (observability) coincide (this will become apparent later when we derive the reachability and observability spaces). But in general the reachability (observability) space has a more complex structure, and unlike the strong reachability (strong observability) space it is not time invariant. This means of course that we need to index the reachability (observability) space. Let R_i' be the reachability matrix for $z'(i,i)=x_i$, and let O_i' be the observability matrix for $z(i,i+1)$. Then, by noting that R_i' is the image space for the map from $\{u_0, \ldots, u_{N-1}\} \rightarrow x_i$ (the boundary value v is set to zero) and that R is the

image space for the map $\{u_{i-n}, \ldots, u_{i+n}\} \rightarrow x_i$, we can deduce that the image space of R is included in the image space of R_i' for $i \in [n, N-n]$ and similarly the null space of O_i' is included in the null space of O for $i \in [n, N-n]$. This of course means that reachability off (observability off) is a weaker condition that reachability on (observability on) i.e. reachability on implies reachability off and observability on implies observability off.

These time varying spaces can be computed as follows.

<u>Theorem 3.6</u>

The reachable space $\mathcal{R}_i'$ (unobservable space $\mathcal{O}_i'$) where $i \in [n, N-n]$ is given by

$$\mathcal{R}_i' = \mathrm{Im}[A^i E^{N-i}(V^i A + V^f E)R \quad R] = \mathrm{Im}[A^i E^{N-i}\{V^i R \quad V^f R\} \quad R] \tag{3.22}$$

$$\mathcal{O}_i' = \mathrm{Ker}\begin{bmatrix} O \\ O(V^i A + V^f E)A^{N-j-1}E^j \end{bmatrix} = \mathrm{Ker}\left[\begin{bmatrix} OV^i \\ OV^f \end{bmatrix} A^{N-j-1} \quad O \quad E^j\right] \tag{3.23}$$

where R (O) is the strong reachability (strong observability) matrix.

<u>Proof:</u>

We prove the result for the reachability case.

Using the expression (3.9) for G(i,j), it is clear that the reachability matrix R_i' is given by

$$R_i' = [A^i(A - E^{N-i}(V^i A + V^f E)E^i)R(N-i) \vdots E^{N-i}(E - A^i(V^i A + V^f E)A^{N-i})R(i)]$$

(3.24)

The above expression basically means that if $w \in Im(R_i')$ then $\exists x,y \in \mathcal{R}$ such that

$$w=A^i(A-E^{N-i}(V^iA+V^fE)E^i)x+E^{N-i}(E-A^i(V^iA+V^fE)A^{N-i})y \qquad (3.25)$$

or equivalently,

$$w=(A^{i+1}x + E^{N-i+1}y)-E^{N-i}A^i(V^iA+V^fE)(E^ix+A^{N-i}y). \qquad (3.26)$$

Thus, using the fact that $\mathcal{R}$ is E and A invariant,

$$w=s-E^{N-i}A^i(V^iA+V^fE)t \qquad \text{where } s,t \in \mathcal{R}. \qquad (3.27)$$

So, we have shown that

$$Im(R_i')\subset Im[A^iE^{N-i}(V^iA+V^fE)R \quad R]. \qquad (3.28)$$

Now, we have to show that any w in the range of $[A^iE^{N-i}(V^iR+V^fR) \quad R]$ is in the range of R_i'. Clearly, we can decompose w as in equation (3.27). What remains to be shown is that there exist x and y in $\mathcal{R}$ such that (3.26) is satisfied. But

$$\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} A^{i+1} & E^{N-i+1} \\ E^i & A^{N-i} \end{bmatrix}^{-1}\begin{bmatrix} s \\ t \end{bmatrix}=\begin{bmatrix} \Gamma^{-1}A^{N-i} & -\Gamma^{-1}E^{N-i+1} \\ -\Gamma^{-1}E^i & \Gamma^{-1}A^{i+1} \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix}. \qquad (3.29)$$

Since $\mathcal{R}$ is E,A,Γ and consequently Γ^{-1} invariant, x and y are in $\mathcal{R}$, which is the desired result.

Now, we show that

$$Im[A^iE^{N-i}(V^iA+V^fE)R \quad R]=Im[A^iE^{N-i}\{V^iR \quad V^fR\} \quad R]. \qquad (3.30)$$

It is clear using the fact that $\mathcal{R}$ is E and A invariant and $V^iE^N+V^fA^N=I$ that

$$Im((V^iA+V^fE)R \quad R)\subset Im\{V^iR \quad V^fR\}. \qquad (3.31)$$

Thus, we have to show that

$$\mathrm{Im}\{V^iR \quad V^fR\}\subset\mathrm{Im}((V^iA+V^fE)R \quad R). \tag{3.32}$$

But,

$$\mathrm{Im}((V^iA+V^fE)R \quad R)\supset\mathrm{Im}((V^iA+V^fE)E^NR \quad (V^iE^N+V^fA^N)AR \quad R)$$
$$\supset\mathrm{Im}(V^f(E^{N+1}-A^{N+1})R \quad R)=\mathrm{Im}(V^fR \quad R) \tag{3.33.1}$$

similarly

$$\mathrm{Im}((V^iA+V^fE)R \quad R)\supset\mathrm{Im}(V^iR \quad R) \tag{3.33.2}$$

and thus (3.32) holds. By using equations (3.31) and (3.32) and the fact that

$$\mathrm{Im}[A^iE^{N-i}(V^iA+V^fE)R \quad R]=\mathrm{Im}[A^iE^{N-i}((V^iA+V^fE)R \quad R) \quad R] \tag{3.34.1}$$

and

$$\mathrm{Im}[A^iE^{N-i}\{V^iR \quad V^fR\} \quad R]=\mathrm{Im}[A^iE^{N-i}(V^iR \quad V^fR) \quad R] \tag{3.34.2}$$

we obtain equation (3.30). Q.E.D.

We already know that the reachable space $\mathfrak{R}_i'$ has constant dimension for all i far enough from the boundaries. However, $\mathfrak{R}_i'$ could "rotate" in space; something that could never happen with time-invariant causal systems where reachability on and off coincide.

<u>Example 3.1</u>

Consider the following TPBVDS

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x_{k+1}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x_k+\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u_k \tag{3.35.1}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} x_0+\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} x_N=0 \tag{3.35.2}$$

$$y_k = x_k. \tag{3.35.3}$$

The system (3.35) is in standard form for all N. We would like to find the strong reachability and reachability spaces.

The strong reachability matrix R can be computed by using equation (3.14). This gives

$$R = \begin{bmatrix} 1 & 1 & \cdots\cdots \\ 0 & 0 & \cdots\cdots \\ 0 & 0 & \cdots\cdots \end{bmatrix}. \tag{3.36}$$

Clearly then, the strong reachability space $\mathcal{R}$ is spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. From Equation (3.22), we find that $\mathcal{R}_i'$ is the space spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ for i even and the space spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ for i odd. We see here that the reachability matrix rotates while keeping the same rank. This system is not reachable because $\mathcal{R}_i'$ is not $\mathbb{R}^3$. It is easy to check that this system is observable and strongly observable.

The question to ask at this point is what is a simple test for reachability (observability). The answer is given by the following theorem.

<u>Theorem 3.7</u>

a) System (3.1) is reachable if and only if the matrix $[EA(V^iA+V^fE)R \quad R]$ (or equivalently $[EA\{V^iR \quad V^fR\} \quad R]$) has

full rank.

b) System (3.1) is observable if and only if the matrix $\begin{bmatrix} 0 \\ 0(V^iA+V^fE)AE \end{bmatrix}$ (or equivalently $\begin{bmatrix} 0 \\ \begin{bmatrix} 0V^i \\ 0V^f \end{bmatrix} AE \end{bmatrix}$) has full rank.

We only prove part a). Part b) can be proven by a similar argument.

<u>Proof:</u>

We know that system (3.1) is reachable if the reachability space $\mathcal{R}_i'$ has dimension n for some i. So, we need to show that $\mathcal{R}_i'$ has dimension n if and only if

$$[EA(V^iA+V^fE)R \quad R]$$

has full rank.

First we show that for all subspaces $\mathcal{D}$ of $\mathbb{R}^n$

$$(E\mathcal{D} + \mathcal{R} = \mathbb{R}^n) \quad \Leftrightarrow \quad (E^k\mathcal{D} + \mathcal{R} = \mathbb{R}^n,\ \forall k>0) \qquad (3.37)$$

where $\Leftrightarrow$ signifies equivalence. Then, by letting $\mathcal{D}=\mathrm{Im}((V^iA+V^fE)R)$ we obtain the desired result.

Clearly, we need to show the two implications in (3.37).

i) We show $\Leftarrow$:

This is clear because multiplying E^{k-1} with vectors of $\mathcal{D}$ that are not in the E invariant space $\mathcal{R}$ could only make the space $\mathcal{D}'$ lose rank where $\mathcal{D}' \oplus \mathcal{R} = E\mathcal{D} + \mathcal{R}$.

ii) We show $\Rightarrow$:

By the above argument, we know that $\mathcal{D} + \mathcal{R} = \mathbb{R}^n$. Let H be an operator such that: $H(\xi) = E\xi + \mathcal{R}$ where ξ is any subspace of $\mathbb{R}^n$. Then,

$E\mathcal{D}+\mathcal{R} = \mathbb{R}^n \iff H(\mathbb{R}^n) = \mathbb{R}^n$. But, this implies that $H^k(\mathbb{R}^n)=\mathbb{R}^n$, which in turn is equivalent to $E^k\mathcal{D}+\mathcal{R} = \mathbb{R}^n$ because $H^k(\mathcal{D}+\mathcal{R}) = E^k\mathcal{D}+\mathcal{R}$ and thus we obtain the desired result.

Similarly, we can show that

$$(A\mathcal{D} + \mathcal{R} = \mathbb{R}^n) \iff (A^j\mathcal{D} + \mathcal{R} = \mathbb{R}^n, \forall j>0) \qquad (3.38)$$

and thus

$$(EA\mathcal{D} + \mathcal{R} = \mathbb{R}^n) \iff (E^kA^j\mathcal{D} + \mathcal{R} = \mathbb{R}^n, \forall j,k>0). \qquad (3.39)$$

Q.E.D.

Thus far, we have seen that the autonomous model (3.1), in general can give rise to time varying reachability and observability matrices. Also, in the case of BVLS, Krener has shown that the class of autonomous systems do not, in general, contain a minimal realization. However, he has shown that if we restrict our attention to systems with stationary weighing patterns then the class of autonomous systems does contain a minimal realization. This is one of our motivations for looking at stationary TPBVDS.

3.4- Stationary TPBVDS

Definition 3.6

The system (3.1) is stationary if and only if

a) $[E,V^i] = 0$ and $[A,V^f] = 0$

where $[\Sigma,\Lambda] = \Sigma\Lambda - \Lambda\Sigma$.

b) $\mathrm{Ker}(E^n) \subset \mathrm{Ker}(V^i)$

c) $\mathrm{Ker}(A^n) \subset \mathrm{Ker}(V^f)$

Condition a) is necessary to guarantee that the Green's function $G(i,j)$ depends only on the difference of the arguments i and j, i.e. $G(i,j) = G(i-j)$. The precise reasons for conditions b) and c) will become apparent later when we study the inward boundary process z', and are also illustrated in the following example.

Example 3.2

Consider the following system

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_{k+1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_k \qquad (3.40.1)$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_0 + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x_N = 0. \qquad (3.40.2)$$

This system is in standard form. The boundary matrices commute with E and A (A is zero here) so condition a) of Definition (3.6) is satisfied. It is easy to verify that condition b) is also verified but condition c) is not.

Now, consider the reachable space of system (3.40). $\mathcal{R}_i'$ is the space spanned by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for $i \neq 0$, and the space

spanned by $\begin{bmatrix}1\\1\end{bmatrix}$ for i=0.

Clearly, the reachable space of x_0 is not included in the reachable space of x_i for i far enough from the boundaries. In fact, this can happen any time the nilpotent Jordan blocks of E and A are not zeroed out by V^i and V^f respectively.

Using the fact that the boundary matrices commute with E and A, we obtain the following expression for the Green's function of the system

$$G(i-j)\overset{\Delta}{=}G(i,j)=\begin{cases} V^i A^{i-j-1} E^{N-i+j} & j<i \\ -V^f E^{j-i} A^{N-1-j+i} & i\leq j \end{cases} \qquad (3.41)$$

Notice that this expression for G(i,j) is much simpler than in the general case (3.9).

By adding conditions b) and c) it might appear that we are ignoring interesting systems that have time invariant Green's function solutions. But for any system that satisfies condition a) there exists an "almost identical" stationary system; "almost identical" means here that for any input sequence u_j, the states of the two systems x_k and x_k' are identical for $k\epsilon[n,N-n]$. In fact, by inspecting Equation (3.41), we see that the almost identical stationary system for a system that only satisfies condition a) can be

obtained by replacing V^i and V^f by $V^{i'}$ and $V^{f'}$ such that

$V^{i'}$ is the lowest rank matrix satisfying $V^{i'}E^n=V^iE^n$,
$V^{f'}$ is the lowest rank matrix satisfying $V^{f'}E^n=V^fE^n$.

Specifically what we are doing is to make sure that $V^{i'}$ and $V^{f'}$ annihilate all the nilpotent blocks of E and A. Since the nth and higher powers of these blocks are zero in any case, the effect of this modification is seen only near the boundaries 0 and N.

Note that the almost identical system for (3.40) has $V^f=0$. The reachability space of x_0 is then equal to 0, and we have a stationary system.

3.4.1-Reachability / Observability: Stationary Case

In the stationary case, we are able to completely characterize the inward boundary process z'(i,j).

In particular,

$$z'(i,j) = V^iE^{N-j+i}x_i+V^fA^{N-j+i}x_j = E^{N-j}A^iv + V^iE^{N-j}(A^{i-1}Bu_0+EA^{i-2}Bu_1+\ldots+E^{i-1}Bu_{i-1})-V^fA^i(A^{N-j-1}Bu_j+EA^{N-j-2}Bu_{j+1}+\ldots+E^{N-j-1}Bu_{N-1}). \qquad (3.42)$$

As before, we define reachability in terms of this process. We define the reachability matrix R'(i,j) as the

reachability matrix of $z'(i,j)$. It is clear that $R'(i,j)$ can be written as

$$R'(i,j)=[V^iE^{N-j}R(i-1) \quad V^fA^iR(N-j-1)] \quad (3.43)$$

where $R(.)$ is the strong reachability matrix defined previously. It should be clear at this point that $z'(i,i)$ is simply x_i and that $Im(R'(i,i))$ is the reachability space $\mathcal{R}'_i$ of x_i. What remains to be shown is that

$$\mathcal{R}'(i,j)=Im(R'(i,j)=\mathcal{R}' \quad (3.44)$$

(i.e. $\mathcal{R}'(i,j)$ does not depend on either i or j) for i and j far enough from the boundaries and that $\mathcal{R}'(i,j)$ is included in $\mathcal{R}'$ for i or j near the boundaries. This is the main difference between stationary and non-stationary systems. In the non-stationary case, the reachable space may rotate while, in the stationary case, the reachable space is constant far enough from the boundaries. It is also true in the stationary case that the reachable space near the boundaries is a subspace of the reachable space far from the boundaries.

