NORTH-HOLLAND SYSTEMS AND CONTROL SERIES

VOLUME 9

TIME-DELAY SYSTEMS

ANALYSIS, OPTIMIZATION AND APPLICATIONS

M. MALEK-ZAVAREI
M. JAMSHIDI

NORTH-HOLLAND

TIME-DELAY SYSTEMS
Analysis, Optimization and Applications

NORTH-HOLLAND SYSTEMS AND CONTROL SERIES
VOLUME 9

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

TIME-DELAY SYSTEMS
Analysis, Optimization and Applications

MANU MALEK-ZAVAREI
AT&T Bell Laboratories
Holmdel, NJ 07733
U.S.A.

MOHAMMAD JAMSHIDI
CAD Laboratory for Systems & Robotics
Department of Electrical and Computer Engineering
The University of New Mexico
Albuquerque, NM 87131
U.S.A.

1987

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

ISBN: 0 444 70204 0

Publishers:

ELSEVIER SCIENCE PUBLISHERS B.V.
P.O. Box 1991
1000 BZ Amsterdam
The Netherlands

Sole distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 Vanderbilt Avenue
New York, N.Y. 10017
U.S.A.

Library of Congress Cataloging-in-Publication Data

Malek-Zavarei, Manu.
Time-delay systems.

(North-Holland systems and control series ; v. 9)
Includes bibliographies.
1. Automatic control. 2. Control theory. 3. Delay lines. I. Jamshidi, Mohammad. II. Title. III. Series.
TJ213.M293 1987 629.8 87-5248
ISBN 0-444-70204-0 (U.S.)

PRINTED IN THE NETHERLANDS

To my Parents (MM-Z)

In Memory of my Parents and
Brother M. Hassan Jamshidi (MJ)

PREFACE

This book is intended to familiarize the reader with an important class of systems: time-delay systems. Such systems arise either as a result of inherent delays in components of the system, or because of deliberate introduction of delay into the system for control purposes.

Delays often occur in the transmission of information or material between different parts of a system. Transportation systems, communications systems, chemical processing systems, metallurgical processing systems, environmental systems and power systems are examples of time-delay systems.

Considerable research has been done on various aspects of time-delay systems in the last 25 years. However, unlike linear nondelay systems on which a host of books have been published, very few books exist on time-delay systems. The existing books are either limited in scope or are too theoretical, often to the point of obscurity. The present book is intended to fill this void. It covers the techniques of analysis and design of time-delay systems without dwelling deeply on mathematical rigor; but this, of course, is not at the expense of sacrificing accuracy. The book is concerned mainly with linear time-delay systems. However, nonlinear time-delay systems are also included whenever applicable. Large-scale time-delay systems are also presented. The book is useful for graduate students in engineering, science and mathematics. It may be used as a textbook for a graduate course on time-delay systems. It would also be useful for researchers in this field since up-to-date surveys on most topics are provided.

The authors have been doing research on time-delay systems since 1969 at various locations including University of Illinois (Urbana, IL), Shiraz University (formerly Pahlavi University, Shiraz, Iran), IBM Thomas J. Watson Research Center (Yorktown Heights, NY), Hughes Aircraft Co. (Canoga Park, CA), the University of New Mexico (Albuquerque, NM) and AT&T Bell Laboratories (Holmdel, NJ). The book includes the authors' course notes as well as the results of their research and other researchers in this field. It consists of four parts and nine chapters as follows:

Part I - Modeling (Chapter 2)

Part II - Analysis (Chapters 3-5)

Part III - Optimization (Chapters 6-8)

Part IV - Applications (Chapter 9)

Chapter 1 serves as an introduction to the book. Time-delay systems are introduced here and examples of their applications are presented. Part I consisting of Chapter 2 presents mathematical description of time-delay systems. System modeling, utilizing both transfer functions (frequency domain) and state-space representation (time-domain), as well as methods of linearization of nonlinear time-delay systems are presented here. Also, modeling of large-scale systems with time delays is considered in this chapter.

Part II consists of Chapters 3, 4 and 5. Chapter 3 is concerned with the analysis of linear time-delay systems. The solution of the first-order time-delay state equation with single delay or multiple delays in the unforced and forced cases is provided here. The concept of adjoint state equation is also presented in Chapter 3. The results of this chapter are used in subsequent chapters dealing with stability, controllability, observability and optimal control. Chapter 4 deals with the stability of time-delay systems. Methods in time domain and frequency domain are presented here. Controllability and observability of time-delay systems are discussed in Chapter 5 where criteria for controllability and observability of stationary and nonstationary linear time-delay systems are developed. The concept of duality is invoked to relate controllability and observability.

Part III of the book consists of Chapters 6, 7 and 8. Optimization of time-delay systems is the topic of Chapter 6. The maximum principle for time-delay systems is stated here and its proof is outlined. The computational difficulties in the application of the maximum principle to the optimal control of time-delay systems is discussed in this chapter and the dynamic programming method of optimal control of time-delay systems is presented. Methods of suboptimal control of time-delay systems are discussed in Chapter 7. Nonlinear and multiple delay systems are treated here as well. Chapter 8 deals with suboptimal control of large-scale time-delay systems. The optimal hierarchical control methods for such systems are considered here.

Part IV consisting of Chapter 9 is concerned with the applications of time-delay systems in different disciplines. Illustrative examples for different applications such as cold rolling mills, traffic control and water resources control systems are presented in this chapter.

To make the book self-sufficient, the required mathematical background is provided in Appendices A and B. These appendices present a review of linear algebra and the transform (Laplace, z and modified z) theories.

The authors are indebted to many people for their various contributions. M. Malek-Zavarei especially thanks Jack Appel, Mo Iwama, John O'Rourke and John Williams for their encouragement and support. M. Jamshidi would like to thank Dean Jerry May of the University of New Mexico's College of Engineering and Russ Seacat, Chairperson of the Electrical and Computer Engineering Department for their leadership and continuous support. The authors would like to thank Professor Madan Singh of the University of Manchester Institute of Science and Technology for reviewing the manuscript and offering many helpful suggestions. They would also like to thank Professor Singh as the editor of North-Holland Systems and Control Series for his suggestion to write the book. We are indebted to many of our colleagues and former students who have contributed to the book in various forms. Special thanks are due Mr. Soo Ryong Lee for a thorough proofreading of the final manuscript.

Finally, the manuscript was prepared using the text processing tools of the UNIX™ operating system. Without these tools, the typing and proofreading of this 196,225-word manuscript would have been indeed an enormous task. We greatly appreciate the efforts of the staff of AT&T Bell Laboratories Text Processing Center and Art Center in preparing the manuscript. Last but not least, we are grateful to our families for their patience and support during the preparation of the manuscript.

Manu Malek-Zavarei
Holmdel, New Jersey

Mohammad Jamshidi
Albuquerque, New Mexico

September 1986

CONTENTS

Preface vii

Chapter 1. PRELIMINARIES

1.1 INTRODUCTION 1
1.2 TYPES OF SYSTEMS 2
1.2.1 Lumped-Parameter and Distributed-Parameter Systems 2
1.2.2 Deterministic and Stochastic Systems 3
1.2.3 Continuous-Time, Discrete-Time and Hybrid Systems 3
1.2.4 Linear and Nonlinear Systems 3
1.2.5 Time-Invariant and Time-Varying System 4
1.3 EXAMPLES OF TIME-DELAY SYSTEMS 5
1.3.1 Cold Rolling Mill 6
1.3.2 Engine Speed Control 7
1.3.3 Spaceship Control 8
1.3.4 A Water Resources System 8
1.3.5 A Hydraulic System 9
1.3.6 A Wind Tunnel System 10
1.4 NOTATION 10
1.4.1 Numbering and Cross Referencing 10
1.4.2 Conventions 11
1.4.3 Abbreviations and Symbols 12
1.5 SCOPE OF THE BOOK 13
REFERENCES 15

PART I: MODELING

Chapter 2. MATHEMATICAL DESCRIPTION OF TIME-DELAY SYSTEMS

2.1 INTRODUCTION 23
2.2 MODELING OF TIME-DELAY SYSTEMS 23
2.3 STATE-SPACE REPRESENTATION 24
2.4 FREQUENCY-DOMAIN REPRESENTATION 28
2.4.1 Continuous-Time Time-Delay Systems 28

2.4.2 Discrete-Time Time-Delay Systems 31
2.4.3 Hybrid Systems 33
2.5 LINEARIZATION OF NONLINEAR TIME-DELAY SYSTEMS 40
2.6 LARGE-SCALE TIME-DELAY SYSTEMS 42
2.6.1 Coupled Time-Delay Systems 42
2.6.2 Modeling of Time-Delay Systems via Singular Perturbation 43
2.6.3 Serial Time-Delay Systems 50
2.6.4 Hierarchical Formulation of a Large-Scale Time-Delay System 51
PROBLEMS 53
REFERENCES 57

PART II: ANALYSIS

Chapter 3. ANALYSIS OF TIME-DELAY SYSTEMS

3.1 INTRODUCTION 61
3.2 HOMOGENEOUS STATE EQUATION – FUNDAMENTAL MATRIX 61
3.3 FORCED STATE EQUATION – COMPLETE SOLUTION 66
3.4 ADJOINT STATE EQUATIONS 70
PROBLEMS 74
REFERENCES 76

Chapter 4. STABILITY OF TIME-DELAY SYSTEMS

4.1 INTRODUCTION 77
4.2 STABILITY OF LINEAR TIME-DELAY SYSTEMS VIA TIME-DOMAIN 79
4.2.1 The Lyapanov Method 79
4.2.2 Uniform Asymptotic Stability 84
4.2.3 Stability of Stochastic Time-Delay Systems 90
4.2.4 Stability of Time-Delay Systems with Variable Delays 93
4.3 STABILITY OF TIME-DELAY SYSTEMS VIA FREQUENCY DOMAIN 97
4.3.1 Pontryagin's Stability Method 99
4.3.2 Graphical Methods 106
4.3.3 Grapho-Analytical Methods 115
PROBLEMS 122
REFERENCES 125

Chapter 5. CONTROLLABILITY AND OBSERVABILITY OF LINEAR TIME-DELAY SYSTEMS

5.1 INTRODUCTION 133
5.2 DEFINITIONS OF CONTROLLABILITY 133
5.3 CONTROLLABILITY CRITERIA 134
5.4 DEFINITIONS OF OBSERVABILITY 142
5.5 OBSERVABILITY CRITERIA 144
PROBLEMS 148
REFERENCES 150

PART III: OPTIMIZATION

Chapter 6. OPTIMIZATION OF TIME-DELAY SYSTEMS

6.1 INTRODUCTION 155
6.2 STATEMENT OF THE PROBLEM 155
6.3 THE MAXIMUM PRINCIPLE 157
6.4 GENERALIZED RICCATI METHOD 160
6.4.1 The Linear Time-Invariant Case 161
6.4.2 The Linear Time-Varying Case 163
6.5 DYNAMIC PROGRAMMING METHOD 166
6.6 TIME-OPTIMAL CONTROL 170
6.6.1 Optimal Controls 173
6.6.2 Optimal Initial Functions 174
PROBLEMS 177
REFERENCES 179

Chapter 7. SUBOPTIMAL CONTROL OF TIME-DELAY SYSTEMS

7.1 INTRODUCTION 181
7.2 SENSITIVITY APPROACH 181
7.2.1 The Linear Time-Invariant Case 182
7.2.2 The Linear Time-Varying Case 190
7.3 NONDELAY CONVERSION APPROACH 202
7.3.1 Linear Systems with Multiple Constant Delays 203
7.3.2 Extensions to Other Systems 209
7.4 SUBOPTIMAL CONTROL OF NONLINEAR TIME-DELAY SYSTEMS 210
7.5 SINGULAR PERTURBATION METHOD 222
7.6 SENSITIVITY TO PARAMETER VARIATIONS 227
7.6.1 Sensitivity with Respect to Plant Parameter Variations 228

7.6.2 Sensitivity with Respect to Approximate Modeling 233
7.6.3 Stability Considerations 236
PROBLEMS 238
REFERENCES 241

Chapter 8. NEAR-OPTIMUM DESIGN OF LARGE-SCALE TIME-DELAY SYSTEMS

8.1 INTRODUCTION 245
8.2 NEAR-OPTIMUM CONTROL OF COUPLED TIME-DELAY SYSTEMS 245
8.3 HIERARCHICAL CONTROL 260
8.3.1 Goal Coordination 260
8.3.2 Interaction Prediction 264
8.3.3 Costate Prediction 267
8.4 HIERARCHICAL CONTROL OF SERIAL TIME-DELAY SYSTEMS 279
8.5 HIERARCHICAL CONTROL OF NONSERIAL TIME-DELAY SYSTEMS 285
8.6 HIERARCHICAL CONTROL OF TIME-DELAY SYSTEMS VIA INTERACTION PREDICTION 301
8.6.1 System Linearization and Decomposition 302
8.6.2 Method of Solution 304
8.6.3 Computational Algorithm 307
8.7 HIERARCHICAL CONTROL OF TIME-DELAY SYSTEMS VIA COSTATE PREDICTION 318
PROBLEMS 338
REFERENCES 341

PART IV: APPLICATIONS

Chapter 9. SOME APPLICATIONS OF TIME-DELAY SYSTEMS

9.1 INTRODUCTION 347
9.2 COLD ROLLING MILLS 347
9.2.1 The Rolling Process 348
9.2.2 A State-Space Model 352
9.2.3 Optimal Control Design 353
9.2.4 Computational Procedure 360
9.2.5 Numerical Solution of the Feedback Gain Matrix 364
9.3 TRAFFIC CONTROL 371
9.3.1 A Hierarchical TD Algorithm 371

9.3.2 Model of an Oversaturated Intersection 378
9.3.3 A Traffic Control Case Study 383
9.4 WATER RESOURCES SYSTEMS............................. 389
9.4.1 Discrete Differential Dynamic Programming............... 390
9.4.2 Khuzestan Water Resources System 395
9.4.3 Computational Results 403
9.5 A HYDRAULIC LEVEL SYSTEM 405
9.6 CONTROL OF TIME-DELAY SYSTEMS VIA SMITH PREDICTOR 409
9.6.1 The Smith Predictor 409
9.6.2 An Adaptive Smith Predictor TD Control System 414
REFERENCES 420

Appendix A. REVIEW OF LINEAR ALGEBRA

A.1 INTRODUCTION 425
A.2 FIELDS AND VECTOR SPACES 425
A.3 LINEAR INDEPENDENCE AND RANK 427
A.4 BASIS AND DIMENSION 428
A.5 LINEAR OPERATORS 429
A.6 INNER PRODUCT AND NORM 430
A.7 JORDAN CANONICAL FORM OF MATRICES 434
A.8 FUNCTIONS OF A SQUARE MATRIX 441
A.9 QUADRATIC FORMS 445
A.10 DIFFERENTIATION WITH RESPECT TO A VECTOR 447
A.11 DIGITAL COMPUTER IMPLEMENTATION 448
PROBLEMS 449
REFERENCES 453

Appendix B. REVIEW OF LAPLACE, z AND MODIFIED z TRANSFORMS

B.1 INTRODUCTION 455
B.2 REVIEW OF THE LAPLACE TRANSFORM 455
B.2.1 Properties of the Laplace Transform 457
B.2.2 Inversion of the Laplace Transform 462
B.2.3 Solution of Differential Equations 468
B.3 REVIEW OF THE z TRANSFORM 471
B.3.1 Properties of the z Transform 475
B.3.2 Inversion of the z Transform........................... 476
B.3.3 Solution of Difference Equations 480

B.4 THE MODIFIED z TRANSFORM 481
B.4.1 Properties of the Modified z Transform 483
B.4.2 Inverse Modified z Transform 485
B.4.3 Solution of Differential-Difference Equations.............. 487
PROBLEMS 489
REFERENCES...492

SUBJECT INDEX...495

CHAPTER 1

PRELIMINARIES

1.1 INTRODUCTION

The purpose of this chapter is to serve as a general introduction to the book. It is intended to provide motivation for the study of time-delay (TD) systems as well as to build a framework within which the properties of such systems can be explained. As such, it may refer to some concepts which are not defined yet but will be defined in later chapters.

Time-delay systems* are those systems in which time delays exist between the application of input or control to the system and their resulting effect on it. They arise either as a result of inherent delays in the components of the system or as a deliberate introduction of time delay into the system for control purposes. Time delays occur often in electronic, mechanical, biological, metallurgical, and chemical systems. They correspond to transport time (as in shock waves in the earth, hormones in the blood stream, fluids in a chemical process, or electromagnetic radiation in space) or to computation times (as in cortical processing of a visual image, analyzing a TV picture by a robot, evaluating the output of a digital control algorithm, or performing a chemical composition analysis) [1.1]. The mathematical formulation of a TD system results in a system of delay-differential equations. A particular class of these equations, the integro-differential equations, was first studied by Volterra [1.2-1.3] who developed a theory for them and investigated time delay phenomena in different systems. Others [1.4-1.20] have made significant contributions to the development of the general theory of functional differential equations of Volterra type.

* Sometimes also referred as *time-lag* or *retarded systems*. We will consistently use *time delay systems* in this book.

Another class of delay-differential equations is the differential-difference equations. Major contributors in this area are Bellman, *et al.* [1.21-1.24], Elsgolts [1.25], Hahn [1.26], Halanay [1.27], Krasovskii [1.28], Wright [1.29-1.31] and Zubov [1.32].

Control processes with time delays were first studied by Callender, *et al.* [1.33]. A historical account of control theory and the relevance of time delays to it is given by Bateman [1.34]. Two excellent bibliographies by Weiss [1.35] and Chosky [1.36] list the contributions to different aspects of TD systems until 1959. A large number of papers have appeared in the literature since then reporting research performed on TD systems.

Despite the considerable amount of research done on different aspects of TD systems, the results of this research is scattered mainly in the form of technical papers published in various journals. The existing books [1.27, 1.37, 1.38] treat TD systems essentially from the mathematical viewpoint with scant reference to the applications of such systems. The book by Marshall [1.39] treats some practical aspects of TD control systems. However, in addition to other shortcomings [1.40], it treats TD systems only in the frequency domain.

In the following sections, first different system classifications will be discussed; then several examples of TD systems will be presented for the purposes of motivation. The notation to be used in the book will be explained in another section and a final section on the scope of the book will conclude this chapter.

1.2 TYPES OF SYSTEMS

From this point on by the word "system" we mean a system model, *i.e.*, a mathematical representation of a physical system. (See Chapter 3.) Systems can be classified according to the type of equations describing them. One important classification is as follows.

1.2.1 Lumped-Parameter and Distributed-Parameter Systems

Lumped-parameter systems are those which can be described by ordinary differential (or difference) equations. In contrast, distributed-parameter systems are those which require partial differential equations for their characterization. A lumped electric circuit, *i.e.*, one in which the element sizes are negligible compared to the wavelength of the highest frequency of operation, is an example of a lumped-parameter system. A transmission line is an example of a distributed-parameter system. Time-delay systems are a compromise between lumped and distributed parameter systems in that they are described by delay differential or difference equations. Some examples of TD systems will be presented in Section 1.3.

Other system classifications are as follows.

1.2.2 Deterministic and Stochastic Systems

In deterministic systems all the parameters can be described exactly. In stochastic systems, however, some (or all) parameters can be described only probabilistically, *i.e.*, as random variables. We will be concerned mainly with deterministic systems in this book.

1.2.3 Continuous-Time, Discrete-Time and Hybrid Systems

In *continuous-time systems*, described by differential equations, all the variables are defined for all values of time in typically a semi-infinite interval $[t_0, \infty]$. In contrast, the variables in *discrete-time systems* are defined only at discrete instants of time. Discrete-time systems are described by difference equations.

In *hybrid systems* (sometimes referred to as *mixed systems*) part of the system is continuous-time and part of it is discrete-time. Such systems are described by differential-difference equations.

1.2.4 Linear and Nonlinear Systems

A linear system is one to which the superposition principle applies. That is, a linear system exhibits proportionality of input and output. This is an oversimplification; a rigorous definition of system linearity will be given in Chapter 2. A linear system can be described by a linear differential or difference equation. A system which is not linear is called a *nonlinear* system.

1.2.5 Time-Invariant and Time-Varying System

A system is called *time-invariant* or *stationary* if all its parameters are constant. Such systems can be described by constant-coefficient differential or difference equations. If one or more of the system parameters vary with time, the system will be called *time-varying* or *nonstationary*. Such systems are described by differential or difference equations with time-varying coefficients.

It should be noted that a given system falls into one of the two categories in each of the above classifications. For example, a system may be distributed, stochastic, continuous-time, linear and time-varying. This book will concentrate mainly on linear, deterministic, nonstationary TD systems. Both continuous-time and discrete-time systems will be considered. Digressions will sometimes be made to cover some aspects of nonlinear TD systems.

1.3 EXAMPLES OF TIME-DELAY SYSTEMS

It can be claimed that any real-life system has some time delay associated with it. The delay (or delays) in a system may be due to one or more of the following causes:

1. Measurement of system variables.
2. Physical properties of the equipment used in the system.
3. Signal transmission (transport delay).

Furthermore, sometimes time delay is built into a system deliberately for proper operation. An example of such a system is that for the control of room temperature. If delay is not built into such a control system, the thermostat-controlled relay will continuously chatter, thus rendering the control impossible. The effect of the time delay on the system dynamics, however, depends on the size of the delay and the system characteristics. Often the delay can be neglected in the system model without considerable effect on the analysis or design. On the other hand, there are instances where the delay is continuous, thus giving rise to a distributed model for the system. For most systems, the use of the exact distributed model is not warranted because of the complications in the analysis of distributed-parameter systems. Time-delay systems often provide acceptable models for such systems.

We are concerned with systems in which the time delay plays an important role. Examples of such systems are control systems, economic systems, political systems, biological and environmental system. Detailed examples of such systems will be presented in Chapter 9. To motivate the reader, in this section we will briefly discuss several examples of TD systems.

1.3.1 Cold Rolling Mill

In a cold rolling mill delays are due to high-speed transport of steel strips between stands. Consider the process of producing a sheet of steel of uniform thickness by rolling tempered steel through a feeder to a multi-stand mill. The sheet thickness is controlled by adjusting the spacing between the rollers according to measurements (by X-ray diffraction) of thickness made on the sheets leaving the roller. (See Figure 1.) This measurement, however, cannot be made as soon as the sheet leaves the roller. It has to be made further down the line where the thickness has reached an equilibrium value due to cooling. Thus there is a delay between the time a sheet leaves the rollers at the *ith* stand and the time it reaches the rollers at the $(i+1)st$ stand at which corrections are applied to the spacing between the rollers.

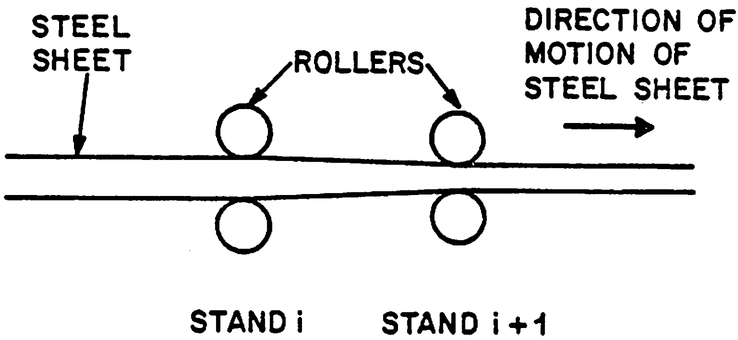

Figure 1. Cold Rolling Mill

1.3.2 Engine Speed Control

A control signal is translated into action by the actuators after some delay due to the inertia in the control equipment. As an example consider the problem of controlling the speed of a steam engine running an electric power generator under varying load conditions. A rudimentary control system for this purpose is the centrifugal governor invented by James Watt in the eighteenth century [1.42, 1.43]. This control system consisted of a set of fly balls or rotating weights suspended from levers which were connected to the steam valve. The fly balls were driven by the engine. An increase in the engine speed would result in a larger centrifugal force which would raise the fly balls and thereby would reduce the steam flow into the engine causing a reduction in its speed. (See Figure 2.) However, there would be a delay between the time that engine speed increased and the time that steam flow was reduced. This delay is due to the inertia of the control equipment.

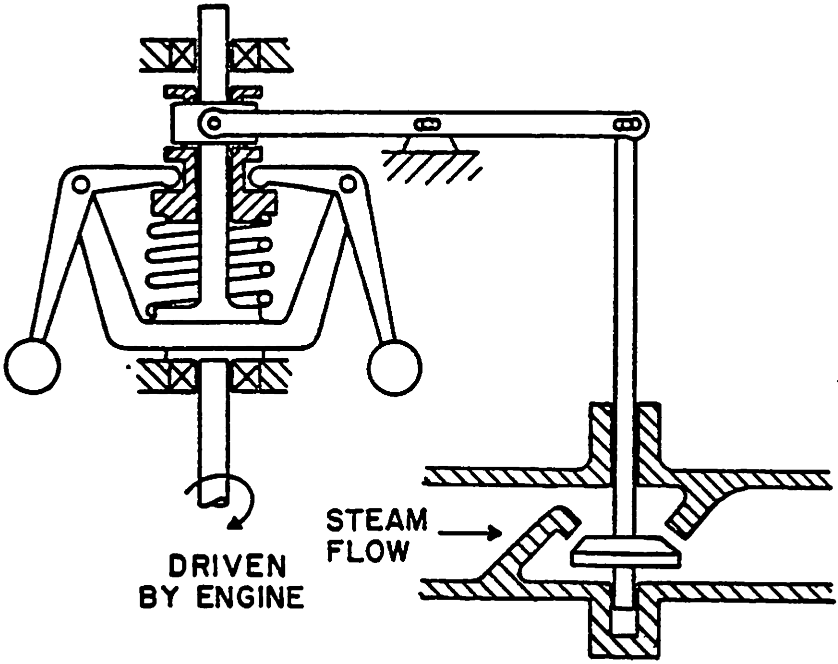

Figure 2. Watt's Centrifugal Governor for the Speed Control of Steam Engine

1.3.3 Spaceship Control

The time delay may be due to the transmission of the control signal over a long distance. Even at the traveling speed of electromagnetic waves in free space (300,000 *Km/s*), this delay could be considerable enough to have effects on the system dynamics. For example, for a spaceship traveling toward the moon, the control signals transmitted from the earth could take as much as one second to reach the spaceship.

1.3.4 A Water Resources System

Consider a water resources system shown in Figure 3. As seen, the system consists of two reservoirs, two power plants, one river and one irrigation area. The mathematical characterization of such a system is based on the continuity of water volume at the reservoir outlets (points A and B in Figure 3) and various system nodes (points C and D) and is generally nonlinear. An accurate model of such a system involves quantities of water returned from irrigation channels which are commonly delayed by one operating period, usually a month.

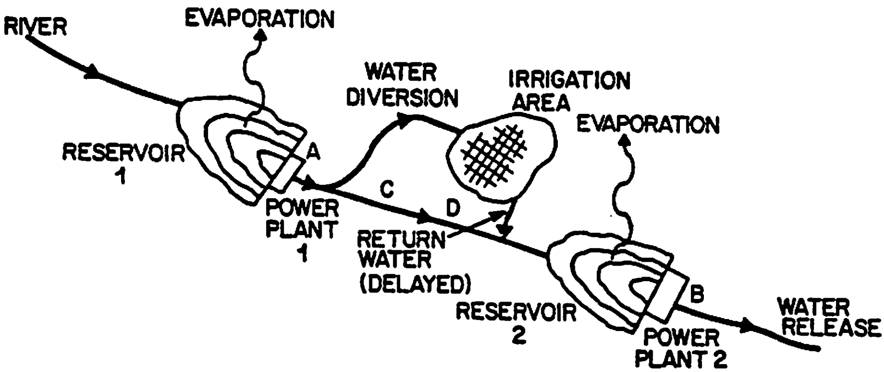

Figure 3. A 2-Reservoir Water Resources System

1.3.5 A Hydraulic System

Consider a hydraulic system shown in Figure 4. As seen in this figure the object is to fill a reservoir up to a level, say h, and keep it at that level. Due to the finite travel time of the water from the valve to the tank (reservoir) a delay exists between the hydraulic actuator and the tank. This system will be reconsidered and analyzed in Chapter 9.

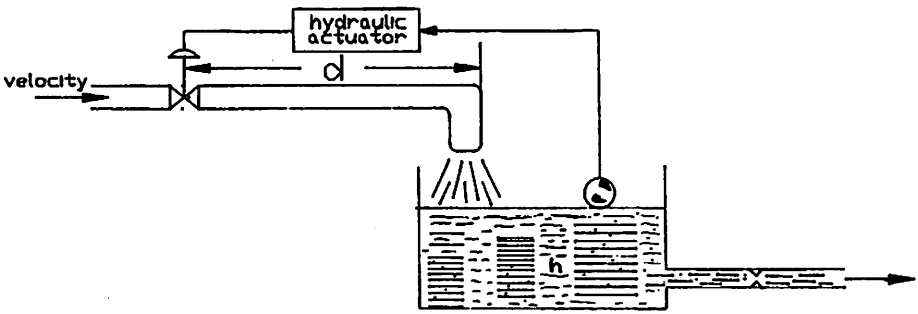

Figure 4. A Hydraulic System

1.3.6 A Wind Tunnel System

Another example of time-delay systems is a continuous-flow cryogenic wind tunnel operating at low temperatures and high Reynolds numbers [1.44]. The tunnel is fan-driven with a Mach number control through fan motor speed regulation and guide-vane angle control. The model of this system, as will be detailed in Chapter 9, is a third-order differential-difference equation involving a delay in the state. The delay represents the transportation time between the guide vanes of the fan and the test section of the tunnel [1.45]. In Chapter 9, a spectrum assignment design scheme via state feedback will be used for this system.