<u>Theorem 3.8</u>

Let $\mathcal{R}'(i,j)$ be the reachability space for the stationary system (3.1). Then for i,j far enough from the boundaries i.e. for $i,j\epsilon[n,N-n]$,

$$\mathcal{R}'(i,j) = \mathcal{R}' = Im[V^iE^NR \quad V^fA^NR] \quad (3.45)$$

where R is the strong reachability matrix. In addition,

$\mathcal{R}'(i,j)$ for i or j near the boundaries, is included in $\mathcal{R}'$.

<u>Proof:</u>

Comparing (3.42) and (3.45) we see that we would like to show that

$$V^i E^{N-j}\mathcal{R} + V^f A^i \mathcal{R} = V^i E^N \mathcal{R} + V^f A^N \mathcal{R}. \tag{3.46}$$

But first, we will show that

$$E^K \mathcal{R} = E^L \mathcal{R} \qquad \text{for } K,L \geq n \tag{3.47}$$

We know that $E\mathcal{R} \subset \mathcal{R}$. If the dimension of $E\mathcal{R}$ is less than the dimension of $\mathcal{R}$ then consider $E^2\mathcal{R}$. Continue this process until for some $k<n$, $\mathrm{Dim}(E^{k+1}\mathcal{R})=\mathrm{Dim}(E^k\mathcal{R})$. This implies that $E^{k+1}\mathcal{R}=E^k\mathcal{R}$. And thus, for $K,L>k$ $E^K\mathcal{R}=E^L\mathcal{R}$.
Consequently, we have that $E^{N-j}\mathcal{R}=E^N\mathcal{R}$, which implies that $V^i E^{N-j}\mathcal{R}=V^i E^N\mathcal{R}$. By the same method we can show that $V^f A^i\mathcal{R}=V^f A^N\mathcal{R}$ and thus we have the desired result.

The second part of the theorem states that $\mathcal{R}'(i,j)$ for i and j near the boudaries is included in $\mathcal{R}'$. This can be easily proven by noting that because of conditions b) and c) of Definition (3.6), $V^i E^K\mathcal{R}=V^i\mathcal{R}$ and $V^f A^K\mathcal{R}=V^f\mathcal{R}$ (that is because $V^i E^k\mathcal{R} \subset V^i\mathcal{R}$ and $\mathrm{Ker}(V^i E^k) = \mathrm{Ker}\ (V^i)$). In fact, this allows us to obtain the following simple expression for $\mathcal{R}'$;

$$\mathcal{R}' = \mathrm{Im}[V^i R \qquad V^f R]. \tag{3.48}$$

In the stationary case we can simplify the expression for the unobservable space as well and show that

$$\mathcal{O}' = \mathrm{Ker}\begin{bmatrix} OV^i E^N \\ OV^f A^N \end{bmatrix} = \mathrm{Ker}\begin{bmatrix} OV^i \\ OV^f \end{bmatrix} \tag{3.49}$$

where O is the strong observability matrix.

<u>Corollary</u>

a) Stationary system (3.1) is reachable (observable) if and only if $\mathcal{R}'=\mathbb{R}^n$ ($\mathcal{O}'=0$).

b) Stationary system (3.1) is reachable (observable) if and only if $[sE-tA \vdots V^iB \vdots V^fB]$ ($\begin{bmatrix} sE-tA \\ CV^i \\ CV^f \end{bmatrix}$) has full rank for all $(s,t)\neq(0,0)$.

To prove statement b) we express $\mathcal{R}'$ in a form slightly different from the one given in Equation (3.48):

$$\mathcal{R}'=\mathrm{Im}[E^{n-1}V^iB \vdots AE^{n-2}V^iB \vdots \ldots \vdots A^{n-1}V^iB \vdots E^{n-1}V^fB \vdots \ldots \vdots A^{n-1}V^fB]$$
$$= \mathrm{Im}[E^{n-1}\{V^iB \vdots V^fB\} \vdots \ldots \vdots A^{n-1}\{V^iB \vdots V^fB\}]. \tag{3.50}$$

Here, we have used the fact that both E and A commute with V^i and V^f.

Now, we can obtain the desired result by following the argument presented to prove statement c) in Theorem (3.3). Note that in this case $\{V^iB \vdots V^fB\}$ plays the role of B.

For system (3.1) we have characterized reachability, strong reachability, observability, and strong observability. As in the case of causal discrete time linear

systems, it is possible to define concepts of controllability and constructibility. These concepts are defined in the same way as for the causal case. For example, system (3.1) is controllable if there exist controls u such that x, far enough from the boundaries, can be driven to zero regardless of the boundary value v. System (3.1) is strongly controllable if there exist controls u on the interval $[i-n, i+n-1]$ such that x_i can be driven to zero regardless of the values of controls off $[i-n, i+n-1]$ and v. For the sake of completeness, we will give conditions for these concepts without any details.

<u>Proposition 3.1</u>

a) System (3.1) is strongly controllable if $\mathrm{Im}((EA)^n) \subset \mathcal{R}$

b) System (3.1) is controllable if $\mathrm{Im}((EA)^n) \subset \mathcal{R}'$

c) System (3.1) is strongly constructible if $\mathcal{O} \subset \mathrm{Ker}((EA)^n)$

d) System (3.1) is constructible if $\mathcal{O}' \subset \mathrm{Ker}((EA)^n)$

In the next section, we will see why all four definitions of rechability and observability are useful for the study of System (3.1).

3.4.2-Minimality for Stationary TPBVDS

Recently, the minimality of boundary value linear systems (BVLS) with real-analytic coefficients has been studied extensively (see[2,5]). It turns out that a minimality condition for the stationary system (3.1) can be obtained directly by analogy with the case of BVLS.

Theorem 3.9

The stationary system (3.1) is minimal if and only if

a) $\mathcal{R}'=\mathbb{R}^n$

b) $\mathcal{O}'=0$

c) $\mathcal{O} \subset \mathcal{R}$.

Furthermore, any well-posed stationary TPBVDS can be reduced to a minimal stationary TPBVDS.

Theorem (3.9) basically states that the stationary system (3.1) is minimal if it is reachable, observable and any mode which is strongly unobservable is strongly reachable.

Proof:

First, we show that any stationary system (assumed to be in standard form) can be reduced to a system satisfying conditions a), b), and c) of Theorem (3.9). This will be done in three steps. The approach is similar in spirit to

the usual decomposition of causal systems into parts which are reachable and observable, reachable but unobservable, etc... However, here we need to worry about two different notions of reachability and observability.

<u>Step 1:</u>

Suppose that $\mathcal{R}'=\mathbb{R}^n$. Then let T be a matrix such that

$$z = T^{-1}x = \begin{bmatrix} z^1 \\ z^2 \end{bmatrix} \quad \text{where } z^1 \epsilon \mathcal{R}',\ z^2 \epsilon \mathcal{R}''$$

and where $\mathcal{R}' \oplus \mathcal{R}'' = \mathbb{R}^n$. This gives the following system for z (note that this change of coordinate does not affect the weighing pattern).

$$E'z_{k+1}=A'z_k+Bu_k \tag{3.51.1}$$

$$V^{i'}z_0+V^{f'}z_N=v' \tag{3.51.2}$$

$$y_k=C'z_k \tag{3.51.3}$$

where $C'=CT$, $E'=T^{-1}ET$, $A'=T^{-1}AT$, $B'=T^{-1}B$, $V^{i'}=T^{-1}V^iT$, $V^{f'}=T^{-1}V^fT$, and $v'=T^{-1}v$.

Note that system (3.51) is in standard form. Minimality is concerned with mapping from inputs u to outputs y, so we can assume that v and consequently v' are zero.

Now, we write the above matrices in matrix block format according to the decomposition of z.

$$A'=\begin{bmatrix} A'_{11} & A'_{12} \\ A'_{21} & A'_{22} \end{bmatrix},\quad E'=\begin{bmatrix} E'_{11} & E'_{12} \\ E'_{21} & E'_{22} \end{bmatrix},\quad V^{i'}=\begin{bmatrix} V^{i'}_{11} & V^{i'}_{12} \\ V^{i'}_{21} & V^{i'}_{22} \end{bmatrix},$$

$$V^{f'}=\begin{bmatrix} V^{f'}_{11} & V^{f'}_{12} \\ V^{f'}_{21} & V^{f'}_{22} \end{bmatrix},\quad C'=[C'_1 \quad C'_2],\quad B'=\begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix}.$$

Using the fact that $TA'=AT$ and that $\mathcal{R}'$ is A and E invariant we can show that

$$A'_{21}=E'_{21}=B_2'=0. \tag{3.52}$$

Moreover, since $y_k=[C'_1 \quad C'_2]\begin{bmatrix} z^1 \\ z^2 \end{bmatrix}$ and z^2 is identically zero (z^2 is in the unreachable space and $v'=0$), we can assume that $C'_2=0$.

Consequently we obtain the following expression for the weighing pattern of system (3.51):

$$W(i,j) = \begin{cases} C'_1 V^{i'}_{11} A'^{\,i-j-1}_{11} E'^{\,N-i+j}_{11} B'_1 & j<i \\ -C'_1 V^{f'}_{11} E'^{\,j-i}_{11} A'^{\,N-1-j+i}_{11} B'_1 & i\leq j \end{cases}. \tag{3.53}$$

But this weighing pattern is identical to the weighing pattern of the following reachable system with lower dimension.

$$E'_{11} z^1_{k+1} = A'_{11} z^1_k + B'_1 u_k \tag{3.54.1}$$

$$V^{i'}_{11} z^1_0 + V^{f'}_{11} z^1_N = v' \tag{3.54.2}$$

$$y_k = C'_1 z^1_k. \tag{3.54.3}$$

Thus, we have shown how to reduce an unreachable system into a reachable system of lower dimension.

<u>Step 2:</u>

Suppose that $\mathcal{O}'\neq 0$. Then, we can find a matrix T such that

$$z=T^{-1}x=\begin{bmatrix} z^1 \\ z^2 \end{bmatrix} \quad \text{where } z^1\epsilon\mathcal{O}'', \ z^2\epsilon\mathcal{O}' \quad \text{where } \mathcal{O}'\oplus\mathcal{O}'' = \mathbb{R}^n.$$

Following the same steps as in the previous case, it can be

shown that (3.52) can be reduced to an observable system of the form (3.54).

<u>Step 3:</u>

Suppose $\mathcal{O}$ is not included in $\mathcal{R}$; then there exists a subspace $\mathcal{Y}\neq 0$ such that $\mathcal{Y}\subset\mathcal{O}$ and $\mathcal{R}\oplus\mathcal{Y}=\mathcal{R}+\mathcal{O}$.

Let T be such that

$$z=T^{-1}x=\begin{bmatrix} z^1 \\ z^2 \end{bmatrix} \quad \text{where } z^1\epsilon\mathcal{Y}', \ z^2\epsilon\mathcal{Y} \quad \text{where } \mathcal{Y}\oplus\mathcal{Y}' = \mathbb{R}^n.$$

$\mathcal{Y}$ is not strongly reachable so $B_2'=0$ and $A_{21}'=E_{21}'=0$. Also, $\mathcal{Y}$ is not strongly observable so $C_2'=0$. Thus, (3.52) can be reduced to a system of the form (3.54) and satisfying condition c).

A similar 3 step procedure has been used by Gohberg and Kaashoek [5] to reduce continuous-time BVLS's to minimal dimension. This procedure is essentially equivalent to the procedure used by Krener [2] to reduce BVLS's to minimal size. His approach is to first do a 4 part Kalman decomposition with respect to the reachability and observability spaces, and then with respect to the strong reachability and strong observability matrices. In the first case, he shows that the unreachable part and the unobservable part do not contribute to the weighting pattern (just as we did in steps 1 and 2). In the second case, he shows that the strongly unreachable <u>and</u> strongly unobservable part does not contribute to the weighting

pattern (step 3).

Now, we need to show that two realizations of the same weighing pattern satisfying conditions a) thru c) of Theorem (3.9) must have the same dimension and consequently are minimal. Let System $\Sigma_1(E_1,A_1,B_1,C_1,V_1^i,V_1^f)$ and System $\Sigma_2(E_2,A_2,B_2,C_2,V_2^i,V_2^f)$ be two such systems with dimensions n_1 and n_2 respectively. In addition, without loss of generality, assume that $\alpha E_i+\beta A_i=I$ for $i=1,2$.

We know that the two realizations have the same weighing pattern, so that

$$C_1V_1^iA_1^{N-1-k}E_1^kB_1= C_2V_2^iA_2^{N-1-k}E_2^kB_2 \tag{3.55}$$

$$C_1V_1^fA_1^{N-1-k}E_1^kB_1= C_2V_2^fA_2^{N-1-k}E_2^kB_2 \tag{3.56}$$

Before proceeding with the rest of the proof we need to prove the following lemma.

<u>Lemma 3.1</u>

Let $\alpha E_i+\beta A_i = I$ for some non-zero α and β. Also suppose that for some matrices M_i, N_i,

$$M_1A_1^kE_1^{N-1-k}N_1 = M_2A_2^kE_2^{N-1-k}N_2 \qquad k\epsilon[0,N-1]. \tag{3.57}$$

Then, if $N>n_i$ (E_i and A_i are $n_i x n_i$),

$$M_1A_1^KE_1^LN_1 = M_2A_2^KE_2^LN_2 \qquad \text{for all } K,L \tag{3.58}$$

<u>Proof</u>

<u>Case 1: K+L < N-1</u>

In this case, we can write $E_i^KA_i^L = E_i^KA_i^L(\alpha E_i+\beta A_i)^{N-1-K-L}$

and expand the expression in the parentheses. In the resulting expression, the powers of E and A add up to N-1 and thus we can use Equation (3.57) to prove Equation (3.58).

<u>Case 2: K+L>N</u>

From Case 1, we know that

$$M_1 A_1^k N_1 = M_2 A_2^k N_2 \qquad k\epsilon[0,N-1]. \tag{3.59}$$

But,

$$\bar{M}_i \bar{A}_i^k \bar{N}_i = M_i A_i^k N_i \qquad k\geq 0 \ , \ i=1,2 \tag{3.60}$$

where $\Sigma_i(\bar{A}_i,\bar{N}_i,\bar{M}_i)$ is the reachable and observable part of the causal system $\Sigma_i(A_i,N_i,M_i)$ for i=1,2 (see for example []).

Using (3.59) and (3.60), we obtain the following:

$$\bar{M}_1 \bar{A}_1^k \bar{N}_1 = \bar{M}_2 \bar{A}_2^k \bar{N}_2 \qquad k\epsilon[0,N-1]. \tag{3.61}$$

Since $\Sigma_i(\bar{A}_i,\bar{N}_i,\bar{M}_i)$ (i=1,2) are reachable and observable, Σ_1 and Σ_2 must be related by a similarity transformation. Thus

$$\bar{M}_1 \bar{A}_1^k \bar{N}_1 = \bar{M}_2 \bar{A}_2^k \bar{N}_2 \qquad k\geq 0 \tag{3.62}$$

and so,

$$M_1 A_1^k N_1 = M_2 A_2^k N_2 \qquad k\geq 0 \quad . \tag{3.63}$$

By using (3.63) and the fact that $\alpha E_i+\beta A_i=I$, we can easily obtain (3.58). Q.E.D.