1.4 NOTATION

In this section first the method of numbering and cross referencing in the book will be explained. Then the notations and abbreviations used throughout the book will be compiled.

1.4.1 Numbering and Cross Referencing

Chapters of the book are consecutively numbered. The sections in each chapter are numbered consecutively by two digits the first of which is the chapter number. Subsections in each section are similarly numbered consecutively by three digits. The equations in each section are numbered on the right-hand side consecutively. Also, the figures in each section are numbered consecutively. For reference to an equation or a figure in the *same* section, only the number of the equation will be referred to. But to refer to an equation or a figure in *another* section, the two-digit section number will also precede the number of the equation or the figure.

Also in each section consecutive numbers are used to identify definitions, theorems and examples. In the *same* section, these are referred to by their numbers. In *other* sections, these are referred to by a three-digit number consisting of their number preceded by the two-digit number of the section in which they appear.

1.4.2 Conventions

Capital letters denote sets or vector spaces, e.g. S, V.

Lower case and Greek letters indicate scalars and scaler-valued functions, e.g. m, α, $f(t)$.

Bold lower case letters indicate vectors, e.g. $\mathbf{x}$, $\mathbf{y}$

Bold capital letters indicate matrices, e.g. $\mathbf{A}$, $\mathbf{B}$

Capital script letters denote fields or transformations, e.g. R.

The Laplace transform of a function is denoted by the corresponding capital letter, e.g. $G(s) = L[g(t)]$.

The z transform of a function is indicated by the corresponding capital letter, e.g. $H(z) = Z[h(t)]$. The modified z transform of a function is indicated by the corresponding capital letter and subscript or argument m, $H(z,m) = Z_m[h(t)]$.

Superscript "′" denotes the transpose of a vector or a matrix, e.g. $\mathbf{x}'$, $\mathbf{A}'$

Superscript "*" denotes the conjugate transpose of a vector or a matrix, e.g. $\mathbf{x}^*$, $\mathbf{A}^*$.

Superscript "-1" denotes the inverse of a matrix or a transformation, e.g. $\mathbf{A}^{-1}$, L^{-1}.

Superscript "$\perp$" denotes the orthogonal complement of a subspace, e.g. $V^{\perp}$.

A dot over a time function denotes its derivative with respect to time, e.g. $\dot{x}$, $\dot{\mathbf{x}}$.

Two or three dots over a time function denotes its second or third derivative with respect to time, respectively, e.g. $\ddot{x}$, $\dddot{x}$. For higher-order derivatives a superscript with the corresponding derivative order in parenthesis may also be used, e.g. $x^{(3)}$, $y^{(4)}$.

A bar over a scalar or a vector denotes its complex conjugate, e.g. $\bar{\alpha}$, $\bar{x}$.

Braces indicates sets , e.g. $\{x\}$.

1.4.3 Abbreviations and Symbols

(.,.), (.,.] and [.,.] denote, respectively, open, semiclosed and closed intervals, respectively, e.g., t in the interval (a,b] means $a<t\leqslant b$.

$\|.\|$: norm of a vector or a matrix, e.g. $\|x\|$, $\|A\|$.
$\triangleq$: equal by definition
$\Rightarrow$: implies
$\Leftarrow$: is implied by
$\Leftrightarrow$: implies and is implied by
$\exists$: there exists
$\forall$: for all
$\ni$: such that
ϵ : belongs to, e.g. $x \epsilon S$
$\notin$: does not belong to
$\supset$: contains, e.g. $S_1 \supset S_2$
$\subset$: is contained in
$\cup$: union
$\cap$: intersection
$\oplus$: direct sum
Δ : end of an example or discussion
$*$: convolution
(.,.) : inner product
j : the imaginary number $\sqrt{-1}$
adj : adjoint of a matrix, e.g. adjA
det(.) : determinant of a matrix, e.g. det(A)

dim (.) : dimension of a vector space, e.g. *dim*(V)

ρ : rank of a matrix, e.g. ρ(A)

γ(.) : nullity of a matrix, e.g. γ(A)

tr (.) : trace of a matrix, e.g. tr(A)

D^{-n} (.) : *n*th-order integration of a function, e.g. $D^{-2}(f)$

diag : diagonal matrix

q.e.d. : quod erat demonstrandum (which was to be proved)

l.t.i. : linear time-invariant

w.r.t. : with respect to

KCL : Kirchoff's current law

KVL : Krichoff's voltage law

exp (.) : exponential of a scalar or a matrix, e.g., $exp(\mathrm{A}t) \triangleq e^{\mathrm{A}t}$

TD : time delay

TPBV : two-point boundary-value (problem)

SISO : single-input single-output

MIMO : multi-input multi-output

LSS : Large-scale system

1.5 SCOPE OF THE BOOK

The main purpose of the book is to review research results on various aspects of TD systems such as analysis, stability, controllability, observability, optimization and optimal control, as well as areas of application. We do not intend to dwell too deeply on mathematical rigor. This, however, by no means implies that accuracy will be compromised. Time domain method will be used predominantly throughout the book. However, reference will also be made to frequency domain techniques and classical control, as appropriate.

The book consists of four parts addressing the following topics in TD systems:

Part I. Modeling

Part II. Analysis

Part III. Optimization

Part IV. Applications

Part I includes Chapter 2 which reviews the mathematical description and modeling of TD systems. Methods of linearization of nonlinear TD systems as well as modeling and decomposition of large-scale TD systems are topics which are discussed in Chapter 2. Methods of approximation of distributed systems as TD systems in both continuous-time and discrete-time cases and the sensitivity of systems to these approximations are also discussed in this chapter.

Part II consisting of Chapters 3, 4 and 5 is concerned with the analysis and structural properties of TD systems. The methods of analysis of TD systems in time and frequency domains are treated in Chapter 3. The solution of the first-order vector-matrix TD state equation in the unforced and forced cases is provided in this chapter. Both single and multiple delay systems are treated. The results of this chapter are used in subsequent chapters on stability, controllability, observability and optimal control. Chapter 4 deals with stability of TD systems. Methods in time domain (Lyapunov theory) and frequency domain (the circle criterion) are presented here. Controllability and observability of TD systems are the topics of Chapter 5. Here, criteria for controllability and observability of stationary and nonstationary linear TD systems are developed. System with time-varying delays as well as multiple delays are also treated in Chapter 5. The concept of duality is invoked to relate controllability and observability.

Part III consisting of Chapters 6, 7 and 8 deals with the optimization of TD systems. The maximum principle for TD systems is stated without a rigorous proof in Chapter 6. However, a sketch of the proof is presented. The computational difficulties in the application of the maximum principle to the optimal control of TD systems is discussed here. Also, the dynamic programming method of optimal control of TD systems is presented in this chapter. Suboptimal control of TD systems is the topic of Chapter 7 where linear as well as nonlinear and multiple delay systems are treated. Also, the sensitivity of the performance index of linear TD systems to parameter variations is studied in Chapter 7. Chapter 8 deals with suboptimal control of large-scale TD systems. The optimal hierarchical control methods (interaction prediction, serial time-delay and costate prediction) as well as decentralized control and stabilization problems are considered here.

Part IV consisting of Chapter 9 is concerned with the applications of TD systems in different disciplines. Illustrative examples, case studies or design problems for different applications such as cold rolling mills, traffic control, water resources control systems, economic and chemical processes are presented in this chapter.

Basic mathematical tools required to understand the topics covered in the book are reviewed in Appendices A and B. Linear algebra is reviewed in Appendix A. Laplace, z and modified z transforms are reviewed in Appendix B.

REFERENCES

[1.1] A. T. Bahill, "A simple adaptive Smith-predictor for controlling time-delay systems," *IEEE Control Syst. Mag.*, Vol. 3, pp. 16-22, 1983

[1.2] *V. Volterra, Lecons sur les* équations intégrales etles équations intégro-differentielles, Gauthier-Villars, Paris, 1913. Also see "Sur la théorie mathématique des phénoménes héréditaires," *F. Math. Paris*, Vol. 7, fasc. III, pp. 249-298, 1928

[1.3] *ibid, Theory of Functionals and Integral and Integro-Differential Equations*, Dover, New York, 1959

[1.4] L. Tonelli, "Sulle equazioni funzionali del tipo di Volterra," *Bull. Calcutta Math. Soc.*, Vol. 20, pp. 31-48, 1928

[1.5] A. N. Tikhonov, "Functional equations of Volterra type and their applications to some problems of mathematical physics," *Byull. Moskov. Univ., Sect. A*, Vol. 1, No. 8, pp. 1-25, 1938

[1.6] N. N. Krasovskii, "On periodical solution of differential equations involving a time lag," *Dokl. Akad. Nauk SSSR*, Vol. 114, pp. 252-255, 1957

[1.7] J. K. Hale, *Asymptotic behavior of the solutions of differential-difference equations*, RIAS Tech. Rept. No. 61-10, 1961

[1.8] *ibid*, "Functional-differential equations with parameters," *Contr. Diff. Eqs.*, Vol. 1, No. 4, pp. 401-410, 1962

[1.9] *ibid*, "Linear functional-differential equations with constant coefficients," *Contr. Diff. Eqs.*, Vol. 2, pp. 291-317, 1963

[1.10] *ibid, Averaging methods for differential equations with retarded arguments and small parameters*, Tech. Rept. No. 64-1, Center for Dynamic Systems, Div. Appl. Math., Brown Univ., Providence, R. I., 1964

[1.11] *ibid*, "Periodic and almost periodic solutions of functional-differential equations," *Arch. Rational Mech. Anal*, Vol. 15, No. 4, pp. 289-304, 1964

[1.12] *ibid*, "Some problems in the qualitative theory of systems with time lag; qualitative methods in the theory of nonlinear vibrations," *Proc. Internat. Symp. Nonlinear Vibrations*, Vol. 11, pp. 394-408, Butterworths, London, 1961

[1.13] R. D. Driver, "Existence theory for a delay-differential system," *Contr. Diff. Eqs.*, Vol. 1, pp. 317-366, 1961

[1.14] *ibid*, "Existence and stability of solutions of a delay-differential system," *Arch. Rational Mech. Anal.*, Vol. 10, pp. 401-426, 1962

[1.15] *ibid*, "Existence and continuous dependence of solutions of a neutral functional-differential equation," *Arch. Rat. Mech. Anal.*, Vol. 19, pp. 144-166, 1965

[1.16] V. Lakshmikantham, "Lyapunov function and a basic inequality in delay-differential, equations," *Arch. Rational Mech. Anal.*, Vol. 10, pp. 305-310, 1962

[1.17] *ibid*, "Functional-differential systems and extensions of Lyapunov's method," *F. Math. Anal. Appl.*, Vol. 8, No. 3, pp. 392-405, 1964

[1.18] D. S. Jones, "Periodic motions in Banach space and applications to functional-differential equations, *Contr. Diff. Eqs.*, Vol. 3, pp. 75-106, 1963

[1.19] *ibid*, "Asymptotic fixed point theorems and periodic systems of functional-differential equations," *Contr. Diff. Eqs.*, Vol. 2, pp. 385-405, 1963

[1.20] *ibid*, "Fundamental inequalities for discrete and discontinuous functional equations," *F. Soc. Ind. Appl. Math.*, Vol. 12, pp. 43-57, 1964

[1.21] R. Bellman and K. L. Cooke, *Differential-Difference Equations*, Academic Press, 1963

[1.22] R. Bellman, *A survey of the theory of the boundedness stability, and asymptotic behavior of solutions of linear and nonlinear differential and difference equations*, Office of Naval Research, Washington, D.C., 1949

[1.23] R. Bellman and J. M. Danskin, *A survey of mathematical theory of time lag, retarded control and hereditary processes*, RAND Corp. Rept. No. R-256, 1956

[1.24] R. Bellman, J. M. Danskin and I. Glicksberg, *A bibliography of the theory and application of differential-difference equations*, RAND Corp. Res. Memo. No. RM-688, 1952

[1.25] L. E. Elsgolts, "Qualitative Methods in Mathematical Analysis," *AMS* Translation of *Math. Monographs*, Vol. 12, 1964

[1.26] W. Hahn, "Bericht uber Differential-Differenzengleichungen mit festen und veranderlichen Spannen," *Fahresber.*, DMV, Vol. 57, pp. 55-84, 1954

[1.27] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags*, Academic Press, New York, 1966

[1.28] N. N. Krasovskii, *Certain Problems in the Theory of Stability of Motion*, Moscow, 1959; English translation: *Stability of Motion*, Stanford Univ. Press, Stanford, Calif., 1963

[1.29] E. M. Wright, "The linear difference-differential equation with constant coefficients," *Proc. Roy. Soc. Edinburgh, Sect. A*, Vol. 62, Part IV (No. 40), pp. 387-393, 1949

[1.30] *ibid*, "The stability of solutions of nonlinear difference-differential equations," *Proc. Roy. Soc. Edinburgh, Sect. A*, Vol. 63, Part I (No. 2), pp. 18-26, 1950

[1.31] *ibid*, "A nonlinear difference-differential equation," *F. Reine Angew. Math.*, Vol. 194, pp. 66-87, 1955

[1.32] W. I. Zubov, *Vibration in Nonlinear Control Systems*, Leningrad, 1962

[1.33] A. Callender, D. R. Hartree and A. Porter, "Time lag in a control system," *Trans. R. Soc.*, Series A, Col. 235, pp. 415-444, 1935 and op. cit., Vol. 161, pp. 460-476, 1937

[1.34] H. Bateman, "The control of an elastic fluid," *Bull. Am. Math. Soc.*, Vol. 51, pp. 601-646, 1964. Also in *Selected Papers on Mathematical Trends in Control Theory*, R. Bellman and R. Kalala (Eds.), (pp. 18-64) Dover, 1964

[1.35] R. Weiss, "Transportation lag - an annotated bibliography," *IRE Trans. on Aut. Control*, Vol. AC-4, pp. 56-68, 1959

[1.36] N. H. Chosky, "Time-lag systems - a bibliography," *IRE. Trans. on Aut. Control*, Vol. AC-5, pp. 66-70, 1960

[1.37] M. N. Oguztoreli, *Time-Lag Control Systems*, Academic Press, 1966

[1.38] P. J. Dhrymes, *Distributed Lags*, Oliver & Boyd, London, 1971

[1.39] J. E. Marshall, *Control of Time-Delay Systems*, IEE Series on Control Engineering, Vol. 10, London, 1979

[1.40] M. Jamshidi and A. T. Bahill, "Book Review: Control of Time-Delay Systems by J. E. Marshall," *IEEE Control Systems Magazine*, Vol. 2, 1982

[1.41] L. A. Zadeh and C. A. Desoer, *Linear System Theory*, McGraw-Hill, 1963

[1.42] J. C. Maxwell, "On Governors," *Proc. Royal Soc. of London*, Vol. 16, pp. 270-283, 1868. Also in *Selected Papers on Mathematical Trends in Control Theory*, R. Bellman and R. Kalala (Eds.), pp. 3-17, Dover, 1964

[1.43] O. Mayr, *The Origins of Feedback Control*, M. I. T. Press, 1970 (Also in *Scientific American*, Vol. 223, pp. 110-118, 1970)

[1.44] E. S. Armstrong and J. S. Tripp, "An application of multivariable design techniques to the control of the National Transonic Facility." *NASA Tech. paper* 1887, 1981

[1.45] A. Z. Manitius, "Feedback controllers for a wind tunnel model involving a delay: analytical design and numerical simulation," *IEEE Trans. Aut. Contr.*, Vol. AC-29, pp. 1058-1068, 1984

Part I MODELING

CHAPTER 2

MATHEMATICAL DESCRIPTION OF TIME-DELAY SYSTEMS

2.1 INTRODUCTION

The first step in studying a physical system is to derive mathematical equations which describe it. This mathematical description should be computationally convenient and, at the same time, should adequately represent the system. This chapter is concerned with developing such mathematical descriptions for TD systems.

If only the terminal behavior of the system is of interest, an input-output description or a transfer function will be adequate to represent the system. But if the internal behavior of the system is also of interest, then a state-space representation of the system will be useful. These concepts will be discussed in this chapter. We will be mainly concerned with the mathematical description of *linear* TD systems. However, linearization of nonlinear TD systems will also be discussed. Methods of modeling and decomposition of large-scale TD systems will be covered in the final sections of this chapter.

2.2 MODELING OF TIME-DELAY SYSTEMS

A *model* is an idealization of a physical system. It is used to organize and/or reduce the computational effort required in the analysis and the design of systems. Developing an adequate model for a complex physical system is an important task. It requires the ability to determine which physical variables or relationships are crucial to the accuracy of the model and which ones can be neglected. This immediately implies that a physical system may have different models depending on the questions of interest about it. For instance, a transistor has different models depending on the amplitude and the frequency of the signals applied to it.

A mathematical representation of a physical system is derived by mathematically describing the interactions among different system components. This is done by the application of physical laws governing such interactions to the models of system components. For example, an electric circuit whose components have been modeled as resistors, capacitors and inductors is mathematically characterized by the applications of the Kirchoff's laws.

Any physical system has some time delay(s) associated with it due to its distributed nature. However, in many cases the delay(s) are negligible and can be ignored in the system model. In other cases, a physical system may have several different delays but their accumulated effects may be included in the system model as a single time delay for simplicity. These approximations of course, introduce some error in the analysis and the design of systems. The magnitudes of these errors often depend directly on the size of the (neglected) delays. Sensitivity of system's behavior to time delay will be discussed in Chapter 7.

In this book we are concerned with physical systems which include delays in their mathematical representations. From this point on, by a *system* we mean the mathematical representation of a physical system.

2.3 STATE-SPACE REPRESENTATION

Associated with each system are two important sets of attributes. One set, denoted by u_1, u_2, ... u_r, indicates the *inputs* to the system or the *control variables*. The inputs represent the stimuli which are applied to the system from external sources. The other set, denoted by y_1, y_2, ... y_m, indicates the *outputs* of the system. The outputs of a system can be measured externally. The inputs and the outputs of a system are related through the dynamic equations describing the system.

There is a third set of attributes associated with each system, namely the *state*. Loosely speaking, the state of a system is a collection of information which contains the history of the system; that is, the knowledge of the state and the input(s) of a system will be adequate to calculate its output(s). More precisely, we have the following definition.

1. **Definition.** A state of a system at time t_o is a collection of information which together with the knowledge of the input(s) for $t \geq t_o$ is sufficient to uniquely determine the output(s) of the system for $t \geq t_o$. Δ

For continuous-time TD systems, the state at time t is defined over an interval $[t',t]$ where t' depends on the delay(s) present in the system. An n-dimensional real-valued vector $\mathbf{x}(t)$ will be used to express the state of the system at time t. Such a vector is called a *state vector* and its components $x_i(t)$, $i = 1,2,\ldots,n$, are called *state variables*. If Δ is the largest time delay in the system, then the knowledge of $\mathbf{x}(t)$ over $(t_1-\Delta,t_1)$ plus the knowledge of the control vector $\mathbf{u}(t)$ over $[t_1,t_2]$ are necessary and sufficient to determine the state $\mathbf{x}(t)$ or the output y (t) for all $t_2 > t_1$ [2.2]. Thus for continuous-time TD systems the *initial state* (or the *initial function)* ϕ(t) must be given for $t\epsilon[t_o-\Delta,t_o]$ where t_o is the initial time of observation of the system.

Note that the state of a system is not necessarily unique. Figure 1 illustrates a system in which the vector $\mathbf{u} = [u_1,u_2,\ldots,u_r]'$ is the *input vector* or the *control vector*, and the vector $\mathbf{y} = [y_1,y_2,\ldots,y_m]'$ is the *output vector*. These vectors belong to vector spaces of proper dimensions. That is, $\mathbf{u}\epsilon U$ and $\mathbf{y}\epsilon Y$ where the vector spaces U and Y, *the input space* and the *output space*, have dimensions r and m, respectively. In the case $r = m = 1$, the system is called a *single-input single-output* (SISO) or a *scalar* system. For r > 1 and/or m > 1 , the system is referred to as a *multi-input multi-output* (MIMO) or *multivariable* system.

The *state space* of a continuous-time TD system is an infinite-dimensional vector space. It is the set of n-dimensional vectors functions defined as follows:

$$\sum = \{\mathbf{x}(\theta),\, t-\Delta \leq \theta \leq t\} \tag{1}$$

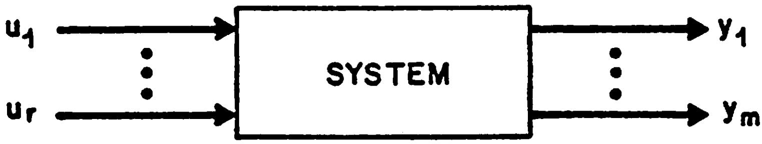

Figure 1. Illustration of a System with r Inputs and m Outputs

where Δ is the largest time delay in the system. Note that the space of n-dimensional vectors $\mathbf{x}$ at each instant of time t (sometimes referred to as the *phase space* of the system) is a subspace of Σ. In fact the state space of a continuous-time TD system is a Banach space of continuous functions over a time interval of length Δ, mapping the interval $[t-\Delta,t]$ into R^n. This is denoted as

$$\Sigma = C([t-\Delta,t],R^n) \tag{2}$$

Using the concept of state makes it possible to describe any TD system by a vector differential and/or difference equation. Such an equation is referred to as the *state equation* of the system. For continuous-time TD systems the state equation in the most general case has the following form:

$$\begin{aligned}\dot{\mathbf{x}}(t) = \mathbf{f}(&\mathbf{x}(t),\mathbf{x}(t-h_{x1}),\mathbf{x}(t-h_{x2}),\ldots,\\ &\mathbf{x}(t-h_{xN}),\mathbf{u}(t),\mathbf{u}(t-h_{u1}),\mathbf{u}(t-h_{u2}),\ldots,\mathbf{u}(t-h_{uR}),t)\end{aligned} \tag{3}$$

where $\mathbf{f}$ is a nonlinear vector-valued function and $h_{xi} > 0$, $i = 1,2$,..., N; $h_{ui} > 0, i = 1,2,\ldots,R$ represent the state and control delays in the system, respectively. The *output equation* expresses the output vector $\mathbf{y}(t)$ as a function of the state and control vectors:

$$y(t) = g(x(t), x(t-h_{x1}), x(t-h_{x2}), \ldots,$$
$$x(t-h_{xN}), u(t), u(t-h_{u1}), u(t-h_{u2}), \ldots, u(t-h_{uR}), t) \tag{4}$$

where g, in general, is a nonlinear vector-valued function. If the system is *time-invariant* (or *stationary*), then the functions f and g in (3) and (4) will not explicitly depend on time t.

For linear TD systems, the state and output equations will have the forms

$$\dot{x}(t) = A(t)x(t) + \sum_{i=1}^{N} A_i(t)x(t-h_{xi}) + B(t)u(t)$$
$$+ \sum_{i=1}^{R} B_i(t)u(t-h_{ui}) \tag{5}$$

$$y(t) = C(t)x(t) + \sum_{i=1}^{N} C_i(t)x(t-h_{xi}) + D(t)u(t)$$
$$+ \sum_{i=1}^{R} D_i(t)u(t-h_{ui}) \tag{6}$$

If, in addition, the system is time-invariant, then matrices A, B, C, D; A_i, C_i, $i = 1,2,\ldots,N$; and B_i, D_i, $i = 1,2,\ldots,R$ will be constant.

The foregoing discussion can be extended to the case of discrete-time TD systems. In such systems the input, the output and the state are described only at discrete instants of time (referred to as the *sampling instants).* We will assume that the sampling period is fixed and that the delays are integral multiples of the sampling period. Then the state and output equations will be similar to the continuous-time case except instead of the first-order derivative of the state vector, we will have advance by one sampling period. That is,

$$x(k+1) = f(x(k), x(k-1), x(k-2), \ldots,$$
$$x(k-N), u(k), u(k-1), u(k-2), \ldots, u(k-R), k) \tag{7}$$
$$y(k) = g(x(k), x(k-1), (k-2), \ldots,$$
$$x(k-N), u(k), u(k-1), u(k-2), \ldots, u(k-R), k) \tag{8}$$

which should be compared to (3) and (4) for the continuous-time case. Note that (7) and (8) are pure difference equations. For stationary linear discrete-time TD systems, the state and output equations will have the forms

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \sum_{i=1}^{N} \mathbf{A}_i(k)\mathbf{x}(k-i) + \mathbf{B}(k)\mathbf{u}(k) + \sum_{i=1}^{R} \mathbf{B}_i(k)\mathbf{u}(k-i) \tag{9}$$

$$\mathbf{y}(k) = \mathbf{C}(k)\mathbf{x}(k) + \sum_{i=1}^{N} \mathbf{C}_i(k)\mathbf{x}(k-i) + \mathbf{D}(k)\mathbf{u}(k) + \sum_{i=1}^{R} \mathbf{D}_i(k)\mathbf{u}(k-i) \tag{10}$$

If, in addition, the system is time-invariant, the matrices will not depend on k (*i.e.* they will be constant matrices). Note that discrete-time TD systems in which the sampling period is not fixed or the delays are not integral multiples of the sampling period, cannot be characterized by pure difference equations. Such systems are described by differential-difference equations. That is, their state and output equations are as in (5) and (6).

In contrast to continuous-time TD systems whose state spaces are infinite-dimensional, the state spaces of discrete-time TD systems are finite-dimensional. This is due to the fact that in discrete-time TD systems the state vector at each sampling instant has a finite number of elements where each element is a function specified at only a finite number of time instants (rather than on an interval). We will further illustrate this point in the next section.

2.4 FREQUENCY-DOMAIN REPRESENTATION

One of the methods of describing a *l.t.i.* system in the input-output form is using the transfer function. The transfer function relates the Laplace or the z transform of the output to the Laplace or the z transform of the input. For a *l.t.i.* system with r inputs and m outputs the transfer function is an $m \times r$ matrix. We will discuss the transfer functions of continuous-time, discrete-time and hybrid *l.t.i.* TD systems separately in the following subsections.

2.4.1 Continuous-Time Time-Delay Systems

For SISO continuous-time systems the transfer function is defined as the ratio of the Laplace transform of the output and the Laplace transform of the input when all the initial conditions are zero. The transfer function can be expressed in terms of the matrices in the state and output equations of the system. For the purpose of illustration, consider a SISO stationary TD system with a single delay in the state. That is,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}u(t),\ t \geqslant t_o \tag{1}$$

$$y(t) = \mathbf{C}\mathbf{x}(t) + Du(t) \tag{2}$$

Taking the Laplace transform of the above equations yields

$$s\mathbf{X}(s)-\mathbf{X}(t_o) = \mathbf{A}\mathbf{X}(s) + \mathbf{A}_1\mathbf{X}(s)e^{-hs} + \mathbf{B}U(s) \tag{3}$$

$$Y(s) = \mathbf{C}\mathbf{X}(s) + DU(s) \tag{4}$$

Considering zero initial state, we have Y(s) = H(s) U(s) where

$$H(s) \triangleq \mathbf{C}(s\mathbf{I}-\mathbf{A}-\mathbf{A}_1e^{-hs})^{-1}\mathbf{B} + D \tag{5}$$

is a scalar function known as the *transfer function* of the system. Note that H(s) in (5), unlike the case of nondelay systems, is not a rational function of s. The *poles* of the system are those (complex) values of s for which *H(s)* tends to infinity. For nondelay systems, *i.e.* when $\mathbf{A}_1 = \mathbf{0}$, the poles of the system will be identical to the eigenvalue of the system matrix A, *i.e.* they will be the roots of $det(s\mathbf{I}-\mathbf{A}) = 0$. For TD systems, however, the poles will be the (complex) values of s which satisfy

$$det(s\mathbf{I}-\mathbf{A}-\mathbf{A}_1e^{-hs}) = 0 \tag{6}$$

For $h \neq 0$, (6) has an infinite number of roots. (These are called the *spectrum* of the TD system.) This observation is consistent with the fact that the state space of (a continuous-time) TD system is infinite-dimensional.