Now, using Lemma (3.1) and the fact that $V_1^i E_1^N + V_1^f A_1^N = V_2^i E_2^N + V_2^f A_2^N = I$, we find that

$$C_1 A_1^{N-1-k} E_1^k B_1 = C_2 A_2^{N-1-k} E_2^k B_2. \tag{3.64}$$

In the case of standard causal systems, the Hankel matrix is very useful in proving minimality. As it turns out, in our case, we can define two Hankel matrices: H' for the inward process z'(i,j) and H for the outward process z(i,j). In the first case, we drive the system with inputs off the interval [i,j] and observe the output on [i,j]. In the second case, we drive the system with inputs on [i,j] and observe the output off [i,j]. Note that Hankel matrices for the two systems must be identical because the two systems are identical from an input-output point of view. To proceed, let us choose i and j sufficiently far from the boundaries and from each other. For simplicity, let us assume that N-1 is divisible by 4 and choose i=(N-1)/4 and j=(3N-3)/4. We then have the following result.

$$H' = O_1^{on} R_1^{off} = O_2^{on} R_2^{off} \tag{3.65}$$

$$H = O_1^{off} R_1^{on} = O_2^{off} R_2^{on} \tag{3.66}$$

where

$$R_j^{on} = [E_j^{(N-1)/2} B_j \,\vdots\, A_j E_j^{(n-3)/2} B_j \,\vdots\, \ldots \,\vdots\, A_j^{(N-1)/2} B_j] \tag{3.67}$$

$$O_j^{on}=\begin{bmatrix} C_j E_j^{(N-1)/2} \\ \vdots \\ C_j A_j^{(N-1)/2} \end{bmatrix} \tag{3.68}$$

$$R_j^{off}=[V_j^i R_j^{on} \quad V_j^f R_j^{on}] \tag{3.69}$$

$$O_j^{off}=\begin{bmatrix} O_j^{on} V_j^i \\ O_j^{on} V_j^f \end{bmatrix}. \tag{3.70}$$

Also, we can write Equation (3.64) as

$$O_1^{on} R_1^{on} = O_2^{on} R_2^{on} \tag{3.71}$$

Since (N-1)/2 is larger than n_1 and n_2, we know that $Im(R_i^{on})=\mathcal{R}^i$, $Im(R_i^{off})=\mathcal{R}'^i$, $Ker(O_i^{on})=\mathcal{O}^i$ and $Ker(O_i^{off})=\mathcal{O}'^i$, where $\mathcal{R}^i$, $\mathcal{R}'^i$, $\mathcal{O}^i$ and $\mathcal{O}'^i$ are the strongly reachable, reachable, strongly unobservable and unobservable spaces of system Σ_i. Now to show that the two systems have the same dimension, first we show that R_1^{on} and R_2^{on} have the same rank r, and that O_1^{on} and O_2^{on} have the same rank o. This can be seen by inspecting Equations (3.65) and (3.66) because the "off" matrices have full rank. Then we use condition c) to obtain the ranks of the two sides of the Equation (3.71); we obtain

$$o + r - n_1 = o + r - n_2 \tag{3.72}$$

and thus

$$n_1 = n_2. \tag{3.73}$$

This completes the proof of Theorem (3.9).

In general, in contrast with the causal case, two minimal realizations of the same weighting pattern are not necessarily related by a similarity transformation (the same thing is true for continuous time BVLS, see [2]). However, it can be shown that two strongly reachable and strongly observable realizations of the same weighting pattern are related by a similarity transformation. The proof is somewhat similar to the proof of Lemma 3.1.

We have seen how our definitions of reachability and observability are used to find conditions on the minimality of the system. We also have shown how we can reduce any stationary system to a minimal system. It must be clear at this point why we cannot use the same argument to characterize minimality in the general case (that is because in the proof of theorem (3.9) we used the fact that $\mathcal{R}'$ and $\mathcal{O}'$ are time-invariant, otherwise the transformation T would become time-varying). So in general, we do not have any systematic way of reducing a non-stationary TPBVDS into a minimal system, or even to test whether a non-stationary system is minimal or not. The only thing we know is that if a TPBVDS satisfies the conditions of Theorem (3.9) it is minimal. These conditions, however, are difficult to verify in the non-stationary case.

3.4.3-Stability for Stationary TPBVDS

Until now we have been considering N as a finite constant. Our system has been defined on a bounded segment of the integers. In this section, we study the stationary TPBVDS over an infinite interval.

In the standard causal case, we know that a system is stable if the effect of the initial condition approachs zero as we get further away from the origin. In the case of the TPBVDS's that we have been studied so far, however, we can not get arbitrarily far away from the boundaries to study stability; that is because our domain is bounded. We could, however, extend the interval over which the system is defined. We already know how to reduce this interval. In Appendix A, when studying the inward boundary process, we showed that we could move in the boudary matrices. We showed that by moving in the boundary matrices, we obtain a new system defined over a smaller interval which has a weighting pattern identical to the weighting pattern of the original system restricted to its domain of definition. The new system has the same dynamics as the original system but the boundary matrices are different. The new boundary matrices are in fact

$$W^i_{jk}=V^iE^{N-k+j} \quad \text{and} \quad W^f_{jk}=V^fA^{N-k+j}. \tag{3.74}$$

It is possible to use the same idea to move out the boundaries. The system obtained by moving out the boudaries must have the same weighting pattern as the original system, which means that if we move back its boundaries to the original locations, we should get the original system back (note that we can move out the boundaries <u>only if</u> the system is stationary i.e. it satisfies all three conditions of Definiton (3.9)). Thus we want to find W^i_{jk} and W^f_{jk} for $j<0$ and $k>N$ such that

$$W^i_{jk}E^{k-N-j}=V^i \quad \text{and} \quad W^f_{jk}A^{k-N-j}=V^f. \tag{3.75}$$

Clearly, W^i_{jk} and W^f_{jk} always exist because E is invertible in the range of V^i and A is invertible in the range of V^f (conditions b) and c) of Definition (3.6)). In fact, they can be computed as follows

$$W^i_{jk}=V^i(\tilde{E}^{k-N-j})^{-1} \quad \text{and} \quad W^f_{jk}=V^f(\tilde{A}^{k-N-j})^{-1} \tag{3.76}$$

where $\tilde{E}$ and $\tilde{A}$ are matrices having the same eigenstructure as E and A but with their zero eigenvalues replaced by one (or any other non-zero scalar). Notice that if either E or A is not invertible, W^i_{jk} and W^f_{jk} are not unique.

We have shown that it is possible to extend the domain of definiton of any stationary system. Another way of looking at it is that any stationary system can be obtained by moving in the boundaries of another stationary system defined over a larger interval. An interesting question that relates to stability is under what conditions we can push

back the boudaries to $\mp\infty$ in a meaningful way so that we can think of our system as a part of a system defined over an infinite interval.

In fact, there are two situations where we might want to study stability. The first situation, as mentioned above, corresponds to the case when the system is actually defined over an infinite interval. the second situation arises when the system is defined over a finite interval with boundary conditions which are physical constraints of the problem, and when we would like to study the effects of increasing the domain of the system (same dynamics, same boundary constraints). For example, consider a system that describes the heat distribution on a ring. This system has boundary conditions (which are periodic) independent of the size of the ring. In this case, we might want to study the effect of increasing the size of the ring.

In the first case, we would like to find systems such that if we move out the boundaries while keeping the weighting pattern the same, the system stays stable. We show that this happens only when the system is separable and both the forward and the backward subsystems are stable. To show this, we need to decompose the modes of our system. Consider the following block diagonalization:

$$E = \begin{bmatrix} I & & \\ & A_b & \\ & & I \end{bmatrix}, \quad A = \begin{bmatrix} A_f & & \\ & I & \\ & & U \end{bmatrix} \tag{3.77}$$

where A_f and A_b have eigenvalues inside the unit circle and U has eigenvalues on the unit circle. This is a trivial generalization of the decompositon in [24].

Thus, the first block is stable in the forward direction, the second block is backward stable and the last block is marginally stable.

Now, we would like to show that this block diagonalization also block diagonalizes V^i and V^f. But first we need to prove the following theorem.

<u>Theorem 3.10</u>

Let

$$S= \begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix} \text{ and } T= \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{3.78}$$

If ST=TS and no eigenvalue of E equals any eigenvalue of F then B=C=0.

<u>Proof:</u>

ST=TS implies that EB=BF and thus $E^kB=BF^k$. This implies in turn that

$$p(E)B=Bp(F) \qquad \text{for any polynomial } p(\cdot) \tag{3.79}$$

Let u_i be any generalized eigenvector of F. Then there exists k such that

$$(\lambda I-F)^k u_i=0 \tag{3.80}$$

so that for $p(x)=(\lambda-x)^k$,

$$p(E)Bu_i=0 \ . \tag{3.81}$$

Now, if $Bu_i \neq 0$, then λ must be an eigenvalue of E which is a contradiction because λ is an eigenvalue of F. If $Bu_i=0$ for all u_i then B must be zero because the u_i's span the whole space, which is the desired result. Q.E.D.

Since V^i and A commute and V^f and E commute, we can use Theorem (3.10) to show that all the off diagonal blocks of V^i and V^f are zero except maybe the (1,3) and (3,1) blocks of V^i and the (3,2) and (2,3) blocks of V^f. Then the fact that $V^iE^N+V^fA^N=I$ implies that the (1,3) and (3,1) blocks of V^i and the (3,2) and (2,3) blocks of V^f are zero, so that V^i and V^f have the following structure:

$$V^i=\begin{bmatrix} V^i_1 & & \\ & V^i_2 & \\ & & V^i_3 \end{bmatrix}, \quad V^f=\begin{bmatrix} V^f_1 & & \\ & V^f_2 & \\ & & V^f_3 \end{bmatrix}. \tag{3.82}$$

Clearly, the boundary matrices W are also block-diagonalized (see Equation (3.76)).

$$W^i=\begin{bmatrix} W^i_1 & & \\ & W^i_2 & \\ & & W^i_3 \end{bmatrix}, \quad W^f=\begin{bmatrix} W^f_1 & & \\ & W^f_2 & \\ & & W^f_3 \end{bmatrix} \tag{3.83}$$

where for simplicity we have dropped the subscript jk.

Notice that as j and k go to minus and plus infinity,

$(W_2^i)_{jk}$ and $(W_1^f)_{jk}$ grow unbounded because $(W_2^i)_{jk}=V_2^i(A_b^{k-j})^{-1}$ and $(W_1^f)_{jk}=V_1^f(A_f^{k-j})^{-1}$ unless V_2^i and V_1^f are zero (for the moment assume that A_b and A_f are invertible, then clearly there inverses are unstable). Also notice that W_3^i does not change and stays equal to V_3^i.

Now consider the outward process $z(j,k)=\begin{bmatrix} z_1(j,k) \\ z_2(j,k) \\ z_3(j,k) \end{bmatrix}$. It is clear (see Appendix B) that, z_1 and z_2 are stricly stable processes whereas z_3 is not. Also, notice that,

$$(x_1)_j=(W_1^f)_{jk}z_1(j,k) \tag{3.84.1}$$

$$(x_2)_k=(W_2^i)_{jk}z_2(j,k) \tag{3.84.2}$$

$$(x_3)_j=(W_3^i)_{jk}z_3(j,k). \tag{3.84.3}$$

Clearly, since if either W_1^f, W_2^i or z_3 grows unbounded x cannot stay bounded, thus the system is stable if and only if

$$V_1^f=0 \tag{3.85.1}$$

$$V_2^i=0 \tag{3.85.2}$$

and no eigenmode is on the unit circle.

This of course is the result that we wanted to show. In the case where A_f and A_b are not invertible, the nilpotent part of A_f and the nilpotent part of A_b must be strictly causal and anticausal respectively. This can be easily deduced from conditions b) and c) of Definition (3.6). Since nilpotent systems are stable, we obtain the desired result.

In the case where we extend the domain of the system while keeping the same boundary conditions, a stable and separable system would clearly be considered as stable, but separability is not a necessary condition. What we would like to find is a necessary and sufficient condition for stability in this case.

Since N is now a parameter, we relax the standard form condition and suppose that V^i and V^f are in standard form for some $N_0 > n$. What we are interested in is the effect of the boundary value v on some x in the middle of the interval as N goes to infinity.

<u>Definition 3.8</u>

The stationary TPBVDS (3.1) is strictly stable if as N goes to infinity the effect of the boundary value v on any x near the mid-section goes to zero, i.e.if

$$\lim_{N\to\infty} (V^i E^N + V^f A^N)^{-1} E^{N/2} A^{N/2} = 0 \quad . \qquad (3.86)$$

Now assume that we have block-diagonalized our system as discussed previously. Then, we obtain the following stability condition.

<u>Theorem 3.11</u>

The stationary TPBVDS in block diagonal form (3.77),(3.82) is strictly stable if and only if

a) V_1^i is non-singular

b) V_2^f is non-singular

c) the system has no eigenvalue on the unit circle.

This result is shown by replacing E, A and the boundary matrices in Equation (3.86) by their block diagonal forms. The result is a block diagonal matrix. Now consider the first block

$$\lim_{N\to\infty} (V_1^i+V_1^f A_f^N)^{-1} A_f^{N/2}=0. \tag{3.87}$$

This of course means that V_1^i must be non-singular. Similarly, we can show b) and c).

So, to test the stability of a stationary TPBVDS, we need to transform the system into the block diagonal form described above and then apply Theorem (3.11).

IV-THE STATIONARY TPBVDS: The Stochastic Problem

In this chapter we study the stationary TPBVDS

$$Ex_{k+1} = Ax_k + Bu_k \tag{4.1.1}$$

$$V^i x_0 + V^f x_N = v \tag{4.1.2}$$

where u is a zero-mean Gaussian random process with covariance matrix $I\delta_{ij}$, and where v is a zero-mean Gaussian random vector, with covariance Q, which is independent of u_k for $k\epsilon[0,N-1]$. The dimensions of all matrices are the same as in (3.1). In addition system (4.1) is assumed to be stationary, in standard form and reachable.

The non-descriptor continuous-time version of system (4.1) was introduced by Krener [1], who called these systems stationary boundary value linear systems. In [25] Krener examines the relation between this class of systems and reciprocal processes, and in particular the problem of realizing reciprocal processes with stationary boundary value linear systems driven by white noise. This problem has not been completely resolved in the continuous time case nor in the discrete time case.

4.1-Introduction:

As in the case of causal systems, the deterministic stationarity of system (4.1), i.e. the fact that the Green's function G(i,j) depends only on the difference of the arguments i and j, does not necessarily imply the existence of a stationary covariance (stochastic stationarity). In fact, we know that a causal system has a stationary covariance only if the initial covariance matrix satisfies the Lyapunov equation, which in turn requires that the system be stable. In the general case of TPBVDS, as well, some conditions must be met in order for the system to have a stationary covariance. These conditions will be derived in the next section. In this section we introduce some preliminary results and establish the notation.

<u>Definition 4.1</u>

The stationary TPBVDS (4.1) is stochastically stationary if and only if

$$\overline{x_{k+i}x_k'} = R_{k+i,k} = R_i \ ,$$

This, of course, is the usual definition of stochastic stationarity.

It should be clear that if (4.1) is stochastically stationary, the <u>variance matrix</u> $P_k \overset{\Delta}{=} R_{k,k}$ of x_k is constant. So, our first step at this point will be to characterize P_k

as much as possible. By multiplying both sides of Equations (4.1.1) and (4.1.2) by their transpose and using the Green's function solution derived in the previous section and taking the expected value, we can show that P_k satisfies the following TPBVDS

$$EP_{k+1}E'-AP_kA' = (V^iE^N)BB'(V^iE^N)'-(V^fA^N)BB'(V^fA^N)' \quad (4.2.1)$$
$$V^iP_0V^{i\prime}-V^fP_NV^{f\prime} = (V^iE^N)Q(V^iE^N)'-(V^fA^N)Q(V^fA^N)' \ . \quad (4.2.2)$$

Now the question is under what conditions does (4.2) completely characterizes P_k. Fortunately, we have developed the necessary tools to respond to this question in the previous chapter. In fact, all we need to do is to apply the well-posedness test for a TPBVDS; but first we need to transform (4.2) into a form similar to system (4.1). This can be done as follows:

$$(E\otimes E)\overline{P}_{k+1}-(A\otimes A)\overline{P}_k=(V^iE^N\otimes V^iE^N)\overline{BB'}-(V^fA^N\otimes V^fA^N)\overline{BB'} \quad (4.3.1)$$
$$(V^i\otimes V^i)\overline{P}_0-(V^f\otimes V^f)\overline{P}_N=(V^iE^N\otimes V^iE^N)\overline{Q}-(V^fA^N\otimes V^fA^N)\overline{Q} \quad (4.3.2)$$

where $\otimes$ represents the Kroneker product and $\overline{P}$, $\overline{Q}$ and $\overline{BB'}$ are vectors obtained from the entries of matrices P, Q and BB' by lexicographic ordering. Note that the right hand sides of the above equations are irrelevant as far as well-posedness

is concerned.

The well-posedness condition in this case reduces to the invertibility of the matrix

$$(V^i \otimes V^i)(E \otimes E)^N - (V^f \otimes V^f)(A \otimes A)^N \ .$$

Thus we obtain the following result.

<u>Theorem 4.1</u>

Equations (4.2.1) and (4.2.2) characterize P_k completely if and only if $\lambda_i \lambda_j \neq \mu_i \mu_j$ for all i and j where λ_i and μ_i are the eigenvalues of matrices $V^i E^N$ and $V^f A^N$ corresponding to eigenvector v_i.

Note that in the causal case the μ_i's are all zero and the λ_i's are all one, so that P_k is always well defined. This is expected because in the causal case we have the initial condition $P_0=Q$, and a forward recursion for P_k. The condition in Theorem (4.1) can also be written as $\lambda_i \neq \mu_j$ because $\lambda_i + \mu_i = 1$ (standard form).

Theorem (4.1) basically states that except under very special circumstances the variance matrix P_k can be uniquely calculated from Equations (4.2.1) and (4.2.2). Now, it is clear that if our system is stochastically stationary, the stationary variance matrix P must satisfy the two algebraic matrix equations obtained by setting $P_k = P_{k+1} = P$ and $P_0 = P_N$ in (4.2.1) and (4.2.2) respectively. In fact, by analogy with

the causal case, one might incorrectly deduce that if (4.2.1) has a positive-definite solution P then the system is stochastically stationary. Unfortunately this is not the case. We derive the correct condition in the next section.

4.2-Stochastically Stationary TPBVDS

In this section we present conditions under which system (4.1) is stochastically stationary and, in addition, obtain a complete characterization of the covariance matrices.

Theorem 4.2

System (4.1) has a constant variance matrix if Q satisfies the following equation

$$EQE'-AQA' = V^iBB'V^{i\prime}-V^fBB'V^{f\prime} \tag{4.4}$$

and in that case the variance matrix P satisfies the following equation

$$EPE'-APA' = (V^iE^N)BB'(V^iE^N)'-(V^fA^N)BB'(V^fA^N)'. \tag{4.5}$$

Proof:

We have to show that $P_i=P_{i+1}$ for all i if Q satisfies Equation (4.4). Let,

$$\Pi_i = \sum_{j=0}^{i} A^{i-j}E^jBB'E'^jA'^{i-j} .$$

Then by using the Green's function solution we can show that

$$P_i=A^iE^{N-i}QE'^{N-i}A'^i+(V^iE^{N-i})\Pi_{i-1}(V^iE^{N-i})'+(V^fA^i)\Pi_{N-1-i}(V^fA^i)'$$

Now, writing P_{i+1} in similar fashion and using the fact that $\Pi_i = A\Pi_{i-1}A'+E^iBB'E'^i$ we can show that $P_i=P_{i+1}$ is equivalent to having

$$A^iE^{N-1-i}(EQE'-AQA')E'^{N-1-1}A'^i =$$

$$A^iE^{N-1-i}(V^iBB'V^{i\prime}-V^fBB'V^{f\prime})E'^{N-1-i}A'^i \tag{4.6}$$

Thus,

$$A^iE^{N-1-i}(EQE'-AQA'-V^iBB'V^{i\prime}+V^fBB'V^{f\prime})E'^{N-1-i}A'^i=0. \tag{4.7}$$

Clearly (4.7) is implied by (4.4). Q.E.D

Notice that (4.4) and (4.7) are in fact equivalent if either E or A is invertible. Consequently, if either E or A is invertible, system (4.1) has constant variance if and only if Q satisfies (4.4).

<u>Example 4.1</u>

Consider the following system

$$\begin{bmatrix}1&0&0\\0&0&1\\0&0&0\end{bmatrix}x_{k+1}=\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}x_k+u_k \tag{4.8.1}$$

$$\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}x_0+\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}x_N=v. \tag{4.8.2}$$

where covariance of v is given by

$$Q=\begin{bmatrix}1 & 0 & 1\\ 0 & 2 & 0\\ 1 & 0 & 1\end{bmatrix}. \tag{4.9}$$

System (4.8) is in standard form and stationary.

It is easy to check that Q satisfies (4.7) but not (4.4). Thus system (4.8) has a constant variance matrix P which can be easily computed. In fact,

$$P=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{4.10}$$

However, system (4.8) is not stochastically stationary because

$$R_{0,N-1} = \overline{x_0 x_{N-1}'} = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{4.11.1}$$

is different from

$$R_{1,N} = \overline{x_1 x_N'} = 0. \tag{4.11.2}$$

Notice also that Equation (4.5) completely characterizes P as long as $\sigma_i\sigma_j \neq 1$ for all i and j, where σ_i is an eigen-mode of the system (i.e. $(E\sigma_i - A)$ is singular). Equation (4.5) clearly reduces to the well-known Lyapunov equation (see for example [23]) in the causal case. In the causal case, a reachable system has a constant variance matrix if and only if A is stable. This means that the eigen-modes σ_i of the system are all less than 1. Clearly in that case $\sigma_i\sigma_j \neq 1$ and we get that Equation (4.5) completely characterizes the variance matrix P.

In cases where (4.5) does not characterize P completely, it is still possible to compute P. In fact, P can be computed from (4.5) by using perturbation methods. For example, we can replace A by $(A+\epsilon I)$ and compute the corresponding P_ϵ; then we can obtain P by letting ϵ go to zero in P_ϵ. The reason this method can always be used to compute P is that, by writing out the expression for P_ϵ using the Green's function solution of the system obtained by replacing A with $A+\epsilon I$, it can be seen that the entries of P_ϵ are rational functions of ϵ, analytic at $\epsilon=0$. This of course means that P_ϵ is a continuous function of ϵ in some neighborhood of $\epsilon=0$.

<u>Example 4.2</u>

Consider the following anticyclic system

$$x_{k+1}=x_k+bu_k \tag{4.12.1}$$

$$(1/2)x_0+(1/2)x_N=0 \tag{4.12.2}$$

where u_k is a white sequence with variance 1. System (4.12) is in standard form.

Clearly, system (4.12) has a constant variance because Q=0 satisfies (4.4). However, to determine the constant variance, Equation (4.5) cannot be used directly because both sides of it are zero in this case. Thus, we have to use a perturbation method.

Consider the following system

$$x_{k+1}=(1+\epsilon)x_k+bu_k \tag{4.13.1}$$

$$\gamma x_0+\gamma x_N=0 \tag{4.13.2}$$

where $\gamma=(1+(1+\epsilon)^N)^{-1}$. System (4.13) is in standard form for all ϵ. For $\epsilon=0$, system (4.13) is equivalent to system (4.12). To compute the variance P of system (4.12), we first compute the variance P_ϵ of system (4.13). P is then equal to the limit of P_ϵ as ϵ goes to zero.

Using Equation (4.5) for system (4.13) we get

$$-2\epsilon P_\epsilon=-2\epsilon Nb^2/4 + o(\epsilon^2) \tag{4.14}$$

so that,

$$P=\lim_{\epsilon->0} P_\epsilon=Nb^2/4. \tag{4.15}$$

This result can also be obtained directly by using the Green's function solution of system (4.12).

Note that, the constant variance matrix P and in general the variance matrices P_i for $i\epsilon[n,N-n]$ are positive definite if the system is reachable. This can easily be seen by writing P_i as follows

$$P_i=A^iE^{N-i}QE'^{N-i}A'^i+R'(i)R'(i)' \tag{4.16}$$

where R' is the reachability matrix defined in Chapter III, i.e.,

$$R'(i) = [V^iE^{N-i}R(i) \quad V^fA^iR(N-i)]. \tag{4.17}$$

Since the system is reachable, R'(i) has full rank for $i\epsilon[n,N-n]$. Which means that R'(i)R'(i)' is positive definite

and so P_i is positive definite.

Theorem 4.3

System (4.1) is stochastically stationary if and only if Q satisfies Equation (4.4).

Proof:

We prove the result for two cases.

Case 1: either E or A invertible

We have to show that Q satisfies (4.4) if and only if $R_{i,j}$ depends only on i-j. Note first that

$$ER_{i+1,k}=\overline{Ex_{i+1}x_k}'=\overline{(Ax_i+Bu_i)x_k}'=AR_{i,k}+B\overline{u_ix_k}' \qquad (4.18)$$

Using (4.18) and the Green's function solution to compute $\overline{u_ix_k}$, we obtain the following equation:

$$ER_{i+1,k}-AR_{i,k} = -BB'[V^fE^{i-k}A^{N-1-(i-k)}]' \qquad i\geq k. \qquad (4.19.1)$$

Similarly, we can show that

$$R_{i,k+1}E'-R_{i,k}A' = V^iA^{i-k-1}E^{N-i+k}BB' \qquad k<i\ . \qquad (4.19.2)$$

Let us first show the "if" part of the theorem. Since Q satisfies (4.4), $R_{k,k}=P$ is constant (Theorem (4.2)). We want to show that $R_{k+s,k}$ does not depend on k. Using Equations (4.19.1) and (4.19.2) and the fact that $R_{k,k}=P$ we obtain the following equations:

$$ER_{k+1,k}-AP = -BB'[V^f A^{N-1}]' \tag{4.20.1}$$

$$PE'-R_{k+1,k}A' = V^i E^{N-1} BB' \quad . \tag{4.20.2}$$

More generally, we have

$$E^s R_{k+s,k}=A^s P -\sum_{j=0}^{s-1} A^{s-j-1} BB'[V^f A^{N-1-j} E^j]' \tag{4.21.1}$$

$$R_{k+s,k}A'^{s}=PE'^{s} -\sum_{j=0}^{s-1} [V^i E^{N-1-j} A^j] BB' E'^{s-j-1} . \tag{4.21.2}$$

Since either E or A is invertible, one of the Equations (4.21.1) or (4.21.2) completely characterizes $R_{k+s,k}$ and clearly this matrix does not depend on k. So, $R_{k+s,k}=R_s$.

Thus far we have shown that if Q satisfies (4.4) then $R_{i,k}=R_{i-k}$. But if $R_{i,k}=R_{i-k}$ then $R_{k,k}$=constant, which, since either E or A is invertible, implies that Q satisfies (4.4). Thus Q satisfies (4.4) if and only if $R_{i,k}=R_{i-k}$ which is the desired result.

<u>Case 2: E and A singular</u>

In Case 1, we took advantage of the fact that when either E or A is invertible, Q satisfies (4.4) if and only if the variance matrix is constant, which in turn was shown to be equivalent to stochastic stationarity. When E and A are both singular, however, we could have a constant variance matrix even when the system is not stochastically stationary or when Q does not satisfy (4.4). In fact, we know that the system has a constant variance P if and only

if Q satisfies (4.7) (which as we said before is equivalent to (4.4) if either E or A is invertible).

First, we block diagonalize our matrices in the same manner as in Section (3.4.3) of Chapter III. This is accomplished by a constant change of coordinate which clearly does not affect stationarity or the validity of Equation (4.4).

Thus, system (4.1) can be written as

$$\begin{bmatrix} E_1 & & \\ & E_2 & \\ & & N_e \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_{k+1} = \begin{bmatrix} N_a & & \\ & A_2 & \\ & & A_3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_k + \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} u_k \qquad (4.22.1)$$

$$\begin{bmatrix} V_1^i & & \\ & V_2^i & \\ & & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_0 + \begin{bmatrix} 0 & & \\ & V_2^f & \\ & & V_3^f \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_N = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \qquad (4.22.2)$$

where N_e and N_a are nilpotent matrices and E_1, E_2, A_2, and A_3 are invertible. Block zeros in V^i and V^f are due to the fact that the null space of E^n must be included in the null space of V^i and that the null space of A^n must be included in the null space of V^f.

We can simplify (4.22) by noting that subsystems 1 and 3 are simply causal and anticausal nilpotent systems. That is because E_1 and N_a have the same Jordan structure (remember that $\alpha E+\beta A=I$) and thus premultiplying subsystem 1 by E_1^{-1} would result in a nilpotent system (because $E_1^{-1}N_a$ is

nilpotent); similarly we can show that subsystem 3 is nilpotent. So, without loss of generality we can assume that (4.22) has the following form

$$\begin{bmatrix} I & & \\ & E_2 & \\ & & N_e \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_{k+1} = \begin{bmatrix} N_a & & \\ & A_2 & \\ & & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_k + \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} u_k \tag{4.23.1}$$

$$\begin{bmatrix} I & & \\ & V_2^i & \\ & & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_0 + \begin{bmatrix} 0 & & \\ & V_2^f & \\ & & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}_N = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}. \tag{4.23.2}$$

(V_1^i and V_3^f are equal to I because the boundary matrices are in standard form).