1. **Example.** Consider the SISO first-order TD system described by

$$\dot{x}(t) = -Kx(t-h) + u(t) \tag{7}$$

$$y(t) = Kx(t) \tag{8}$$

where K is a positive constant. Equations (7) and (8) are the state and output equations of a system with an integrator with gain K in the forward path and a time delay h in the feedback path whose block diagram is shown in Figure 1. The transfer function of this system is

$$H(s) \triangleq \frac{Y(s)}{U(s)} = \frac{K}{s+Ke^{-sh}} \tag{9}$$

The poles of the system are the roots of

$$s+Ke^{-sh} = 0 \tag{10}$$

Let

$$s = \sigma + j\omega = \rho e^{j\phi} \text{ where } \rho = \sqrt{\sigma^2+\omega^2}\,,\ \phi = \tan^{-1}\frac{\omega}{\sigma} \tag{11}$$

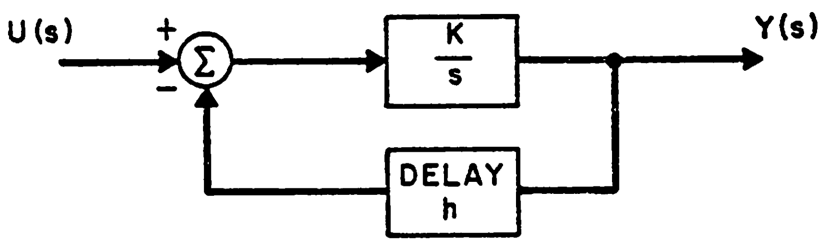

Figure 1. Block Diagram for Example 2.4.1

Then from (10) we have

$$\rho e^{j\phi} e^{(\sigma+j\omega)h} = -K = K e^{j\pi(1+2k)},\ k = 0, \pm 1, \pm 2, \cdots \tag{12}$$

Using (11) and equating the magnitudes and the exponents in (12) yields

$$\sqrt{\sigma^2+\omega^2}\, e^{\sigma h} = K \tag{13}$$

$$\tan^{-1}\frac{\omega}{\sigma} + \omega h = \pi(1+2k),\ k = 0, \pm 1, \pm 2, \cdots$$

The values of $s = \sigma + j\omega$ which satisfy (10) can be found from (13). Note that an infinite number of such values exist. Δ

The above discussion on the transfer function of SISO continuous-time TD systems can be extended to the case of MIMO systems with multiple delays in state and control. (See Problem 2.3.) Note that for MIMO TD systems with r inputs and m outputs, the transfer function $\mathbf{H}(s)$ is an $m \times r$ matrix, referred to as the *transfer function matrix.*

2.4.2 Discrete-Time Time-Delay Systems

The development of the transfer function for the discrete-time TD systems is analogous to that for the continuous-time TD systems. The transfer function of a stationary SISO TD system (sometimes referred to as the z-transfer function) is defined as the ratio of the z transform of the output and the z transform of the input when all initial conditions are zero. It can be expressed in terms of the matrices in the state and output equations of the system. To illustrate this consider a SISO stationary TD system with a single delay of ℓ sampling periods in the state. That is,

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{A}_1\mathbf{x}(k-\ell) + \mathbf{B}u(k),\ k \geqslant k_o \tag{14}$$

$$y(k) = \mathbf{C}\mathbf{x}(k) + Du(k) \tag{15}$$

Note that in (14) ℓ is an integer. Taking the z transform of (14) and (15) yields

$$z\mathbf{X}(z) - z\mathbf{x}(k_o) = \mathbf{A}\mathbf{X}(z) + \mathbf{A}_1 z^{-\ell}\mathbf{X}(z) + \mathbf{B}U(z) \tag{16}$$

$$Y(z) = \mathbf{C}\mathbf{X}(z) + DU(z) \tag{17}$$

Considering zero initial state, we have $Y(z) = H(z)U(z)$ where

$$H(z) \triangleq \mathbf{C}(z\mathbf{I}-\mathbf{A}-\mathbf{A}_1 z^{-\ell})^{-1}\mathbf{B} + \mathbf{D} \tag{18}$$

An important difference between H(z) in (18) and the analogous $H(s)$ for the continuous-time TD systems in (5) is that $H(z)$ is a *rational* function. Here again, the poles of the system are defined as those (complex) values of z for which $H(z)$ tends to infinity. That is, the values of z that satisfy

$$det(z\mathbf{I}-\mathbf{A}-\mathbf{A}_1 z^{-\ell}) = 0 \tag{19}$$

Note that (19) has a finite number of roots. This verifies the previous observation of the fact that the state space of a discrete-time TD system is finite-dimensional.

The above discussion can be extended to MIMO systems with multiple delays in state and control. (See Problem 2.4.) Note that for MIMO TD systems with r inputs and m outputs, the transfer function $\mathbf{H}(z)$ is an $m \times r$ matrix. The case where the delays are not integer multiples of the sampling period will be discussed in Section 2.4.3.

2. **Example.** Consider the MIMO discrete-time TD system described by the following state and output equations:

$$\mathbf{x}(k+1) = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}(k-\ell) + \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} u(k) \tag{20}$$

$$\mathbf{y}(k) = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}(k), \quad k \geqslant 0 \tag{21}$$

where $\ell \geqslant 1$ is an integer. The transfer function matrix of this system is

$$\begin{aligned} \mathbf{H}(z) &= \mathbf{C}(z\mathbf{I}-\mathbf{A}-\mathbf{A}_1 z^{-\ell})^{-1}\mathbf{B} \\ &= \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} z+1-z^{-\ell} & -1 & 0 \\ 0 & z+1-z^{-\ell} & 0 \\ 0 & 0 & z+3-z^{-\ell} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{2}{z+1-z^{-\ell}} \\ \dfrac{2}{z+1-z^{-\ell}} \end{bmatrix} \end{aligned} \tag{22}$$

For $\ell=2$, the poles of the system are the roots of $z^3 + z^2-1 = 0$ which are

$$z_1 = -0.877 + j0.745,\ z_2 = -0.877 - j0.745,\ z_3 = 0.755$$

2.4.3 Hybrid Systems

Hybrid, or mixed, systems refer to those in which part of the system is continuous-time and part of it is discrete-time. Such systems include *sampled-data systems* and *digital control systems*. In general, a system in which at one or more points signals are received only at discrete intervals of time is called a sampled-data system. Digital control systems which may be considered as a special class of sampled-data systems are those in which at one or more points, signals are expressed in a numerical code for digital transducers (usually digital computers).

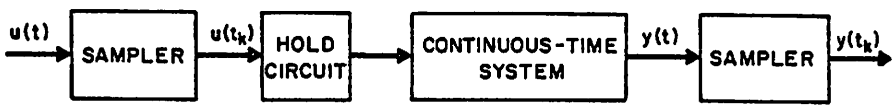

Figure 2. An Open-Loop Sampled-Data System

Figure 2 shows an open-loop sampled-data system. In this system the input $u(t)$ is sampled at regular intervals to obtain input samples $u(t_k) \triangleq u(k)$. These samples, which are held constant during each subinterval by a hold circuit, are applied to the continuous-time system which produces continuous output $y(t)$. The output is similarly sampled to obtain samples $y(k)$. This scheme is used in systems where a digital computer is employed to calculate the proper input at each instant of time. Figure 3 shows the block diagram of a digital control system. The A/D block is an analog-to-digital converter (sampler) and the D/A block is a digital-to-analog converter (hold circuit). The digital computer receives output samples, compares them with a specified reference and calculates the input at each instant t_k according to a control strategy for which it is programmed.

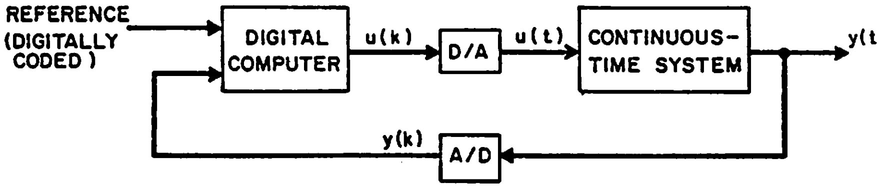

Figure 3. Block Diagram of a Digital Control System

The modified z transform technique can be used to study hybrid systems. Let us first consider the sampled-data system shown in Figure 4. Here, $\mathbf{u}^*(t)$ and $\mathbf{y}^*(t)$ represent the samples of the input $\mathbf{u}(t)$ and output $\mathbf{y}(t)$. The Laplace transforms of signals are also indicated in Figure 4. The Laplace transform of the output $y(t)$ is

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{U}^*(s) \tag{23}$$

But the Laplace transform of the sampled output $y^*(t)$ is

$$\mathbf{Y}^*(s) = \mathbf{G}^*(s)\mathbf{U}^*(s) \tag{24}$$

where $G^*(s)$ indicates the Laplace transform of the sampled function $\mathbf{g}^*(t)$. (See Problem 2.8.) Equation (24) is referred to as the *pulse transform* of (23).

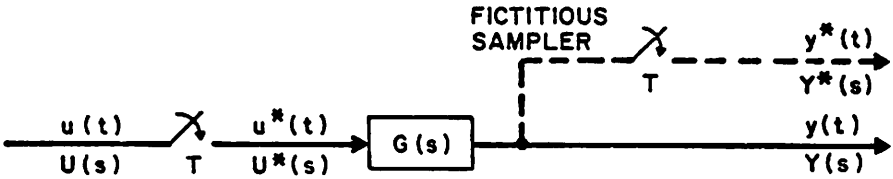

Figure 4. A *l.t.i* Continuous-Time System with Sampled Data

The z transform of the output $y(t)$ can be obtained by changing the variable s to $T^{-1}\ell nz$ in (24):

$$Y(z) = G(z)U(z) \tag{25}$$

which describes the output at only the sampling instants. (See (2.4.4).) The relationship $Y(z) = [Y^*(s)]_{s=T^{-1}\ln z}$ is sometimes indicated as $Y(z) = Z[Y(s)]$ for convenience. It is equivalent to using the corresponding entries for $y(t)$ in a table of both the Laplace and the z transforms.

In hybrid systems, where the output is a function of *continuous* time, the modified z transform can be used to determine the output. If a fictitious delay $\Delta T = (1-m)T$ is inserted in the system of Figure 4, as shown in Figure 5, we can write

$$Y_m^*(s) = G_m^*(s)U_m^*(s) \tag{26}$$

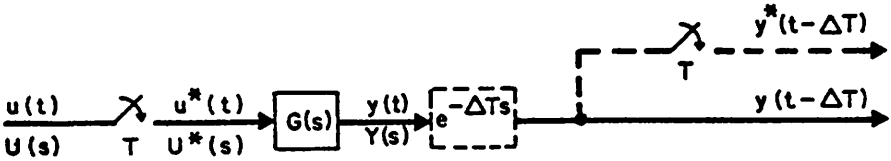

Figure 5. A *l.t.i.* Sampled-Data System with a Fictitious Time-Delay

where $Y_m{}^*(s) \triangleq L[y(t-T+mT)]$ and $0 \leqslant m \leqslant 1$. Equation (26) is referred to as the *modified pulse transform* of (23). Therefore, the modified z transform of the output can be written as

$$Y(z,m) = G(z,m)U(z) \tag{27}$$

where $G(z,m) = [G_m{}^*(s)]_{s=T^{-1}\ln z}$ is sometimes referred to as the *modified z-transfer function* of the system. By varying the value of $m = 1-\Delta$ in (27) between 0 and 1 and calculating the inverse modified z transform of $Y(z,m)$, we can obtain virtually all information about $y(t)$.

Now consider the sampled-data feedback system shown in Figure 6. We have

$$E(s) = U(s)-G(s)H(s)E^*(s) \tag{28}$$

whose pulse transform is

$$E^*(s) = U^*(s)-GH^*(s)E^*(s) \tag{29}$$

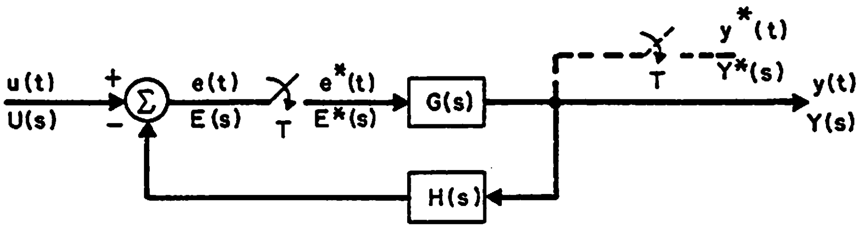

Figure 6. A *l.t.i.* Closed-Loop Sampled-Data System

where $GH^*(s)$ indicates the Laplace transform of the sampled signal $(L^{-1}[G(s)H(s)])^*$. Therefore we have

$$Y(s) = G(s)E^*(s) = \frac{G(s)U^*(s)}{1+GH^*(s)} \tag{30}$$

Taking the pulse transform of (30) yields

$$Y^*(s) = G^*(s)E^*(s) = \frac{G^*(s)U^*(s)}{1+GH^*(s)} \tag{31}$$

Replacing s by $T^{-1}\ell nz$ in (31) the z transform of the output $y(t)$ can be obtained:

$$Y(z) = \frac{G(z)U(z)}{1+GH(z)} \tag{32}$$

where $GH(z)$ indicates the z transform of $(L^{-1}[G(s)H(s)])^*$, *i.e.*, $GH(z) = Z[GH(s)]$. The term $\frac{G(z)}{1+GH(z)}$ in (32) represents the overall (closed-loop) z-transfer function of the system in Figure 6.

Again, the modified z transform technique can be used to determine the output between the sampling instants. To find the modified z transform of the output in Figure 6, introduce a fictitious time delay of ΔT in the forward path and a fictitious time advance of ΔT in the feedback path as shown in Figure 7. Then we have

$$Y(s) = G(s)E^*(s)e^{-\Delta Ts} \tag{33}$$

whose modified z transform yields

$$Y(z,m) = G(z,m)E(z) \tag{34}$$

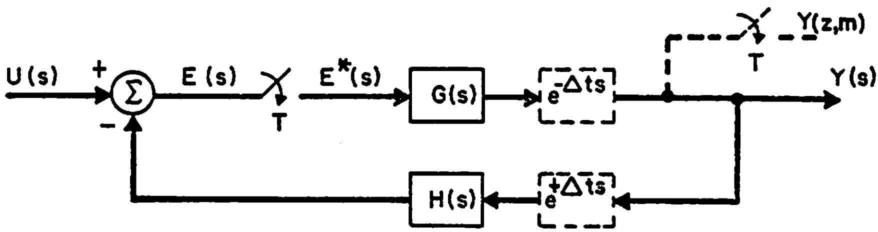

Figure 7. A Closed-Loop Sampled-Data System with Fictitious Delay and Advance for Modified z-Transform Analysis

Note that the delay and the advance compensate each other around the loop. Thus (28) and (29) hold and we have

$$E(z) = \frac{U(z)}{1+GH(z)} \tag{35}$$

Therefore, from (34),

$$Y(z,m) = \frac{G(z,m)}{1+GH(z)}\,U(z) \tag{36}$$

Expression $\frac{G(z,m)}{1+GH(z)}$ represents the modified z-transfer function of the system of Figure 7.