Now, suppose that the boundary covariance

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \tag{4.24}$$

satisfies (4.4). In that case we know that (4.23) has a constant variance matrix P. Let

$$R_{i,k} = \begin{bmatrix} R_{i,k}^{11} & R_{i,k}^{12} & R_{i,k}^{13} \\ R_{i,k}^{21} & R_{i,k}^{22} & R_{i,k}^{23} \\ R_{i,k}^{31} & R_{i,k}^{32} & R_{i,k}^{33} \end{bmatrix} \quad \text{with } i \geq k \tag{4.25}$$

be the covariance matrix of (4.23). Using equations (4.21.1) and (4.21.2), as in the previous case, it is easy to see that, except for $(m,n)=(3,1)$, $R_{i,k}^{mn}=R_{i-k}^{mn}$. This is because (4.21.1) and (4.21.2) characterize all the entries of $R_{k+1,k}$ except (3,1). Notice that this entry is simply the cross

correlation between subsystem 1 (the nilpotent causal system) and subsystem 3 (the nilpotent anticausal system). As shown in Lemma (4.1) to follow, these two subsystems have stationary cross correlation if and only if

$$Q_{13}N_e'-N_aQ_{13}=0. \tag{4.26}$$

But (4.26) is implied by (4.4) (in fact (4.26) is the (1,3) entry of (4.4)). Thus, if Q satisfies (4.4) then $R_{i,k}=R_{i-k}$. On the other hand, if $R_{i,k}=R_{i-k}$ then $R_{k,k}$ must be constant which means that Q must satisfy (4.7); Q must also satisfy (4.26). But the condition imposed on Q by (4.4) differs from (4.7) only by the fact that (4.7) disregards the (1,3) and (3,1) entries of (4.4). These entries are simply Equation (4.26) and its transpose. Thus stochastic stationarity implies that Q must satisfy (4.4). So Q satisfies (4.4) if and only if system (4.23) is stochastically stationary. Q.E.D.

<u>Lemma 4.1</u>

The system

$$\begin{bmatrix} I & \\ & N_2 \end{bmatrix}\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_{k+1}=\begin{bmatrix} N_1 & I \end{bmatrix}\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_k+\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}u_k \tag{4.27.1}$$

$$\begin{bmatrix} I & \\ & 0 \end{bmatrix}\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_0+\begin{bmatrix} 0 & \\ & I \end{bmatrix}\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_N=\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{4.27.2}$$

where N_1 and N_2 are nilpotent is stochastically stationary if and only if

$$Q_{11}-N_1Q_{11}N_1'=B_1B_1' \tag{4.28.1}$$

$$Q_{22}-N_2Q_{22}N_2'=B_2B_2'. \tag{4.28.2}$$

and

$$Q_{12}N_2'=N_1Q_{12}. \tag{4.28.3}$$

<u>Proof:</u>

We would like to find a necessary and sufficient condition for the stochastic stationarity of (4.27). Clearly, a necessary condition is that x^1 and x^2 be individually stationary.

Let $Q=\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$ and $R_{i,k}=\begin{bmatrix} R^{11}_{i,k} & R^{12}_{i,k} \\ R^{21}_{i,k} & R^{22}_{i,k} \end{bmatrix}$. Then, clearly $R^{11}_{i,k}=R^{11}_{i-k}$ and $R^{22}_{i,k}=R^{22}_{i-k}$ if and only if (4.28.1) and (4.28.2) are satisfied. Equations (4.28.1) and (4.28.2) are the usual Lyapunov equations for subsystems 1 and 2. Note that regardless of the value of Q_{12} if Q_{11} and Q_{22} satisfy (4.28), system (4.27) has a constant variance matrix

$$P=\begin{bmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{bmatrix}. \tag{4.29}$$

But, existence of a constant variance does not necessarily imply stochastic stationarity (Equations (4.28.1) and (4.28.2) are essentially equivalent to equation (4.7)). The fact that Equations (4.28.1) and (4.28.2) are satisfied implies that x^1 and x^2 are individually stationary. Thus

$$R^{11}_{i,k}=R^{11}_{i-k} \quad \text{and} \quad R^{22}_{i,k}=R^{22}_{i-k}. \tag{4.30}$$

It is also easy to check that for $i<k$

$$R^{21}_{i,k}=\overline{x^2_i x^1_k}$$
$$=B_2B_1'N_1'^{k-i-1}+N_2B_2B_1'N_1'^{k-i-2}+\ldots+N_2^{k-i-1}B_2B_1'. \quad (4.31)$$

The right hand side of (4.31) is only a function of k-i so,

$$R^{21}_{i,k}=R^{21}_{i-k}. \quad (4.32)$$

Thus, the only part of $R_{i,k}$ that may not be stationary is $R^{12}_{i,k}$ (for $i<k$). But, clearly

$$R^{12}_{i,k}=0 \qquad \text{for either } i \text{ or } k\epsilon[n,N-n] \quad (4.33)$$

since x^1_i and x^2_k are functions of nonoverlaping intervals of the white input noise and i and k are far enough from the boundaries so that the boundary condition terms have disappeared.

When $i\epsilon[0,n]$ and $k\epsilon[N-n,N]$, $R^{12}_{i,k}$ may not be zero if v_1 and v_2 are correlated (i.e. $Q_{12}\neq 0$). We want $R^{12}_{i,k}$ to be stationary which means that

$$R^{12}_{0,N-1}=R^{12}_{1,N} \quad (4.34.1)$$
$$R^{12}_{0,N-2}=R^{12}_{1,N-1}=R^{12}_{2,N} \quad (4.34.2)$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

or,

$$Q_{12}N_2'=N_1Q_{12} \quad (4.35.1)$$
$$Q_{12}N_2'^2=N_1Q_{12}N_2'=N_1^2Q_{12} \quad (4.35.2)$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

It is clear that if (4.35.1) is satisfied, then all

Equations (4.35) are satisfied. Thus a necessary and sufficient condition for $R^{12}_{i,k}$ to be stationary is (4.28.3).

Thus system (4.27) is stochastically stationary if and only if Q satisfies (4.28.1), (4.28.2) and (4.28.3) which is the desired result. Also notice that, these three equations are equivalent to equation (4.4). Consequently, we obtain the expected result: (4.27) is stochastically stationary if and only if Q satisfies (4.4).

Let R_j denote the covariance matrix $R_{k+j,k}$ of a stochastically stationary system.

<u>Theorem 4.4</u>

If Q satisfies Equation (4.4) then the system is stochastically stationary and R_j can be computed using the following second order TPBVDS

$$ER_{j+1}E'+AR_{j+1}A' = AR_jE'+ER_{j+2}A'. \qquad (4.36)$$

with appropriate boundary conditions (we discuss boundary conditions later).

This equation is analogous to the second order differential equation obtained by Krener [25] for the covariance of the continuous-time stationary two point boundary value system.

<u>Proof:</u>

Equation (4.36) follows from (4.19.1). What remains to be shown is that (4.36) is well-posed. First, we need to prove the following Lemma.

<u>Lemma 4.2</u>

The mth order descriptor system

$$Q_m x_{j+m} + Q_{m-1} x_{j+m-1} + \dots + Q_0 x_j = Bu_j \tag{4.37}$$

is well-posed for some appropriate boundary conditions if and only if there exists z such that the polynomial matrix $[Q_m z^m + Q_{m-1} z^{m-1} + \dots + Q_0]$ is invertible.

<u>Proof:</u>

Using state augmentation, we can rewrite (4.37) as follows

$$\begin{bmatrix} I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & Q_m \end{bmatrix} \bar{x}_{j+1} = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ -Q_0 & & \cdots\cdots & & -Q_{m-1} \end{bmatrix} \bar{x}_j + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ Bu_j \end{bmatrix}. \tag{4.38}$$

Now, all we need to do is to see under what condition the two matrices in (4.38) form a regular pencil. It is easy to show that this happens if and only if the condition in Lemma (4.1) is satisfied. Q.E.D.

Now using Lemma (4.1) well posedness of (4.36) becomes equivalent to invertibility of the polynomial matrix $[-z^2(E \otimes A) + z(E \otimes E + A \otimes A) - A \otimes E]$. But, this matrix is equal to

$(zE-A)\otimes(E-zA)$. Since E and A form a regular pencil, we can always find a z such that $(zE-A)$ and $(E-zA)$ are both invertible, which implies that their Kroneker product is invertible. Q.E.D.

Equation (4.36), of course, does not completely characterize R_j; we need boundary conditions. One boundary condition that we already know is $R_0=P$ which can be computed from the Generalized Lyapunov Equation (4.5) either directly or by the perturbation method described previously. We can also obtain two more boundary conditions by multiplying Equation (4.1.2) by x_0 and x_N respectively and taking expectations. We obtain the following two equations

$$V^i P+V^f R_N = QE'^N \tag{4.39.1}$$

$$PV^{f\,\prime}+R_N V^{i\,\prime}=A^N Q. \tag{4.39.2}$$

These conditions are not guaranteed to completely characterize R_j in conjunction with (4.36); we may have to find other boundary conditions. For example, consider the case where $V^i=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ and $V^f=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$. In this case, it is clear that (4.39.1) and (4.39.2) do not completely characterize R_N (only the two main diagonal elements of R_N can be computed from (4.39.1) and (4.39.2)), and clearly the knowledge of $R_0(=P)$ and a partial knowledge of R_N do not result in a well-posed boundary condition.

Note, however, that the boundary for system (4.36)

consists of R_0, R_1, R_{N-1} and R_N, and if we know these 4 values we certainely have (probably more than) enough boundary conditions to solve (4.36). We can of course always find these directly from the Green's function solution for x. We do not explore the problem of obtaining appropriate boundary conditions any further because (4.36) is probably not the best method for computing the R_i's, anyway. A better method is based on the following two recursions obtained directly from (4.19.1) and (4.19.2):

$$ER_{j+1} = AR_j - BB'[V^f E^j A^{N-1-j}]' \tag{4.40.1}$$

$$R_{j+1}A' = R_j E' - V^i A^j E^{N-1-j} BB'. \tag{4.40.2}$$

Note that both of these recursions start from $R_0=P$ which as we have seen can be easily computed. When E is non-singular, we use (4.40.1) and when A is non-singular, we use (4.40.2). However, when E and A are both non-singular, (4.40.1) and (4.40.2) can only partially characterize R_j. It is easy to show that the part of R_j that cannot be characterized by these equations corresponds to the cross-correlation between the forward nilpotent and the backward nilpotent subsystems (we have already encountered this problem in the proof of Theorem (4.3)). However, we have already derived a simple expression for this cross-correlation (see Equation (4.31)). Of course, in order to use this equation we have to first decompose our system into a form similar to (4.22).

Example 4.3

We would like to obtain the covariance matrix R_j with $j\epsilon[0,N]$ for system (4.12). We have already computed the variance matrix (see Example (4.2))

$$R_0=P=Nb^2/4. \tag{4.41}$$

From Equation (4.36), we obtain the following recursion for R_j:

$$R_{j+2}=2R_{j+1}-R_j \tag{4.42}$$

for which we need R_1 as well as R_0 to start the recursion. R_1 can be easily computed from (4.20.1):

$$R_1=R_0-b^2/2=(N-2)b^2/4 \tag{4.43}$$

So,

$$R_2=2R_1-R_0=(N-4)b^2/4 \tag{4.44}$$

and in general,

$$R_j=(N-2j)b^2/4. \tag{4.45}$$

4.3- Stability

The relationship between the existance of a positive definite solution for the Lyapunov equation and stability for causal systems is well-known. In this section, we investigate this relationship in the more general case of stationary TPBVDS's. We know that for causal and reachable systems, the Lyapunov equation has a positive definite solution P if and only if the system is strictly stable. In

the case of TPBVDS's, the generalized Lyapunov equation (Equation 4.5) is explicitly a function of N, the length of the interval. This implies that an unstable, strongly reachable system may be stochastically stationary and thus have a constant matrix variance P (see Example 4.2), so the existance of a positive-definite solution to the generalized Lyapunov equation does not in general imply strict stability. As we will see the matrix P, in this case, diverges as N goes to infinity. This, does not happen in the causal case because P is independent of N. Considering these differences, it should be expected that our results may not be as simple as they are in the causal case.

First, notice that in the case of TPBVDS, the generalized Lyapunov equation may have a positive definite solution even when the system cannot be made stationary by proper choice of the boundary value covariance Q -i.e. there may be a nonnegative solution to (4.5) when there is <u>no</u> nonnegative solution for Q to (4.4). In this case the solution for P is <u>not</u> the constant variance matrix. However, we will show that as N goes to infinity, for any strictly stable stationary TPBVDS (as defined in Chapter III), the covariance matrices of the x's near the center of the interval approach a constant matrix P which <u>is</u> a solution to the generalized Lyapunov equation with N set to infinity. We will also show a partial converse. Specifically, we will

prove that if, for any $N \geq 0$, the generalized Lyapunov equation of a reachable system has a positive semi-definite solution P with unique main-diagonal elements, then the system is stable.

Let us now consider the uniquness of the solution of the generalized Lyapunov equation. It follows from (4.3.1) that the generalized Lyapunov equation has a unique solution if and only if the matrix

$$\Delta = E \otimes E - A \otimes A \tag{4.46}$$

is invertible. Since E and A have identical Jordan structure, the eigenvalues of Δ are of the form $\lambda_i \lambda_j - \mu_i \mu_j$ where λ_i's and μ_i's are the eigenvalues of E and A respectively, corresponding to the eigenvector v_i. Since λ_i and μ_i cannot be both zero because of regularity of the pencil $\{E,A\}$, Δ is singular only in the following two cases:

(1) There exists an eigenmode σ_i on the unit circle.

(2) There exist eigenmodes σ_i and σ_j such that $\sigma_i \sigma_j = 1$, or $\sigma_i = 0$ and $\sigma_j = \infty$.

The second statement can be easily shown by noting that

$$\sigma_i = \lambda_i / \mu_i \tag{4.47}$$

which implies that the eigenvalues of Δ can be written as $\sigma_i \sigma_j - 1$. When $\sigma_i = 0$ and $\sigma_j = \infty$ then E and A are both singular, a case where we already know that the Lyapunov equation does not uniquely specify P. The first statement is also easy to see because if there is a non-real eigenmode σ_i on the unit

circle, then its complex conjugate $\sigma_j=\sigma_i^*$ is also an eigenmode in which case clearly $\sigma_i\sigma_j-1=0$. And if σ_i equals 1 or -1, then $\sigma_i\sigma_i-1=0$ which means that Λ is not invertible. It is easy to show that in the first case where there is an eigenmode on the unit circle, the generalized Lyapunov equation does not characterize all the main-diagonal elements of P. Whereas, in the second case the main-diagonal elements of P are completely characterized by the generalized Lyapunov equation (although the off-diagonal elements are not). The reason this distinction is important is that in the first case (the case where the system has eigenmodes on the unit circle) clearly the system is not stable, however, in the second case the system could very well be stable eventhough the generalized Lyapunov equation does not have a unique solution.

Theorem 4.5

Let system (4.1) be reachable, then, if, for any N, the generalized Lyapunov equation has a positive semi-definite solution P with unique main-diagonal elements, then system (4.1) is strictly stable.

Proof

Notice first that since the main-diagonal elements of P are uniquely determined by the generalized Lyapunov

equation, system (4.1) does not have any eigenmodes on the unit circle. Thus, by making a change of coordinate, we can decompose the system as follows:

$$E=\begin{bmatrix} I & \\ & A_b \end{bmatrix},\ A=\begin{bmatrix} A_f & \\ & I \end{bmatrix},\ V^i=\begin{bmatrix} V_f^i & \\ & V_b^i \end{bmatrix},\ V^f=\begin{bmatrix} V_f^f & \\ & V_b^f \end{bmatrix},\ B=\begin{bmatrix} B_f \\ B_b \end{bmatrix}$$

where the eigenvalues of A_b and A_f are within the unit circle. To show stability we need to show that V_f^i and V_b^f are invertible. Using the above decomposition, we can decompose the generalized Lyapunov equation as follows

$$P_f-A_fP_fA_f=V_f^iB_fB_f'V_f^{i\,\prime}-V_f^f(A_f)^NB_fB_f'(A_f')^NV_f^{f\,\prime} \qquad (4.48.1)$$

$$A_bP_bA_b-P_b=V_b^i(A_b)^NB_bB_b'(A_b')^NV_b^{i\,\prime}-V_b^fB_bB_b'V_b^{f\,\prime} \qquad (4.48.2)$$

$$P_{fb}A_b'-A_fP_{fb}=V_f^iB_fB_b'(A_b')^NV_b^{i\,\prime}-V_f^f(A_f)^NB_fB_b'V_b^{f\,\prime} \qquad (4.48.3)$$

where

$$P=\begin{bmatrix} P_f & P_{fb} \\ P_{bf} & P_b \end{bmatrix}. \qquad (4.49)$$

Clearly, if P is positive definite, so is P_f. Since we also know that A_f is strictly stable, from (4.48.1) we can deduce that

$$v_k'(V_f^iB_fB_f'V_f^{i\,\prime}-V_f^f(A_f)^NB_fB_f'(A_f')^NV_f^{f\,\prime})v_k \geq 0 \qquad (4.50)$$

where v_k is any left eigenvector of A_f. What we would like to show is that V_f^i is invertible for which we need the following lemma.

<u>Lemma 4.3</u>

Let A and V be nxn matrices such that

$$AV=VA \qquad (4.51)$$

then if V is singular, there exists a right (left) eigenvector of A in the right (left) null space of V.

<u>Proof</u>

We will prove this result for the right eigenvector of A, the proof for the left eigenvector is similar.

Let

$$\xi \epsilon Ker(V) \tag{4.52}$$

then

$$V\xi=0. \tag{4.53}$$

So,

$$VA\xi=AV\xi=0 \tag{4.54}$$

which implies that

$$A\xi \epsilon Ker(V). \tag{4.55}$$

Thus Ker(V) is A-invariant, which implies that A has at least one eigenvector inside the null space of V. Q.E.D.

Suppose that V_f^i is not invertible, then using Lemma (4.3) we can deduce that there exists a left eigenvector of A_f, v_k, such that

$$v_k' V_f^i=0. \tag{4.56}$$

We also know that the system is reachable so

$$v_k'[V_f^i B_f \quad V_f^f B_f] \neq 0. \tag{4.57}$$

Thus,

$$v_k' V_f^f B_f \neq 0 \tag{4.58}$$

which since

$$\mathrm{Ker}(A_f^N) \subset \mathrm{Ker}(V_f^f) \tag{4.59}$$

implies that

$$V_f^f (A_f)^N B_f \neq 0. \tag{4.60}$$

Clearly, (4.56) and (4.60) are inconsistent with (4.50) which implies that V_f^i is invertible. Similarly we can show that V_b^f is invertible. Q.E.D.

Notice that the converse of Theorem (4.5) is not in general true. For example consider a strictly stable system which is reachable but not strongly reachable. We have seen in Chapter III that the part of the system which is reachable but not strongly reachable corresponds to an undriven system (this part is "reached" through the boundary condition just as an undriven causal system can be reached through the initial condition, however, in this case, the boundary condition is a function of the inputs reflected through the boundary condition as well as the boundary value). The only way in which an undriven system can have constant covariance is if the state of this system is constant. So, it is clear that a reachable but not strongly reachable system can be stochastically stationary only if the system has an eigenmode on the unit circle (in fact all the eigenmodes of the reachable but not strongly reachable part of the system must be on the unit circle if the process

is stochastically stationary). This of course is not consistent with our defintion of strict stability. However, this is not a problem if we only consider strongly reachable systems.

But then there is another problem. Equation (4.48.3) is not always consistent for strictly stable systems. In fact, consider the case where A_f and A_b are the same scalar, clearly then except for special cases of the input matrix B and/or the boundary matrices V, (4.48.3) is never satisfied. In general, this happens when A_b and A_f have a common eigenvalue which means that the system has eigenmodes σ_i and σ_j such that $\sigma_i\sigma_j=1$. Considering these problem, we obtain the following theorem.

<u>Theorem 4.6</u>

Let system (4.1) be strongly reachable and its eigenmodes σ_j satisfy

$$\sigma_j\sigma_k \neq 1 \tag{4.61}$$

for all j and k. Then, there exists $N^* > 0$ such that the generalized Lyapunov equation has a positive definite solution P for all $N > N^*$.

<u>Proof</u>

We first show that for N large enough (including infinity), the solutions of (4.48.1) and (4.48.2), P_f and

P_b, are positive definite and that the solution P_{fb} of (4.48.3) goes to zero as N goes to infinity.

Note that the right hand side of (4.48.3) goes to zero as N goes to infinity. Since A_f and A_b do not have any common eigenvalues (thanks to (4.61)), P_{fb} must also go to zero as N goes to infinity.

Now, note that (4.48.1) can also be written as

$$P_f - A_f P_f A_f = B_f B_f' - V_f^f (A_f)^N B_f B_f' - B_f B_f' (A_f')^N V_f^{f\,\prime}. \qquad (4.62)$$

Since the system is stable, $V_f^f (A_f)^N$ goes to zero as N goes to infinity and clearly for large enough N the right hand side of (4.62) is positive semi-definite. Moreover, $(A_f, \hat{B})$ is reachable in the causal sense for large enough N where

$$\hat{B}\hat{B}' = B_f B_f' - V_f^f (A_f)^N B_f B_f' - B_f B_f' (A_f')^N V_f^{f\,\prime}. \qquad (4.63)$$

This is true because for large enough N,

$$v_k' (B_f B_f' - V_f^f (A_f)^N B_f B_f' - B_f B_f' (A_f')^N V_f^{f\,\prime}) v_k > 0 \qquad (4.64)$$

since the system is strongly reachable, i.e.

$$v_k' B_f B_f' v_k > 0. \qquad (4.65)$$

Now combining (4.62) and (4.63), and the fact that $(A_f, \hat{B})$ is reachable in the causal sense, we get the standard causal Lyapunov equation which we know has a positive definite solution. Notice that even at $N=\infty$, (4.62), and thus (4.48.1), has a positive definite solution. Similarly, we can show that (4.48.2) also has a positive definite solution for large enough N (including $N=\infty$).

So far, we have shown that as N becomes larger, P_f and

P_b approach positive definite matrices and P_{fb} approachs zero. Thus, as N becomes larger, the eigenvalues of P converge to the eigenvalues of P_f and P_b which are all strictly greater than zero. This means that for N larger than some N^*, the eigenvalues of P are all positive. This of course is what we wanted to show. Q.E.D.

<u>Example 4.4</u>

Consider the following system

$$x_{k+1}=(1/2)x_k+u_k \tag{4.66.1}$$

$$rx_0+4rx_N=v \tag{4.66.2}$$

where

$$r=(1+4(1/2)^N)^{-1}. \tag{4.67}$$

System (4.66) is in standard form and stable. The generalized Lyapunov equation in this case is given by

$$(3/4)p=r^2(1-16(1/4)^N). \tag{4.68}$$

Clearly, (4.68) has a positive solution p only if N is larger than 2. So, in this case

$$N^*=2. \tag{4.69}$$

The strong reachability assumption of Theorem (4.6) was needed to show that (4.64) holds. Notice that for $N=\infty$, without the strong reachability assumption, (4.48.1) and (4.48.2) have positive semi-definite solutions and (4.48.3) has a zero solution. So that, the Lyapunov equation in this

case has a positive semi-definite solution. Also, note that since the right hand side of (4.48.3) is zero for $N=\infty$, we do not needed the assumption that $\sigma_j\sigma_k$ is different from 1.

<u>Corollary</u>

Let system (4.1) be strictly stable. Then the generalized Lyapunov equation with $N=\infty$ has a positive semi-definite solution P^*. In addition, if system (4.1) is strongly reachable, then P^* is positive definite.

The generalized Lyapunov equation at $N=\infty$ for a strictly stable system is given by

$$EPE'-APA'=R \tag{4.70}$$

where

$$R=\begin{bmatrix} B_fB_f & \\ & -B_bB_b \end{bmatrix}. \tag{4.71}$$

The solution to (4.70) is

$$P^*=\begin{bmatrix} P_f^* & \\ & P_b^* \end{bmatrix} \tag{4.72}$$

where P_f^* and P_b^* are solutions of the following Lyapunov equations

$$P_f-A_fP_fA_f'=B_fB_f' \tag{4.73.1}$$

and

$$P_b-A_bP_bA_b'=B_bB_b'. \tag{4.73.2}$$

<u>Theorem 4.7</u>

Let system (4.1) be strictly stable. Then for <u>any</u> choice of the boundary covariance Q, the variance matrices of the x's near the center of the interval [0,N] converge to P^* (solution of the generalized Lyapunov equation with $N=\infty$) as N goes to infinity.

<u>Proof</u>

Let $P_{k,N}$ be the variance matrix of the solution x_k of system (4.1) defined over the interval [0,N]. Then what we have to show is that

$$\lim_{N\to\infty} P_{(N/2)+j,N} = P^* \tag{4.74}$$

for all constant j (independent of N) where for simplicity we have assumed that N is even.

From the Green's function solution of (4.1) we can show that

$$\begin{aligned} P_{(N/2)+j,N} = {} & A^{(N/2)+j}E^{(N/2)-j}QE'^{(N/2)-j}A'^{(N/2)+j} + \\ & (V^i E^{(N/2)-j})\Pi_{(N/2)+j-1}(V^i E^{(N/2)-j})' + \\ & (V^f A^{(N/2)+j})\Pi_{(N/2)-j-1}(V^f A^{(N/2)+j})' \end{aligned} \tag{4.75}$$

where

$$\Pi_k = \sum_{m=0}^{k} A^{k-m}E^m BB'E'^m A'^{k-m}. \tag{4.76}$$

Assuming that the system is decomposed as before, it is easy to show that

$$\lim_{N\to\infty} P_{(N/2)+j,N} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}\Pi_\infty\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}\Pi_\infty\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}. \tag{4.77}$$

But since $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$ commute with E and A, we get

$$\lim_{N\to\infty} P_{(N/2)+j,N} = \lim_{k\to\infty} \sum_{m=0}^{k} A^{k-m} E^{m} \begin{bmatrix} B_f B_f' & \\ & B_b B_b' \end{bmatrix} E'^{m} A'^{k-m}. \quad (4.78)$$

So,

$$\lim_{N\to\infty} P_{(N/2)+j,N} =$$

$$\begin{bmatrix} \sum_{m=0}^{\infty} (A_f)^m B_f B_f' (A_f')^m & \\ & \sum_{m=0}^{\infty} (A_b)^m B_b B_b' (A_b)^m \end{bmatrix} =$$

$$\begin{bmatrix} P_f^* & \\ & P_b^* \end{bmatrix} = P^*. \quad (4.79)$$

This completes the proof of Theorem (4.7). Q.E.D.

Theorem (4.7) essentially means that any strictly stable stationary TPBVDS, regardless of the boundary value covariance, converges to a stochastically stationary system as N goes to infinity. This stochastically stationary system is separable, with independent forward and backward subsystems.

<u>Example 4.5</u>

Consider system (4.66). This system cannot be made stationary by proper choice of the boundary value covariane q. That is because the stochastic stationarity test is that q satisfies (4.4) or,

$$(3/4)q=r^2(-15) \tag{4.80}$$

which clealry cannot be satisfied by any non-negative q. Now consider the generalized Lyapunov equation with $N=\infty$:

$$(3/4)p^*=1 \tag{4.81}$$

which implies that

$$p^*=4/3. \tag{4.82}$$

Now, we compute the variance of $x_{N/2}$, $P_{N/2+k}$, as N goes to infinity,

$$P_{N/2}=(1/2)^N q+r^2 \sum_{m=0}^{N/2-1+k} (1/4)^m+16(1/2)^{N/2+2k} \sum_{m=0}^{N/2-1-k} (1/4)^m. \tag{4.83}$$

So,

$$\lim_{N\to\infty} P_{N/2+k} = 4/3 =p^* \tag{4.84}$$

which is the expected result.

V- Conclusion

5.1- Contributions

In this thesis, we have developed a system theory for two point boundary value descriptor systems (TPBVDS). We have first considered the deterministic problem (Chapter III) and then the stochastic problem (Chapter IV). The major contributions of our work in the deterministic case are:

(1) The derivation of a simple necessary and sufficient condition for well-posedeness and the concept of standard form which played a key role in deriving the Green's function solution for the TPBVDS.

(2) The derivation of a Generalized Cayley-Hamilton theorem.

(3) The extension of the concepts of inward and outward boundary processes introduced by Krener for his continuous-time boundary value linear systems to the case of the TPBVDS's.

(4) The study of various concepts of reachability and observability and their characterization in terms of reachability and observability spaces, which were then used to obtain reachability and observability tests.

(5) The characterization of stationary TPBVDS's and "almost stationary" TPBVDS's.

(6) Obtaining necessary and sufficient conditions for

minimality of stationary TPBVDS's and a procedure to reduce any non-minimal stationary TPBVDS to a minimal stationary TPBVDS.

(7) The study of different concepts of stability for stationary TPBVDS's.

In the stochastic case, our contributions are:

(8) The derivation of a necessary and sufficient condition for stochastic stationarity of stationary TPBVDS driven by white noise.

(9) The derivation of the Generalized Lyapunov equation which must be satisfied by the variance matrix (covariance matrix evaluated at zero) of any stochastically stationary TPBVDS.

(10) The development of a perturbation method that guarantees that the variance matrix of a stochastically stationary TPBVDS can always be computed from its Generalized Lyapunov equation.

(11) The derivation of various recursions for the covariance matrices of stochastically stationary TPBVDS.

(12) The study of the relationships between the notion of stability and the existance of a solution for the generalized Lyapunov equation.

5.2- Suggestions for Further Research

The followings are some open questions and possible topics of future research:

(1) The study of time varying TPBVDS.

The general case of time varying TPBVDS is probably not of much interest, however, it is possible that there may be interesting structures if we consider some special classes of time varying TPBVDS's, e.g. time-varying models equivalent to autonomous but non-stationary TPBVDS's.

(2) The study of the relationship between TPBVDS with stationary weighting patterns and stationary TPBVDS's.

We know that a minimal non-stationary TPBVDS could have a stationary weighting pattern. We suspect that such a TPBVDS can be realized by a stationary TPBVDS of the same dimension. Moreover, we believe that such stationary TPBVDS can be constructed by modifying the strongly unreachable and strongly unobservable parts of the system (remember that any strongly reachable and strongly observable TPBVDS with stationary weighting pattern is stationary).

(3) Further studies of the problem of stability.