The modified z transform method can similarly be used to analyze hybrid TD systems in which the delay is not an integral multiple of the sampling period. If the system has a delay of kT where k is a noninteger, write k as $\ell + \Delta$ where ℓ is the largest integer less than n and $0 < \Delta < 1$. Then the integer delay part ℓT produces a factor $z^{-\ell}$ in the modified z transform and the nonintegral delay ΔT can be handled separately. In fact we have

$$Z[f(t-kT)] = z^{-\ell}Z[f(t-\Delta T)] \tag{37}$$

The term $Z[f(t-\Delta T)]$ where $0 < \Delta < 1$ is called the *delayed z transform* of *f(t)*. It is sometimes indicated as $F(z,\Delta)$ where $F(z,\Delta) = [F(z,m)]_{m=1-\Delta}$

3. **Example.** Consider the TD sampled-data feedback system shown in Figure 8. Let $G(s) = \dfrac{1}{s+1}$ and $k = 1.4$. Thus, using the above notation we will have $\ell = 1$ and $\Delta = 0.4$. By (36), the closed-loop transfer function of the system is $H(z) = \dfrac{G_1(z,m)}{1+G_1(z)}$ where $G_1(s) = G(s)e^{-kTs} = \dfrac{e^{-1.4Ts}}{s+1}$. Thus

$$G_1(z,m) = z^{-1}G(z,m) = z^{-1}Z_m\left[\frac{1}{s+1}\right] = Z^{-1}\left[\frac{e^{-mT}}{z-e^{-T}}\right] \tag{38}$$

and

$$G_1(z) = [G_1(z,m)]_{m=1-0.4} = G_1(z,0.6) = \frac{e^{-0.6T}}{z-e^{-T}} \tag{39}$$

Therefore,

$$H(z) = \frac{e^{-mT}}{z(z-e^{-T}+e^{-0.6T})} \tag{40}$$

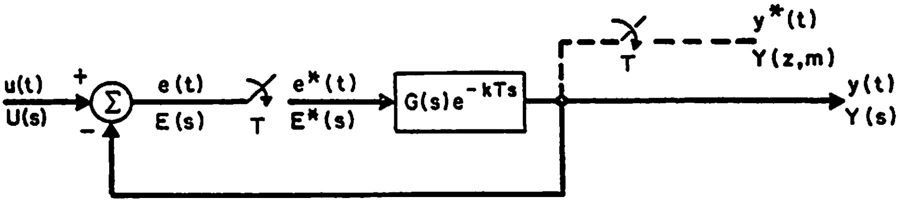

Figure 8. A TD Sampled-Data Feedback System

2.5 LINEARIZATION OF NONLINEAR TIME-DELAY SYSTEMS

Consider the state-space representation of a nonlinear TD system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t-h_x), \mathbf{u}(t), \mathbf{u}(t-h_u)) \tag{1}$$

with initial functions

$$\mathbf{x}(t) = \mathbf{x}_0(t), \ -h_x \leqslant t \leqslant 0 \tag{2a}$$

$$\mathbf{u}(t) = \mathbf{u}_0(t), \ -h_u \leqslant t \leqslant 0 \tag{2b}$$

where $\mathbf{x}$ and $\mathbf{u}$ are $n \times 1$ and $r \times 1$ state and control vectors, respectively, and h_x and h_u are delays in state and control. Assume that the vector-valued function $\mathbf{f}(\cdot)$ is continuously differential *w.r.t.* all its arguments.

Suppose that a nominal input $\mathbf{u}_n(t)$ and the resulting nominal state $\mathbf{x}_n(t)$ are known. They may be available from the system's previous history or as a result of a computer solution. It is, however, understood that $\mathbf{x}_n(t)$ and $\mathbf{u}_n(t)$ satisfy,

$$\dot{\mathbf{x}}_n(t) = \mathbf{f}(\mathbf{x}_n(t), \mathbf{x}_n(t-h_x), \mathbf{u}_n(t), \mathbf{u}_n(t-h_u)) \tag{3}$$

It is desired to find a solution $\mathbf{x}(t)$ as a result of application of a control $\mathbf{u}(t)$ which is slightly different from $\mathbf{u}_n(t)$. Let these approximate vectors be defined by

$$\mathbf{u}(t) = \mathbf{u}_n(t) + \delta\mathbf{u}(t) \tag{4a}$$

$$\mathbf{x}(t) = \mathbf{x}_n(t) + \delta\mathbf{x}(t) \tag{4b}$$

where $\delta\mathbf{x}(t)$ and $\delta\mathbf{u}(t)$ are $n\times 1$ and $r\times 1$ vectors of state and control *perturbations*, respectively. Using (4), (1) can be rewritten as

$$\begin{aligned}\dot{\mathbf{x}}(t) = \dot{\mathbf{x}}_n(t) + \delta\dot{\mathbf{x}}(t) = \mathbf{f}(&\mathbf{x}_n(t) + \delta\mathbf{x}(t),\ \mathbf{x}_n(t-h_x) \\ &+ \delta\mathbf{x}(t-h_x),\ \mathbf{u}_n(t) + \delta\mathbf{u}(t),\ \mathbf{u}_n(t-h_u) + \delta\mathbf{u}(t-h_u))\end{aligned} \tag{5}$$

The Taylor series expansion of function $\mathbf{f}(\cdot)$ about the nominal trajectories and truncating after the first terms lead to:

$$\begin{aligned}&\mathbf{f}(\mathbf{x}_n(t) + \delta\mathbf{x}(t),\ \mathbf{x}_n(t-h_x) + \delta\mathbf{x}(t-h_x),\ \mathbf{u}_n(t) + \delta\mathbf{u}(t),\ \mathbf{u}_n(t-h_u) + \delta\mathbf{u}(t-h_u)) \\ &\cong \mathbf{f}(\mathbf{x}_n(t),\ \mathbf{x}_n(t-h_x),\ \mathbf{u}_n(t),\ \mathbf{u}_n(t-h_u)) + \nabla_x\mathbf{f}(\cdot)|_n\delta\mathbf{x}(t) \\ &+ \nabla_{x(t-h_x)}\mathbf{f}(\cdot)|_n\delta\mathbf{x}(t-h_x) + \nabla_u\mathbf{f}(\cdot)|_n\delta\mathbf{u} + \nabla_{x(t-h_x)}\mathbf{f}(\cdot)|_n\delta u(t-h_u)\end{aligned} \tag{6}$$

where $\nabla_x\mathbf{f}(\cdot)$ is the *Jacobian matrix* of $\mathbf{f}(\cdot)$ with respect to $\mathbf{x}(t)$. It is noted that in (6) it has been assumed that vectors $\delta\mathbf{x}(t)$ and $\delta\mathbf{u}(t)$ are sufficiently small so that terms $\delta\mathbf{x}_i^2(t)$, $\delta\mathbf{u}_j^2(t)$, $\delta\mathbf{x}_i(t)\delta\mathbf{u}_j(t)$, $\delta\mathbf{x}_i^2(t-h_x)$, $\delta\mathbf{u}_j^2(t-h_u)$, $\delta\mathbf{x}_i(t-h_x)\delta\mathbf{u}_j(t-h_u)$, etc. are negligible. Now, considering (3) and (6), (5) will lead to [3.8]

$$\begin{aligned}\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) &+ \mathbf{A}_1(t)\delta\mathbf{x}(t-h_x) + \mathbf{B}(t)\delta\mathbf{u}(t) \\ &+ \mathbf{B}_1(t)\delta\mathbf{u}(t-h_u)\end{aligned} \tag{7}$$

where matrices $\mathbf{A}(t)$, $\mathbf{A}_1(t)$, $\mathbf{B}(t)$ and $\mathbf{B}_1(t)$ are defined by

$$\mathbf{A}(t) = \nabla_{\mathbf{x}}\mathbf{f}(\cdot)|_n\ ,\ \mathbf{A}_1(t) = \nabla_{\mathbf{x}(t-h_x)}\mathbf{f}(\cdot)|_n \tag{8a}$$

$$\mathbf{B}(t) = \nabla_{\mathbf{u}}\mathbf{f}(\cdot)|_n\ ,\ \mathbf{B}_1(t) = \nabla_{\mathbf{u}(t-h_u)}\mathbf{f}(\cdot)|_n. \tag{8b}$$

and the initial functions are given by

$$\delta\mathbf{x}(t) = \mathbf{x}_0 - \mathbf{x}_n(t), -h_x \leqslant t \leqslant 0 \tag{9a}$$

$$\delta\mathbf{u}(t) = \mathbf{u}_0(t) - \mathbf{u}_n(t), -h_u \leqslant t \leqslant 0 \tag{9b}$$

The system described by (7) - (9) represents a linear time-varying TD system whose more general form was described by (2.3.5). This perturbed system provides the most common scheme to study the properties and behavior of nonlinear systems (delayed or non-delayed) in the neighborhood of a known nominal solution [2.9]. This linear perturbation system will be used in Chapter 7 to find near-optimum control for nonlinear TD systems.

2.6 LARGE-SCALE TIME-DELAY SYSTEMS

In this section a brief overview of the mathematical description of large-scale TD systems is presented, while their optimization and control will be treated in Chapter 8.

Large-scale TD systems can be modeled in several ways depending on the structural behavior of the system. In the next four sections four possible mathematical representations of large-scale systems with time delays will be given.

2.6.1 Coupled Time-Delay Systems

Consider a large-scale TD system consisting of a set of N coupled subsystems

$$\begin{aligned}
\dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1,\mathbf{u}_1,\epsilon C_1(\mathbf{x},\mathbf{x}(t-h_x),\mathbf{u},\mathbf{u}(t-h_u),t) \\
\dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_2,\mathbf{u}_2,\epsilon C_2(\mathbf{x},\mathbf{x}(t-h_x),\mathbf{u},\mathbf{u}(t-h_u),t),t) \\
&\ \ \vdots \\
\dot{\mathbf{x}}_N &= \mathbf{f}_N(\mathbf{x}_N,\mathbf{u}_N,\epsilon C_N(\mathbf{x},\mathbf{x}(t-h_x),\mathbf{u},\mathbf{u}(t-h_u),t),t)
\end{aligned} \tag{1}$$

where $\mathbf{x} = [\mathbf{x}_1'\mathbf{x}_2' \ldots \mathbf{x}_N']'$ and $\mathbf{u} = [\mathbf{u}_1'\mathbf{u}_2' \ldots \mathbf{u}_N']'$ are state and control vectors, respectively, ϵ is a scalar coupling parameter, h_x and h_u are the delays, not necessarily *short*. The problem within the context of a near-optimum control is solved in Chapter 8. The linear version of the N coupled TD systems (1) is given by

$$\begin{aligned}
\dot{\mathbf{x}}_1(t) &= \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1\mathbf{u}_1(t) + \epsilon C_1(\mathbf{x}(t),\mathbf{x}(t-h_x),\mathbf{u}(t),\mathbf{u}(t-h_u),t) \\
\dot{\mathbf{x}}_2(t) &= \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{B}_2\mathbf{u}_2(t) + \epsilon C_2(\mathbf{x}(t),\mathbf{x}(t-h_x),\mathbf{u}(t),\mathbf{u}(t-h_u),t) \\
&\ \ \vdots \\
\dot{\mathbf{x}}_N(t) &= \mathbf{A}_N\mathbf{x}_N(t) + \mathbf{B}_N\mathbf{u}_N(t) + \epsilon C_N(\mathbf{x}(t),\mathbf{x}(t-h_x),\mathbf{u}(t),\mathbf{u}(t-h_u),t)
\end{aligned} \tag{2}$$

with initial functions

$$\mathbf{x}_i(t) = \phi_i(t), \ -h_x \leqslant t \leqslant 0 \tag{3a}$$

$$\mathbf{u}_i(t) = \eta_i(t), \ -h_u \leqslant t \leqslant 0 \tag{3b}$$

for $i=1,2,\ldots,N$.

In this way, as $\epsilon \to 0$, a large-scale TD system can be decoupled into N nondelayed isolated subsystems which are much smaller in scale (order) while being unretarded. As $\epsilon > 0$, the delay involved with subsystems are accounted for through its interactions with the other N-1 subsystems by virtue of $\epsilon C_i(\cdot)$ functions. It is noted that, as for ϵ, C_i needs not be small in magnitude. In Chapter 7 this coupling property will be used to find near-optimum control for a large-scale TD System[2.10].

2.6.2 Modeling of Time-Delay Systems via Singular Perturbation

Consider a linear multi-delay system

$$\dot{x}(t) = A_o x(t) + Bu(t) + \sum_{i=1}^{N} A_i x(t-ih) \tag{4}$$

with initial function

$$x(t) = \phi(t), \qquad -Nh \leqslant t \leqslant 0, \tag{5}$$

where A_i, $i = 0,1,\ldots,N$ and $\mathbf{B}$ are $n \times n$ and $n \times m$ constant matrices and h is the delay. The problem is to reduce (4)-(5) to an unretarded singularly perturbed system. The approximation, suggested independently by Sannuti [2.11] and Inoue et al [2.12], begins by dividing the delay h into k equal subintervals Δh such that $h = k\Delta h$, then let

$$\begin{aligned} y_1(t) &= x(t-\Delta h) \\ y_2(t) &= x(t-2\Delta h) \\ &\vdots \\ y_k(t) &= x(t-k\Delta h) = x(t-h). \end{aligned} \tag{6}$$