We have shown that there are connections between the concept of stability presented in Chapter III and the Generalized Lyapunov equation derived in Chapter IV. Further research is needed to fully understand these connections.

(4) The development of a stochastic realization theory.

Krener has considered the problem of realizing continuous-time Gaussian reciprocal processes as outputs of stationary boundary value linear systems. A natural extension of this problem would be to realize discrete Gaussian reciprocal processes with stochastically stationary TPBVDS since the non-causal structure of these models fits perfectly the non-causal nature of reciprocal processes.

(5) The development of numerically stable algorithms to compute the solution of TPBVDS.

This problem has not been considered in our work. However, we suspect that our study of stability is crucial for the development of stable algorithms.

APPENDIX A: The Inward-Boundary Value Process z'

In this section, we present a recursive method for computing the inward-boundary value process z'. Then we show that z' has all the properties discussed in Chapter III. Finally, we consider the special case where E and A are both invertible.

A.1-Introduction

In Chapter III, we defined the inward-boundary value process z'. We claimed that z' always exists and that system (3.14) has the same solution as system (3.1) over its domain of definition.

We have seen that the outward process z(j,k) is obtained by eliminating x_{j+1} through x_{k-1} in Equation (3.2.1). The process z'(j,k) is obtained in a similar fashion; in this case however, we eliminate x_0 through x_{j-1} and x_{k+1} through x_N by performing row cancellations on the system matrix S (defined in Section 3.1 of Chapter III).

A.2-Recursive Method to Compute z'

As seen in Chapter III, our TPBVDS can be written as

$$S\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} Bu_0 \\ \vdots \\ Bu_{N-1} \\ v \end{bmatrix} \tag{A.1.1}$$

where S is the system matrix given by

$$S=\begin{bmatrix} -A & E & & & \\ & -A & E & & \\ & & \cdot & \cdot & \\ & & & -A & E \\ V^i & & & & V^f \end{bmatrix}. \tag{A.1.2}$$

We know that the outward boundary value process z can be computed by performing elementary row operations on the matrix S. In fact by simply premultiplying (A.1.1) by

$$\Pi = [0 \ldots 0 \ A^{j-i-1} \ EA^{j-i-2} \ .. \ E^{j-i-1} \ 0 \ldots 0] \tag{A.2.1}$$

we obtain

$$-A^{j-i}x_i - E^{j-i}x_j =$$
$$A^{j-i-1}Bu_i + EA^{j-i-2}Bu_{i+1} + \ldots + E^{j-i-1}Bu_{j-i} = z(i,j) \tag{A.2.2}$$

where z(i,j) is the outward process on [i,j].

The computation of the inward process z' is more complicated but similar to that of the outward process z. In this case we premultiply both sides of (A.1) by

$$\Omega = \begin{bmatrix} 0 & & & & \\ & I & & & \\ & & \cdot & & \\ & & & I & \\ T & & & & P \end{bmatrix} \tag{A.3}$$

where [T P] has full row rank and

$$[T \quad P]\begin{bmatrix} -A \\ V^i \end{bmatrix} = 0. \tag{A.4}$$

T and P always exist because $\begin{bmatrix} -A \\ V^i \end{bmatrix}$ has full column rank

(otherwise the matrix S would not be invertible). This gives

$$\begin{bmatrix} 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & -A & E & & \\ \cdot & & \cdot & \cdot & \\ \cdot & & & -A & E_f \\ 0 & TE & & & PV^f \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} 0 \\ Bu_1 \\ \vdots \\ Bu_{N-1} \\ Pv+TBu_0 \end{bmatrix} \tag{A.5}$$

which can be written as

$$S' \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} Bu_1 \\ \vdots \\ Bu_{N-1} \\ Pv+TBu_0 \end{bmatrix} \tag{A.6}$$

where

$$S' = \begin{bmatrix} -A & E & & & \\ & -A & E & & \\ & & \cdot & \cdot & \\ & & & -A & E_f \\ TE & & & & PV^f \end{bmatrix}. \tag{A.7}$$

The matrix S' has full rank because S has full rank and

$$\mathrm{rank}(\Omega)=\mathrm{rank}(S)-n \tag{A.8.1}$$

which means that

$$\mathrm{rank}(S')=\mathrm{rank}(S)-n \tag{A.8.2}$$

and S' has n less columns and n less rows than S. Thus S' is the system matrix of the following well-posed system

$$Ex_{k+1}=Ax_k+Bu_k \tag{A.9.1}$$

$$TEx_1+PV^f x_N=Pv+TBu_0. \tag{A.9.2}$$

Notice that (A.9) is defined over [1,N]. The boundary matrices of (A.9) are not necessarily in standard-form so that we need to premultiply (A.9.2) by

$$\Lambda=(TE^N+PV^f A^{N-1})^{-1}. \tag{A.10}$$

It should be clear that

$$W^i_{1N}=\Lambda TE \tag{A.11.1}$$

$$W^f_{1N}=\Lambda PV^f \tag{A.11.2}$$

are the boundary matrices which appear in Equation (3.13) of Chapter III, and

$$z'(1,N)=\Lambda Pv+\Lambda TBu_0. \tag{A.11.3}$$

Thus, we have shown how to move the left boundary inward by one step. Clearly, we can apply the same method again and move the left boundary inward to any point on the interval [0,N]. Similarly, we can move the right boundary inward. All we need to do is to premultiply (A.1) by

$$\Omega' = \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & 0 \\ 0 & & & T'P' \end{bmatrix} \tag{A.12}$$

where [T' P'] has full row rank and

$$[T' \;\; P'] \begin{bmatrix} -E \\ V^f \end{bmatrix} = 0. \tag{A.13}$$

The boundary matrices W and z' can be computed as in the previous case.

By using the recursive method described above, we can compute W^i_{jk}, W^f_{jk} and z'(j,k) for all $j,k\epsilon[0,N]$. It should also be clear that (3.14) has the same solution as (3.1) over its domain of definition.

A.3-Reachable Space of z'

In this section, we derive an expression for $\mathfrak{R}'(j,k)$

(the reachable space of z'(j,k)) and prove Theorem (3.5).

In the previous section, we described a method for recursively computing the boundary matrices W and the process z' at any two point on [0,N]. In fact, it is possible to directly compute these boundary matrices W and and z' as follows.

Using the expression (3.11) for the outward process z and Equation (A.1), we can obtain the following expression

$$\begin{bmatrix} -A^j & E^j & & \\ & -A^{k-j} & E^{k-j} & \\ & & -A^{N-k} & E^{N-k} \\ V^i & & & V^f \end{bmatrix} \begin{bmatrix} x_0 \\ x_j \\ x_k \\ x_N \end{bmatrix} = \begin{bmatrix} z(0,j) \\ z(j,k) \\ z(k,N) \\ v \end{bmatrix}. \tag{A.14.1}$$

Premultiplying (A.14.1) by

$$\Omega_{jk} = \begin{bmatrix} 0 & I & 0 & 0 \\ T^i_{jk} & 0 & T^f_{jk} & P_{jk} \end{bmatrix} \tag{A.14.2}$$

where $[T^i_{jk} \; T^f_{jk} \; P_{jk}]$ has full rank and

$$-T^i_{jk}A^j + P_{jk}V^i = 0 \tag{A.15.1}$$

$$T^f_{jk}E^{N-k} + P_{jk}V^f = 0, \tag{A.15.2}$$

we get

$$\begin{bmatrix} -A^{k-j} & E^{k-j} \\ T^i_{jk}E^j & -T^f_{jk}A^{N-k} \end{bmatrix} \begin{bmatrix} x_j \\ x_k \end{bmatrix} = \begin{bmatrix} z(j,k) \\ T^i_{jk}z(0,j) + T^f_{jk}z(k,N) + P_{jk}v \end{bmatrix}. \tag{A.16}$$

From (A.16), we see that the boudary matrices at j and k are given by

$$W^i_{jk} = \Lambda_{jk}T^i_{jk}E^j \tag{A.17.1}$$

$$W^f_{jk}=-\Lambda_{jk}T^f_{jk}A^{N-k}, \tag{A.17.2}$$

where

$$\Lambda_{jk}=(S^i_{jk}E^k-S^f_{jk}A^{N-j})^{-1}. \tag{A.17.3}$$

The identity (A.16) also shows that the inward boundary value process is given by

$$z'(j,k)=\Lambda_{jk}(T^i_{jk}z(0,j)+T^f_{jk}z(j,N)+P_{jk}v). \tag{A.18}$$

Notice that the introduction of Λ_{jk} is necessary to guarantee that the boundary matrices are in standard form.

It is clear from (A.18) that the reachable space of $z'(j,k)$ is

$$\mathcal{R}'(j,k)=\Lambda_{jk}(T^i_{jk}\mathcal{R}+T^f_{jk}\mathcal{R}) \tag{A.19}$$

for j and k far enough from the boundaries (so that $z(0,i)$ and $z(j,N)$ can be arbitrary elements of $\mathcal{R}$).

Now, assume that j-1 is far from the boundaries as well. We would like to show that in this case $\mathcal{R}'(j-1,k)$ has the same dimension as $\mathcal{R}'(j,k)$. The first thing we do is to find $T^i_{j-1,k}$, $T^f_{j-1,k}$ and $P_{j-1,k}$. In fact we show that a possible choice for $T^i_{j-1,k}$, $T^f_{j-1,k}$ and $P_{j-1,k}$ is

$$T^i_{j-1,k}=T^i_{jk}\tilde{A} \tag{A.20.1}$$

$$T^f_{j-1,k}=T^f_{jk} \tag{A.20.2}$$

$$P_{j-1,k}=P_{jk}, \tag{A.20.3}$$

where $\tilde{A}$ has the same eigenstructure as A except that its zero eigenvalues have been replaced by 1. $\tilde{A}$ can be obtained by transforming A into Jordan form by using some

transformation U, then replacing all the zero diagonal elements (eigenvalues) with one and finally transforming back this matrix using U^{-1}. The result is $\tilde{A}$. Clearly, if A is invertible $\tilde{A}=A$.

What we need to show is that $[T^{i}_{j-1,k}\ T^{f}_{j-1,k}\ P_{j-1,k}]$ has full rank (which is clearly true since $\tilde{A}$ is invertible) and that $T^{i}_{j-1,k}$, $T^{f}_{j-1,k}$ and $P_{j-1,k}$ satisfy (A.15.1) and (A.15.2). Equation (A.15.2) is clearly satisfied. Thus, we have to show that

$$-T^{i}_{jk}\tilde{A}A^{j-1}+P_{jk}V^{i}=0 \tag{A.21}$$

when we know that

$$-T^{i}_{jk}A^{j}+P_{jk}V^{i}=0. \tag{A.22}$$

Suppose that A has been transformed into the following Jordan form by using some transformation U (this corresponds to a change of coordinate which clearly does not affect the rank of the reachability matrix)

$$A=\begin{bmatrix} J & \\ & N \end{bmatrix} \tag{A.23}$$

where J is invertible and in Jordan form and N is nilpotent and in Jordan form. Then clearly

$$\tilde{A}=\begin{bmatrix} J & \\ & N+I \end{bmatrix}. \tag{A.24.1}$$

Since j-1 is far from the boundaries, $N^{j-1}=0$. So,

$$A^{j-1}\tilde{A}=\begin{bmatrix} J^{j-1} & \\ & 0 \end{bmatrix}\begin{bmatrix} J & \\ & N+I \end{bmatrix}=A^{j} \tag{A.24.2}$$

which directly implies (A.21), which is the desired result.

Thus, we can write the reachable space of z'(j-1,k) as

$$\mathcal{R}'(j-1,k)=\Lambda_{j-1,k}(T^{i}_{jk}\tilde{A}\mathcal{R}+T^{f}_{jk}\mathcal{R}). \tag{A.25}$$

Notice that if either j-1 or k is not far enough from the boundaries, the reachable space of z(0,j-1) or the reachable space of z(k,N) may not be the whole space $\mathcal{R}$ and thus (A.25) may not hold.

Now, we prove that

$$\tilde{A}\mathcal{R}=\mathcal{R}. \tag{A.26}$$

This of course implies that the dimension of $\mathcal{R}'(j-1,k)$ equals that of $\mathcal{R}'(j,k)$ because all Λ's are invertible.

To prove (A.26), again we assume that A is in the Jordan form (A.22). Since

$$\tilde{A}=A+\begin{bmatrix}0 & \\ & I\end{bmatrix} \tag{A.27}$$

all we need to show is that

$$\begin{bmatrix}0 & \\ & I\end{bmatrix}\mathcal{R}\subset\mathcal{R}. \tag{A.28}$$

We know that

$$\mathcal{R}=\mathrm{Im}\begin{bmatrix}B_1 & JB_1 & \cdots\cdots J^{n-1}B_1 & \\ B_2 & NB_2 & \cdots N^{\mu-1}B_2 & 0\cdots 0\end{bmatrix} \tag{A.29}$$

where μ is the nilpotency degree of N. Let J be $n_1 x n_1$ and N be $n_2 x n_2$; then clearly $n_1+n_2=n$ and $\mu\leq n_2$. The inclusion (A.28) is equivalent to the following statement:

$$\text{if } \begin{bmatrix}\xi_1\\ \xi_2\end{bmatrix}\epsilon\mathcal{R}, \quad \text{then } \begin{bmatrix}0\\ \xi_2\end{bmatrix}\epsilon\mathcal{R}.$$

But if $\begin{bmatrix}\xi_1\\ \xi_2\end{bmatrix}\epsilon\mathcal{R}$, there exist vectors u_i where $0\leq i\leq\mu-1$ such that

$$\xi_2 = \sum_{i=0}^{\mu-1} N^i B_2 u_i . \tag{A.30}$$

We would like to show that we can find u_i where $\mu \leq i \leq n-1$ such that

$$\sum_{i=0}^{n-1} J^i B_1 u_i = 0, \tag{A.31}$$

which means that

$$\sum_{i=\mu}^{n-1} J^i B_1 u_i = -\sum_{i=0}^{\mu-1} J^i B_1 u_i \tag{A.32}$$

or equivalently,

$$\sum_{i=\mu}^{n-1} J^{i-\mu} B_1 u_i = J^{-\mu}(-\sum_{i=0}^{\mu-1} J^i B_1 u_i). \tag{A.33}$$

Clearly $\sum_{i=0}^{\mu-1} J^i B_1 u_i$ is in the reachable space of (J, B_1), and since J is invertible the right hand side of (A.33) is in the reachable space of (J, B_1). Thus, since $n-1-\mu > n_2-1$, we can easily deduce from the Cayley-Hamilton Theorem that we can always find u_i where $\mu \leq i \leq n-1$ such that (A.33) is satisfied. This, of course, is the desired result.

So far, we have shown that $\mathcal{R}'(j-1,k)$ has the same dimension as $\mathcal{R}'(j,k)$. Similarly, we can show that $\mathcal{R}'(j,k+1)$ has the same dimension as $\mathcal{R}'(j,k)$ where $j,k+1 \in [n, N-n]$. It is clear then that $\mathcal{R}'(j,k)$ has the same dimension for all $j,k \in [n, N-n]$. This completes the proof of Theorem (3.5).