Differentiating the first equation in (6) yields

$$\dot{y}_1(t) = \dot{x}(t-\Delta h) \cong (x(t)-x(t-\Delta h))/\Delta h \tag{7}$$

or

$$\Delta h \dot{y}_1(t) = x(t) - y_1(t) . \tag{8a}$$

In a similar fashion

$$\begin{aligned}\Delta h \dot{\mathrm{y}}_2 &= \mathrm{y}_1 - \mathrm{y}_2 \\ &\vdots \\ \Delta h \dot{\mathrm{y}}_k &= \mathrm{y}_{k-1} - \mathrm{y}_k \,,\end{aligned} \tag{8b}$$

where each y_i, $i = 1,\ldots,k$ is itself an n-dimensional vector. If an $\ell = kn$ dimensional vector $\mathrm{Y} \triangleq (\mathrm{y}_1' \mathrm{y}_2' \cdots \mathrm{y}_k')'$ and a small scalar parameter $\mu \triangleq \Delta h = h/k$ are defined, then (8) can be rewritten as

$$\mu \dot{\mathrm{Y}} = \tilde{\mathrm{A}}_1 \mathrm{x} + \tilde{\mathrm{C}}_1 \mathrm{Y} \tag{9}$$

where

$$\tilde{\mathrm{A}}_1 = \begin{bmatrix} I \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \tilde{\mathrm{C}}_1 = \begin{bmatrix} -I & 0 & \cdot & \cdot & \cdot & 0 \\ I & -I & & & & \cdot \\ 0 & I & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & I & -I \end{bmatrix} \tag{10}$$

are $\ell \times n$ and $\ell \times \ell$ dimensional matrices and $\mathrm{I} = \mathrm{I}_n$ is an $n \times n$ identity matrix. The above formulation can be similarly applied to $\mathrm{x}(t-2h)$, $\mathrm{x}(t-3h)$,... up to $\mathrm{x}(t-Nh)$. Thus we define $\mathrm{w}_1 \triangleq \mathrm{x}(t-2\Delta h)$, ..., $\mathrm{w}_k \triangleq \mathrm{x}(t-2h)$, $\mathrm{p}_1 \triangleq \mathrm{x}(t-(N-1)\Delta h)$, $\mathrm{v}_1 \triangleq \mathrm{x}(t-N\Delta h)$,..., $\mathrm{v}_k \triangleq \mathrm{x}(t-N\mathrm{h})$. Then $\dot{\mathrm{x}}(t-2h)$,..., $\dot{\mathrm{x}}(t-Nh)$ can be expressed in a compact form as in (9):

$$\begin{aligned}\mu \dot{\mathrm{w}} &= \tilde{C}_{21} \mathrm{Y} + \tilde{C}_2 \mathrm{w} \\ &\vdots \\ \mu \dot{\mathrm{v}} &= \tilde{C}_{N,N-1} \mathrm{p} + \tilde{C}_N \mathrm{v},\end{aligned} \tag{11}$$

where

$$\tilde{C}_{j,j-1} = \begin{bmatrix} I & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & I & 0 & & & \cdot \\ 0 & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & I \end{bmatrix}, \quad \tilde{C}_j = \begin{bmatrix} -I & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -I & 0 & & & \cdot \\ 0 & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & -I \end{bmatrix} \tag{12}$$

for $j = 2,\ldots,N$. If an integer $N\ell = Nkn$ is defined to be the order of a large dimensional vector $\mathrm{z} \triangleq (\mathrm{Y}'\mathrm{w}' \ldots \mathrm{p}'\mathrm{v}')'$, then the linear TD system (4) is reduced to

$$\dot{\mathbf{x}} = \mathbf{A}_o\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{A}_1\mathbf{y}_k + \mathbf{A}_2\mathbf{w}_k + \cdots + \mathbf{A}_{N-1}\mathbf{p}_k + \mathbf{A}_N\mathbf{v}_k \tag{13a}$$

$$\mu\dot{\mathbf{z}} = \tilde{\mathbf{A}}\mathbf{x} + \tilde{\mathbf{C}}\mathbf{z} \tag{13b}$$

where $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{C}}$ defined by

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{\mathbf{A}}_1 \\ \hline \mathbf{0} \\ \hline \vdots \\ \mathbf{0} \end{bmatrix}, \quad \tilde{\mathbf{C}} \triangleq \begin{bmatrix} \tilde{\mathbf{C}}_1 & \mathbf{0} & \cdot & \cdot & \mathbf{0} \\ \tilde{\mathbf{C}}_{21} & \tilde{\mathbf{C}}_2 & & & \cdot \\ \mathbf{0} & & \cdot & & \cdot \\ \cdot & & & \cdot & \mathbf{0} \\ \mathbf{0} & \cdot & & \cdot\ \tilde{\mathbf{C}}_{N,N-1} & \tilde{\mathbf{C}}_N \end{bmatrix} \tag{14}$$

are $N\ell\times n$ and $N\ell\times N\ell$ dimensional matrices whose elements are defined by the corresponding block identity matrices in (10)-(12). The initial conditions are

$$\mathbf{x}(0) = \Phi(0)\triangleq\mathbf{x}_o \tag{15a}$$

$$\begin{aligned} \mathbf{z}(0) &= \{\mathbf{Y}'(0) \mid \mathbf{w}'(0)\cdots\mathbf{p}'(0) \mid \mathbf{v}'(0)\}' \\ &= \{\Phi(-\mu)\cdots\Phi(-k\mu) \mid \Phi(-2\mu)\cdots\Phi(-2k\mu) \\ &\quad \Phi(-N\mu)\cdots\Phi(-Nk\mu)\}\ . \end{aligned} \tag{15b}$$

Thus (13) is a large-scale singularly-perturbed system representing an approximation of the TD system (4). For a given value of h, a large enough k can be chosen so that $\mu = h/k$ is a small parameter. If $\mu = 0$, by the elimination of faster variables $\mathbf{w}$ and $\mathbf{v}$ in favor of slow variable $\mathbf{x}$, (11) is reduced from an $\nu = (N\ell+n) = (Nk+1)$ n-dimensional to an n-dimensional system

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} \tag{16}$$

where

$$\mathbf{A} = \mathbf{A}_o + \sum_{i=1}^{N} \mathbf{A}_i\mathbf{D}_{ik} \tag{17}$$

$\mathbf{D}_{ik}$ is an $n\times n$ dimensional matrix extracted from the kth n-dimensional row of $\mathbf{D}$ matrix, $\mathbf{z} = -\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{A}}\mathbf{x}\triangleq\tilde{\mathbf{D}}\mathbf{x}$. The accuracy of this approximation, which depends on k, will be discussed in due course in Chapter 8. In summary, an n-dimensional time-delay system is first transformed into a νth order unretarded large-scale singularly perturbed system.

1. **Example.** Consider a three-stand cold rolling mill shown in Figure 1. A linear two-delay model of this system is given by [2.13]:

$$\dot{\mathbf{x}} = \mathbf{A}_o(r)\mathbf{x} + \mathbf{B}(r)\mathbf{u} + \mathbf{A}_1\mathbf{x}(t-h_x) + \mathbf{A}_2\mathbf{x}(t-2h_x) \tag{18}$$

$$\mathbf{x}(t) = \phi(t), \quad -2h_x < t < 0 \tag{19}$$

where $\mathbf{A}_0(r)$, $\mathbf{A}_i$, $i = 1,2$ and $\mathbf{B}(r)$ are 44×44, 44×44 and 44×21 dimensional matrices. Parameter r is the coiler's radius, assumed to be slow varying. Table 1 shows the choice of 44 state and 21 control variables. The delay h_x denotes the transit time of the strip from the outlet of one stand to the inlet of the next one. The parameter r is the winding or pay-off reel radius. For this system where three stands are involved, two delay terms appear in the model [2.13,2.14]. For an n-stand cold rolling mill, there will be n-1 time delays. It is desired to find a singularly-perturbed model for this system.

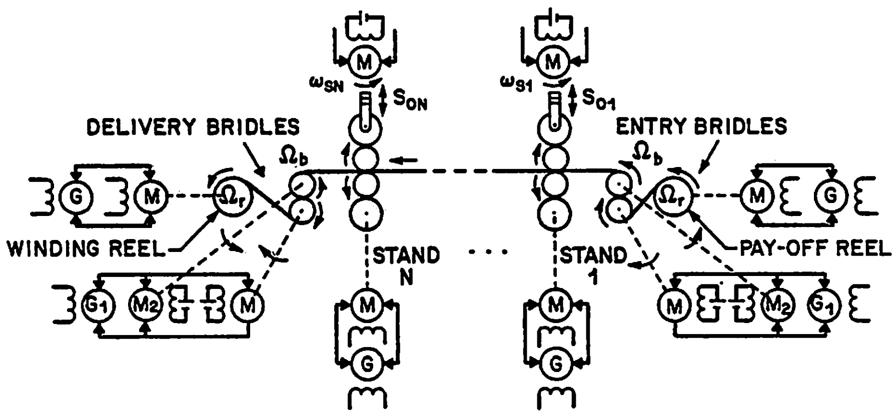

Figure 1. An N-Stand Cold Rolling Mill

Following the procedure discussed above, the delay h_x is divided into k sub-intervals and two set of vectors $\mathbf{Y}_i$, $\mathbf{W}_i$, $i = 1,2,...,k$ are defined to be $\mathbf{X}(t-ih_x)$ and $\mathbf{X}(t-2ih_x)$, respectively. The system (18)-(19) would be approximated by

$$\dot{\mathbf{X}} = \mathbf{A}_o(r)\mathbf{X} + \mathbf{B}(r)\mathbf{u} + \mathbf{A}_1\mathbf{Y}_k + \mathbf{A}_2\mathbf{W}_k \tag{20a}$$

$$\mu\dot{\mathbf{Y}} = \tilde{\mathbf{A}}_1\mathbf{X} + \tilde{C}_1\mathbf{Y} \tag{20b}$$

$$\mu\dot{\mathbf{W}} = \tilde{C}_{21}\mathbf{Y} + \tilde{C}_2\mathbf{W} \tag{20c}$$

$$\mathbf{X}(0) = \phi(0) \tag{21a}$$

$$\mathbf{Y}_i(0) = \phi(-i\mu) \tag{21b}$$

$$\mathbf{W}_i(0) = \phi(-2i\mu) \tag{21c}$$

where $\mathbf{Y} = (\mathbf{Y}_1'\mathbf{Y}_2' \cdots \mathbf{Y}_k')'$, $\mathbf{W} = (\mathbf{W}_1'\mathbf{W}_2' \cdots \mathbf{W}_k')'$ and $\mu = \Delta h_x = h_x/k$.

Table 1. A Choice of State and Control Vectors for Cold Rolling Mill

		Vector		
State	Physical	$\hat{i}_{f0j}\ \hat{i}_{f1j}\ \hat{i}_{fj}\ \Omega_{bj}\ \Omega_j\ \bar{i}_{f1}\ \hat{i}_{a1j}\ \hat{i}_{a2j}$	$\hat{i}_{aj}\ \hat{s}_j\ \hat{V}_{sj}\ \hat{i}_{f1j}\ \hat{i}_{f2j}\ \Omega_j\ \bar{i}_{0j}\ \hat{i}_{a}sj\ \hat{i}_{a1j}\ \hat{i}_{a2j}$	$n = 44$
	Mathematical	$x_1\ x_2\ x_3\ x_4\ x_5\ x_6\ x_7\ x_8$	$x_9\ x_{10}\ x_{11}\ x_{12}\ x_{13}\ x_{14}\ x_{15}\ x_{16}\ x_{17}\ x_{18}$	
Control	Physical	$\hat{e}_{oj}\ \ \hat{e}_{lj}\ \ \hat{e}_j$	$\hat{e}_{f1j}\ \ \hat{e}_{f2j}\ \ \hat{e}_{asj}\ \ \hat{e}_{a1j}\ \ \hat{e}_{a2j}$	$r = 21$
	Mathematical	$u_1\ \ u_2\ \ u_3$	$u_4\ \ u_5\ \ u_6\ \ u_7\ \ u_8$	

System (20) presents $\nu = (2k+1)n$ ordinary differential equations. For example, if $k = 10$ with $n = 44$, then $\nu = 924$. In an attempt to find an optimum control law for (20) a matrix Riccati equation of 924×924 dimension or equivalently $1/2\nu(\nu+1) = 427350$ scalar equations have to be solved which is impossible by the most efficient schemes [2.15] with the most powerful computers. A detail study of the three-stand cold rolling mill [2.13,2.14] reveals that among its 44 state variables, twelve of them associated with screw down mechanism and flux equation are so slow that they can be considered constant. Among the other 32 variables, another 15, associated with motors, armature circuits and transducers, are considered fast. The remaining 17 variables such as strip tensions, motors speeds, etc. are the most dominant states of the system. After considering the last observation the system (20) is rewritten as

$$\dot{\mathbf{x}} = \hat{\mathbf{A}}_o(r)\mathbf{x} + \mathbf{C}_o(r)\mathbf{z} + \mathbf{C}_3\mathbf{x} + \mathbf{B}_o(r)\mathbf{u} + \mathbf{d}_o \tag{22a}$$

$$\mu\dot{\mathbf{z}} = \hat{\mathbf{A}}_1\mathbf{x} + \mathbf{C}_4\mathbf{z} + \mathbf{B}_1\mathbf{u} + \mathbf{d}_1 \tag{22b}$$

$$\mu\dot{\mathbf{Z}} = \tilde{\mathbf{A}} + \tilde{\mathbf{C}}\mathbf{Z} \tag{22c}$$

Where $\mathbf{Z} = (\mathbf{Y}'\mathbf{W}')'$, $\mathbf{x}$ is 17th order dominant (slow) state vector, $\mathbf{z}$ is 15th order fast state vector and all vectors and matrices are of appropriate dimensions and are defined before as well as elsewhere [2.13]. Now if $\mu = 0$ eliminating $\mathbf{z}$ and $\mathbf{Z}$ then (22) reduces to

$$\dot{\mathbf{x}} = \mathbf{A}(r)\mathbf{x} + \mathbf{B}(r)\mathbf{u} \tag{23}$$

where $\mathbf{A}(r)$ and $\mathbf{B}(r)$ are 17×17 and 17×12 dimensional matrices are given by

$$\mathbf{A}(r) = \left[\begin{array}{cccccccc|ccccccccc}
-0.318 & & & & & & & & & & & & & & & & \\
 & -0.318 & & & & & & & & & & & & & & & \\
 & & -0.33 & & & & & & & & & & & & & & \\
 & & & -0.11 & & -0.532 & & & & & & 0 & & & & & \\
a_{51} & & a_{53} & & a_{55} & a_{56} & & & & & & & & & & & \\
 & & & 23.4 & a_{65} & -0.805 & -0.805 & & & & & & & & & & \\
 & & & & & -7.65 & -19.0 & & & & & & & & & & \\
 & & & & & 0.173 & & -0.1195 & 00.0224 & & & & & & & & \\
\hline
 & & & & & & & -3.95 & -19.0 & & & & & & & & \\
 & & & & & & & 0.041 & & -0.0111 & & & & & & & -0.0216 \\
 & & & & & & & & & -4.3 & -18.8 & & & & & & 6.53 \\
 & & & & & & & & & & & -0.318 & & & & & \\
 & & & & & & & & & & & & -0.318 & & & & \\
 & & 0 & & & & & & & & & & & -0.33 & & & \\
 & & & & & & & & & & & & 2.21 & & -0.11 & & -0.205 \\
 & & & & & & & & & & & a_{16,12} & & a_{16,14} & & a_{16,16} & a_{16,17} \\
 & & & & & & & & & & -0.973 & & & & -0.19 & a_{17,16} & -0.973
\end{array}\right]$$

where $a_{51}(r) = (N_1(r))/D_1(r)$, $a_{53}(r) = a_{55}(r) = 0.1a_{51}(r) = -0.7a_{51}(r)$, $a_{56}(r) = (N_2(r))/D_1(r)$, $a_{65}(r) = 6.6r$, $a_{16,12}(r) = (N_1(r))/D_2(r)$, $a_{16,14}(r) = 0.1a_{16,12}(r)$, $a_{16,17}(r) = (N_2(r))/D_2(r)$, and $a_{17,16}(r) = 7.96r$ with $N_1(r) = 1.98(r+0.152r^3)$, $D_1(r) = 1 + 0.0185r^4$, $N_2(r) = -8.52r$, $D_2(r) = 1 + 0.0483r^4$; and

$$
B(r) = \left[\begin{array}{c|c} B_1(r) & 0 \\ \hline 0 & B_2(r) \end{array}\right] = \left[\begin{array}{cccc|ccccccc}
-0.318 & & & & & & & & & & \\
 & b_{33}(r) & & & & & & & & 0 & \\
 & & -1.22 & -1.22 & & & & & & & \\
 & & & & -0.826 & -0.826 & & & & & \\
\hline
 & & & & & & -0.424 & -0.424 & & & \\
 & & & & & & & & 0.318 & & \\
 & & & & & & & & & 0.318 & \\
 & & & & & & & & & & b_{14,22}(r) \\
 & 0 & & & & & & & 0 & & 0 \\
 & & & & & & & & & & 0 \\
 & & & & & & & & & & 0
\end{array}\right]
$$

where $b_{33}(r)=0.105/(r(1-5.25r^2)), b_{14,22}(r)=0.105/(r(1-8.78r^2))$. This is the desired reduced-order model.