A.4-Case of Invertible E and A

When E and A are both invertible, we obtain simple expressions for the inward process z' and the boundary matrices W. We also obtain a simple expression for the Green's function solution of system (3.1).

The first step is to find T^i_{jk}, T^f_{jk} and P_{jk} satisfying (A.15.1) and (A.15.2). It is easy to see that a possible choice is given by

$$T^i_{jk}=V^iA^{-j} \tag{A.34.1}$$

$$T^f_{jk}=-V^fE^{k-N} \tag{A.34.2}$$

and

$$P_{jk}=I. \tag{A.34.3}$$

Then by computing Λ_{jk} and using (A.16.1), (A.16.2) and (A.17), we can show that the boundary matrices W and the inward process z' can be written as

$$W^i_{jk}=A^jE^{N-k}V^i(EA^{-1})^j \tag{A.35.1}$$

$$W^f_{jk}=-A^jE^{N-k}V^f(AE^{-1})^{N-k} \tag{A.35.2}$$

and

$$z'(j,k)=A^jE^{N-k}(v+V^iA^{-j}z(0,j)+V^fE^{k-N}z(k,N)). \tag{A.36}$$

Since

$$x_j=z'(j,j) \tag{A.37}$$

we obtain the following simple expression for the Green's function solution

$$G(j,k)=\begin{cases}-A^{j}E^{N-j}V^{f}E^{k-N}A^{N-k-1} & k\geq j\\ A^{j}E^{N-j}V^{i}E^{k}A^{-k-1} & j>k.\end{cases} \qquad (A.38)$$

This expression can also be derived directly from Equation (3.9); its derivation, however, is not trivial.

APPENDIX B: Recursive Solutions for TPBVDS's

Unlike for causal systems, the solution of a TPBVDS cannot be computed using a simple recursion. This of course is expected since the solution x at any time k is a function of the inputs over the whole interval [0,N]. In this section, we present two algorithms to compute the solution of TPBVDS's and propose ideas on how a third algorithm can be constructed.

The first algorithm we discuss, the two-filter solution, was proposed by Adams [24]. His algorithm processes the inputs from left to right and from right to left. The second algorithm presented in this section processes the inputs starting from the center and going outwards to the boundaries and then starting from the boundaries and moving in toward the center. Finally, we discuss an algorithm based on the computation of the inward and outward processes z' and z.

B.1- The Two-Filter Solution

Adams formulates the general solution of a TPBVDS as a linear combination of two stable recursions, one forward and the other backward. His formulation is presented below as it appeared in [22], with only a few changes in the notation.

Since {E,A} comprise a regular pencil, there exist

nonsingular matrices F and T such that

$$FET^{-1} = \begin{bmatrix} I & 0 \\ 0 & A_b \end{bmatrix} \overset{\Delta}{=} E' \tag{B.1.1}$$

and

$$FAT^{-1} = \begin{bmatrix} A_f & 0 \\ 0 & I \end{bmatrix} \overset{\Delta}{=} A' \tag{B.1.2}$$

where all eigenvalues of A_f and A_b lie within the unit circle (we assume that the system has no eigen-mode on the unit circle). The above decomposition splits the system into two subsystems:

$$x_{f,k+1} = A_f x_{f,k} + B_f u_k \tag{B.2.1}$$

and

$$x_{b,k} = A_b x_{b,k+1} - B_b u_k \tag{B.2.2}$$

where

$$\begin{bmatrix} x_{f,k} \\ x_{b,k} \end{bmatrix} = Tx_k \tag{B.3.1}$$

and

$$\begin{bmatrix} B_f \\ B_b \end{bmatrix} = FB. \tag{B.3.2}$$

Given the above transformation the boundary condition takes the form

$$[V_f^i \vdots V_b^i] \begin{bmatrix} x_{f,0} \\ x_{b,0} \end{bmatrix} + [V_f^f \vdots V_b^f] \begin{bmatrix} x_{f,N} \\ x_{b,N} \end{bmatrix} \tag{B.4}$$

where

$$[V_f^i \vdots V_b^i] = V^i T^{-1} \overset{\Delta}{=} V^{i\prime} \tag{B.5.1}$$

and

$$[V_f^f \vdots V_b^f] = V^f T^{-1} \overset{\Delta}{=} V^{f\prime}. \tag{B.5}$$

Define $x^0_{f,k}$ as the solution to (B.2.1) with zero initial condition and $x^0_{b,k}$ as the solution to (B.2.2) with zero final condition. Then it is easy to see that

$$x_{f,k}=(A_f)^k x_{f,0}+x^0_{f,k} \tag{B.6.1}$$

and

$$x_{b,k}=(A_b)^{N-k} x_{b,k}+x^0_{b,k}. \tag{B.6.2}$$

Substituting for $x_{f,N}$ and $x_{b,0}$ from (B.6.1) and (B.6.2) into (B.4) and solving for $x_{f,0}$ and $x_{b,N}$ gives

$$\begin{bmatrix} x_{f,0} \\ x_{b,N} \end{bmatrix} = (F_N)^{-1}[v-V^f_f x^0_{f,N}-V^0_b x^0_{b,0}] \tag{B.7}$$

where

$$F_N = V^{i'}(E')^N + V^{f'}(A')^N. \tag{B.8}$$

Finally, substituting for $x_{f,0}$ and $x_{b,k}$ from (B.7) and (B.8), it can be shown that the solution to (B.2) is given by

$$\begin{bmatrix} x_{f,k} \\ x_{b,k} \end{bmatrix}=(E')^{N-k}(A')^k(F_N)^{-1}[v-V^f_f x^0_{f,N}-V^0_b x^0_{b,0}]+\begin{bmatrix} x^0_{f,k} \\ x^0_{b,k} \end{bmatrix}. \tag{B.9}$$

Applying the inverse of the transformation in (B.3.1), the original process x_k is recovered by

$$x_k=T^{-1}\begin{bmatrix} x_{f,k} \\ x_{b,k} \end{bmatrix}. \tag{B.10}$$

In this way, Adams has constructed a stable forward/backward two filter recursive implementation of the general solution of a TPBVDS. Notice that the invertibility of the matrix F_N is not an issue, since F_N is invertible if the system is well-posed (in fact invertibility of F_N is our test for

well-posedness). The stability of $(F_N)^{-1}$, however, is an issue. If $(F_N)^{-1}$ grows unbounded as N goes to ∞, the two filter solution is not stable. But, if $(F_N)^{-1}$ grows unbounded so does the solution x which means that the system is unstable. Of course, we cannot expect to find a stable algorithm to compute the solution of an unstable system. This concept of stability, by the way, is consistent with our previous concept of stability (Definition 3.8 in Chapter III).

B.2- The Outward-Inward Solution

The idea behind this algorithm is to first compute the outward process z by starting from the center of the interval and then stepping outward one step at the time. Once we reach the boundaries, we use z(0,N) and the boundary condition to compute x_0 and x_N, and finally, we use backward substitution starting from x_0 and x_N and moving inward one step at the time to compute all the x's.

For simplicity, we assume that N is odd and that E and A commute. In this case our outward recursion starts from z((N-1)/2,(N+1)/2) and is given by

$$\begin{aligned} z([(N-1)/2]-k-1,[(N+1)/2]+k+1) &= \\ EAz([(N-1)/2]-k,[(N+1)/2]+k) &+ \\ A^{2k+2}Bu_{[(N-1)/2]-k-1} &+ E^{2k+2}Bu_{[(N-1)/2]+k} \end{aligned} \tag{B.11}$$

with the initial condition

$$z((N-1)/2,(N+1)/2) = Bu_{(N-1)/2}. \tag{B.12}$$

Using the above recursion, we compute $z(j,N-j)$ for $j=0,\ldots,(N-1)/2$. Now we can use $z(0,N)$ to solve for x_0 and x_N as follows:

$$-A^N x_0+E^N x_N=z(0,N) \tag{B.13.1}$$

and

$$V^i x_0+V^f X_N=v \tag{B.13.2}$$

so that,

$$\begin{bmatrix} x_0 \\ x_N \end{bmatrix} = \begin{bmatrix} -A^N & E^N \\ V^i & V^f \end{bmatrix}^{-1} \begin{bmatrix} z(0,N) \\ v \end{bmatrix}. \tag{B.14}$$

It is easy to show that the matrix $\begin{bmatrix} -A^N & E^N \\ V^i & V^f \end{bmatrix}$ is invertible if and only if the system is well-posed, i.e. $V^iE^N+V^fA^N$ is invertible.

To construct the inward stage of the solution, we note that

$$-A^{N-2j}x_j+E^{N-2j}x_{N-j}=z(j,N-j) \tag{B.15.1}$$

and

$$\delta_j Ex_j+Ax_{N-j}=\delta_j Ax_{j-1}+Ex_{N-j+1}+\delta_j Bu_{j-1}-Bu_{N-j} \tag{B.15.2}$$

for all $j\epsilon[1,(N-1)/2]$ and all scalar δ_j. Equation (B.15.2) follows simply from the descriptor dynamics. Using (B.15.1) and (B.15.2) to solve for x_j and x_{N-j} we get

$$\begin{bmatrix} x_j \\ x_{N-j} \end{bmatrix} = \begin{bmatrix} -A^{N-2j} & E^{N-2j} \\ \delta_j E & A \end{bmatrix}^{-1} \begin{bmatrix} z(j,N-j) \\ \delta_j A x_{j-1} + E x_{N-j+1} + \delta_j B u_{j-1} - B u_{N-j} \end{bmatrix} \quad (B.16)$$

where δ_j is chosen such that $\begin{bmatrix} -A^{N-2j} & E^{N-2j} \\ \delta_j E & A \end{bmatrix}$ is invertible (if the system has no eigenmode on the unit circle, δ_j can be taken equal to 1). Of course (B.16) is the desired recursion which starts from the values of x_0 and x_N computed previously. Notice that both the outward and the inward recursions are time-varying.

A variation of the above algorithm can be constructed by noting that the solution x can be constructed from the outward process z and the inward process z'. More precisely, we can compute z(j,N-j) as done in the outward-inward algorithm and z'(j,N-j) from the recursive solution described in Appendix A. Then the solution x can be computed as

$$\begin{bmatrix} x_j \\ x_{N-j} \end{bmatrix} = \begin{bmatrix} -A^{N-2j} & E^{N-2j} \\ W^i_{j,N-j} & W^f_{j,N-j} \end{bmatrix}^{-1} \begin{bmatrix} z(j,N-j) \\ z'(j,N-j) \end{bmatrix}. \quad (B.17)$$

Notice that this algorithm resembles the two-filter solution because the inward and outward boundary processes z and z' can be computed independently. The outward-inward solution, however, requires that the outward recursion be completed before the inward recursion can start. The reason for this

is that the final state of the outward recursion is used to compute the initial state of the inward recursion.

References

(1) Krener, A. J., Boundary Value Linear Systems, Asterisque, 1980.

(2) Krener, A. J., Acausal Realization Theory, Part I: Linear Deterministic Systems, to be published.

(3) Gohberg, I. and Kaashoek, M. A., Time Varying Linear Systems with Boundary Conditions and Integral Operators, I. The Transfer Function and its Properties, Int. Eq. and Op. Th. 7(1984) 325-391.

(4) Gohberg, I. and Kaashoek, M. A., Time Varying Linear Systems with Boundary Conditions and Integral Operators, II. Similarity and Reduction. Preprint.

(5) Gohberg, I. and Kaashoek, M. A., On minimality and stable minimality of time varying linear systems with well posed boundary conditions, Proc. 7th Internat. Symposium on the Math. Theory of Networks and Systems, Stockholm, Sweden, June 1985.

(6) Luenberger, D. G., Time-Invariant Descriptor Systems, Automatica, vol. 14, 1978.

(7) Luenberger, D. G., Dynamic Equations in Descriptor Form, IEEE Trans. on Automat. Contr., vol. AC-22, no.3, June 1979.

(9) Lewis, F. L., Adjoint Matrix, Cayley-Hamilton Theorem, and Fadeev's Method for the Matrix Pencil (sE-A), Proc. 22nd IEEE Conf. Dec. and Contr., San Antonio, TX, 1983.

(10) Lewis, F. L., Descriptor Systems: Expanded Descriptor Equations and Markov Parameters, IEEE Trans. on Automat. Contr., vol. AC-28, no.5, May 1983.

(11) Lewis, F. L., Descriptor Systems: Decomposition into Forward and Backward Subsystems, IEEE Trans. on Automat. Contr., vol. AC-29, no.1, Jan. 1984.

(12) Lewis, F. L., Fundamental, Reachability, and Observability Matrices for Discrete Descriptor Systems, IEEE Trans. on Automat. Contr., vol. AC-30, no.5, May 1985.

(13) Bender, D. J., Lyapunov-Like Equations and Reachability/Observability Gramians for Descriptor Systems, submitted for publication.

(14) Cobb, D., Controllability, Observability, and Duality in Singular Systems, submitted for publication.

(15) Verghese, G. C.; Levy, B. C.; and Kailath, T., A Generalized State-Space for Singular Systems, IEEE Trans. on Automat. Contr., vol. AC-26, no.4, August 1981.

(16) Van Dooren, P. M., The Generalized Eigenstructure Problem in Linear System Theory, IEEE Trans. on Automat. Contr., vol. AC-26, no.1, Feb. 1981.

(17) Yip, E. L. and Sincovec, F., Solvability, Controllability, and Observability of Continuous Descriptor Systems, IEEE Trans. on Automat. Contr., vol. AC-26, no.3, June 1981.

(18) Verghese, G. C., Infinite Frequency Behaviour in Generalized Dynamincal Systems, Ph.D. Thesis, Stanford University, Dec. 1978.

(19) Rosenbrock, H. H., Structural Properties of Linear Dynamincal Systems, Int. J. Contr., vol. 20, no.2, 1974.

(20) Adams, M. B., Linear Estimation of Boundary Value stochastic Processses, Sc.D. Thesis, Dept. of Aero. and Astro. MIT, Cambridge, MA., Jan. 1983.

(21) Adams, M. B.; Willsky, A. S. and Levy, B. C., Linear Estimation of Boundary Value Stochastic Processes, Part I and II, IEEE Trans. Automat. Contr., vol. AC-29, 1984.

(22) Grantmacher, F. R., The Theory of Matrices. New York: Chelsea, 1974.

(23) Kailath, T., Linear Systems. Prentice-Hall, 1980.

(24) Adams, M. B.; Willsky, A. S. and Levy, B. C., Linear Smoothing for Descriptor Systems, Memo. LIDS-P-1367, Laboratory for Information and Decision Systems, MIT. Cambridge, MA., March 1984.

(25) Krener, A. J., Reciprocal Processes and the Stochastic Realization Problem for Acausal Systems, Proc. 7th Internat. Symposium on the Math. Theory of Networks and Systems, Stockholm, Sweden, June 1985.

(26) Bernstein, S., Sur les liaisons entre les grandeurs aleatoires, Proc. of Int. Cong. of Math., Zurich (1932) 288-309.

(27) Carmichael, J. P., Masse, T. C., and Theodorescu, R., Processus gaussiens stationaires reciproque sur un intervalle, C. R. Acad. Sc.

Paris, Serie I 295(1982) 291-293.