2.6.3 Serial Time-Delay Systems

In this section the linear large-scale system with a serial structure and time-delay linear interaction shown in Figure 2 is discussed. The state and output equations of this system are represented by,

$$\dot{\mathbf{x}}_1(t) = \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1\mathbf{u}_1(t) + \mathbf{v}_1(t) \tag{24a}$$

$$\mathbf{y}_1(t) = \mathbf{C}_1\mathbf{x}_1(t) \tag{24b}$$

$$\dot{\mathbf{x}}_2(t) = \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{B}_2\mathbf{u}_2(t) + D_2\mathbf{x}_1(t-\tau_1) + \mathbf{v}_2(t) \tag{24c}$$

$$\mathbf{y}_2(t) = \mathbf{C}_2\mathbf{x}_2(t) \tag{24d}$$

$$\vdots$$

$$\dot{\mathbf{x}}_N(t) = \mathbf{A}_N\mathbf{x}_N(t) + \mathbf{B}_N\mathbf{u}_N(t) + D_N\mathbf{x}_{N-1}(t-\tau_{N-1}) + \mathbf{v}_N(t) \tag{24e}$$

$$\mathbf{y}_N(t) = \mathbf{C}_N\mathbf{x}_N(t) \tag{24f}$$

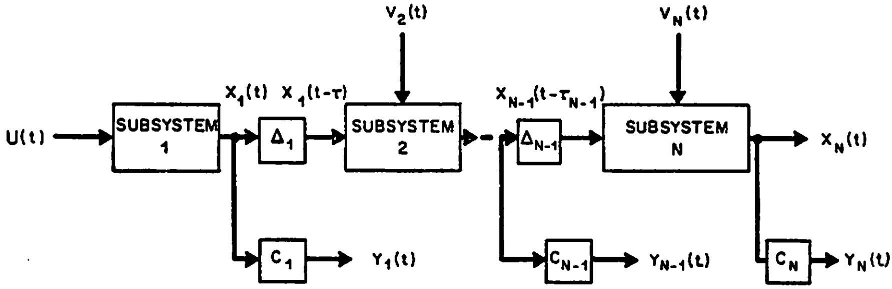

Figure 2. A Serial TD System

where $\mathbf{x}_i(t)$, $\mathbf{y}_i(t)$ and $\mathbf{u}_i(t)$ are n_i-, r_i- and m_i-*dimensional* state, output and control vectors, respectively, $v_i(t)$ is the ith subsystem external input and τ_i, $i = 1,2,\ldots,N-1$ are the delays between the subsystems. A hierarchical control of this system is given in Section 8.3.

2.6.4 Hierarchical Formulation of a Large-Scale Time-Delay System

In this section it is demonstrated how a given TD system is decomposed into N subsystems suitable for optimal hierarchical control as will be discussed in Chapter 8. The discussions of this section is applicable to both continuous-time and discrete-time linear or nonlinear TD systems. For the sake of discussion, consider the following linear large-scale discrete-time TD system:

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{C}(k)\mathbf{x}(k-h_x) + \mathbf{D}(k)\mathbf{u}(k-h_u) \qquad (25)$$

$$\mathbf{x}(k) = \phi(k), -h_x < k < 0 \qquad (26a)$$

$$\mathbf{u}(k) = \eta(k), -h_u < k < 0 \qquad (26b)$$

where $\mathbf{x}(k)$ and $\mathbf{u}(k)$ are $n\times 1$ and $r\times 1$ state and control vectors, respectively and

$\mathbf{A}(\cdot)$, $\mathbf{B}(\cdot)$, $\mathbf{C}(\cdot)$ and $\mathbf{D}(\cdot)$ are $n \times n$, $n \times r$, $n \times n$ and $n \times r$ matrices. Let us assume that the state and control can be decomposed as

$$\mathbf{x} = [\mathbf{x}_1'\mathbf{x}_2' \cdots \mathbf{x}_N']' \tag{27a}$$

$$\mathbf{u} = [\mathbf{u}_1'\mathbf{u}_2' \cdots \mathbf{u}_N']' \tag{27b}$$

where $\mathbf{x}_i(k) \epsilon R^{n_i}$, $\mathbf{u}_i(k) \epsilon R^{r_i}$ for $i = 1,2, \ldots, N$ and

$$\sum_{i=1}^{N} n_i = n \ , \ \sum_{i=1}^{N} r_i = r \tag{28}$$

The initial functions $\phi(t)$ and $\eta(t)$ can be similarly partitioned. The matrices **A**, **B**, **C** and **D** are partitioned by denoting the ith block row of the above matrices by $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{C}_i$ and $\mathbf{D}_i$ and the ij block of these matrices by $\mathbf{A}_{ij}$, $\mathbf{B}_{ij}$, $\mathbf{C}_{ij}$ and $\mathbf{D}_{ij}$, $i, j = 1, 2, \ldots, N$. Let $\hat{\mathbf{A}}_i$ be $\mathbf{A}_i$ with $\mathbf{A}_{ii}$ set to zero, *i.e.*

$$\hat{\mathbf{A}}_i = [\mathbf{A}_{i1}, \mathbf{A}_{2}, \cdots, \mathbf{A}_{i,i-1}, 0, \mathbf{A}_{i,i+1}, \ldots, \mathbf{A}_N] \,. \tag{29}$$

Similar notation applies to matrices $\mathbf{B}_i$, $\mathbf{C}_i$ and $\mathbf{D}_i$. With the above notation, the *ith* subsystem, $i = 1,2, \ldots, N$, becomes

$$\mathbf{x}_i(k+1) = \mathbf{A}_i(k)\mathbf{x}_i(k) + \mathbf{B}_i(k)\mathbf{u}_i(k) + \mathbf{z}_i(k) \tag{30}$$

where

$$\begin{aligned} \mathbf{z}_i(k) = {} & \hat{\mathbf{A}}_i(k)\mathbf{x}(k) + \hat{\mathbf{B}}_i(k)\mathbf{u}(k) + \mathbf{C}_i(k)\mathbf{x}(k-h_x) \\ & + \mathbf{D}_i(k)u(k-h_u) \end{aligned} \tag{31}$$

The formulation of (30)-(31) is now appropriate for a hierarchical control of large-scale systems with time delay which can be solved by many multilevel techniques. In Chapter 8 various large-scale hierarchical control schemes such as goal coordination, interaction prediction and costate prediction will be extended to the TD case. These schemes are considered also in References [2.16-2.19].

PROBLEMS

2.1 Consider a single-reservoir water resources system shown in Figure P2.1. where the stored water is assumed to be the state variable, $u_i(t)$, $i = 1,2,3$ are the control variables while $y_1(t)$ is a known river input flow, and $y_2(t)$ is the output flow. Develop a state-space model for this system. Find an output equation in terms of $y_2(t)$ to accompany the state equation.

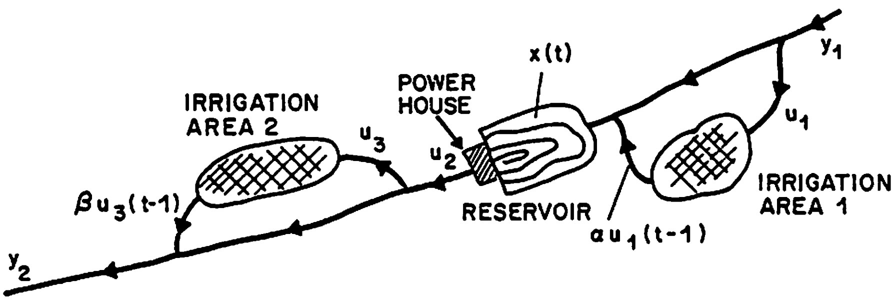

Figure P2.1

2.2 Determine the transfer function of the MIMO continuous - time TD system whose state and output equations are given by (2.3.5) and (2.3.6).

2.3 Determine the transfer function of the MIMO discrete-time TD system whose state and output equations are given by (2.3.9) and (2.3.10).

2.4 Find the transfer function for a sample-and-hold unit shown in Figure P2.4. Under which conditions is $H(s) = \frac{1}{sT}(1-e^{-sT})$ valid?

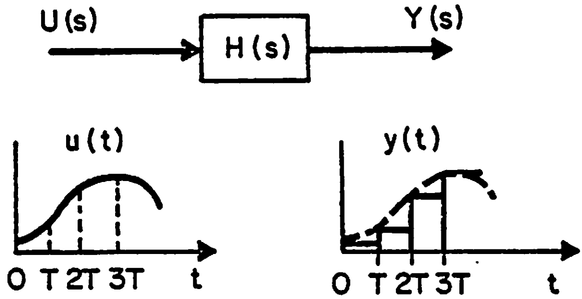

Figure P2.4

2.5 Consider the SISO discrete-time TD system shown in Figure P2.5 where K is a positive constant and ℓ is a positive integer.

i. Write a set of state and output equations for this system.

ii. Determine the closed-loop transfer function of the system.

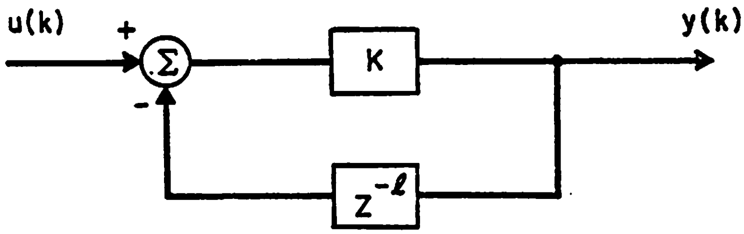

Figure P2.5

2.6 Consider the continuous-time TD system shown in Figure P2.6 where K is a positive constant.

i. Write a set of state and output equations for this system.

ii. Determine the closed-loop transfer function and calculate the poles.

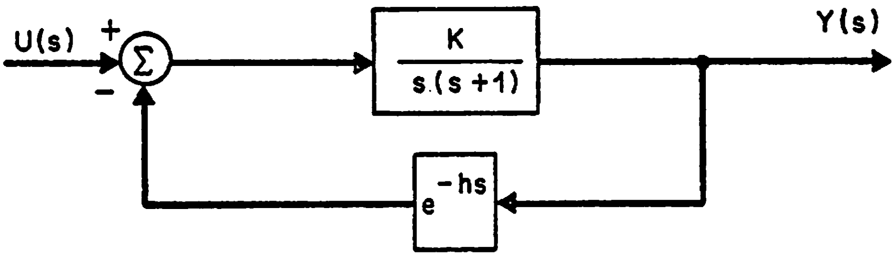

Figure P2.6

2.7 If $g(t) = L^{-1}[G(s)]$, show that $G^*(s) = \sum_{k=0}^{\infty} g(kt)e^{-kTs}$ and verify (2.4.24).

2.8 Consider the system shown in Figure 2.4.5. Show that

$$\sum_{k=0}^{\infty} y(kt-\Delta T)z^{-k} = \sum_{k=0}^{\infty} g(kt-\Delta T)z^{-k} \sum_{k=0}^{\infty} u(kT)z^{-k}$$

2.9 Verify (2.4.29) and (2.4.31).

2.10 Derive (2.4.34) from (2.4.33).

2.11 Find the z-transfer function of the sampled-data TD system shown in Figure P2.11 where $T = 1$ and

$$G_1(s) = \frac{e^{-s}}{s+1}$$

$$G_2(s) = \frac{e^{-1.2s}}{s+2}$$

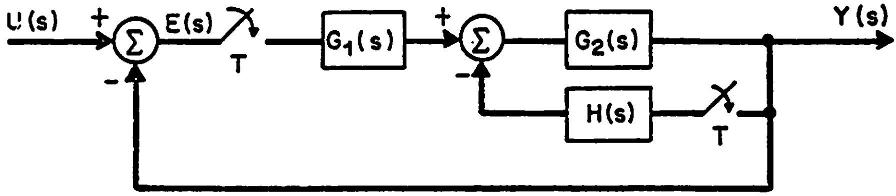

Figure P2.11

2.12 Find the step response of the sampled-data TD system shown in Figure P2.12 where $T = 1$ and

$$G_1(s) = \frac{e^{-1.4s}}{s+1}$$

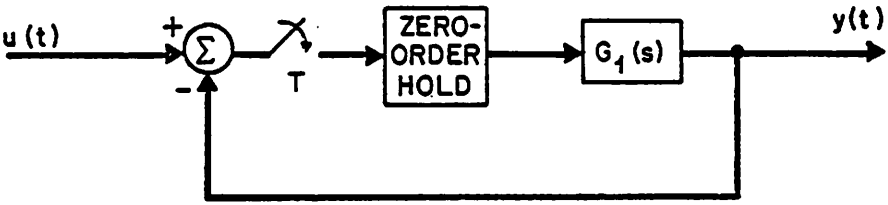

Figure P2.12

2.13 Find a linearized TD model for the following system

$$\dot{\mathbf{x}} = \begin{bmatrix} -2x_1^2(t-0.1) - x_2(t)u_1(t) - 5x_3^3(t) \\ -4x_1^2(t-0.1) + x_1(t-0.1)x_3(t) \\ -x_3(t)u_2(t) - x_2(t)u_2(t) \end{bmatrix}$$

about the *nominal state* $\mathbf{x}_n(t) = [0,1,0]'$ and *nominal control* $\mathbf{u}_n(t) = [1,1]'$. Assume all initial functions are zero.

2.14 Consider the following TD system

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\mathbf{u}(t) + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\mathbf{x}(t-0.1)$$

with $\mathbf{x}(t) = [1,-1]$, $-0.1 \leqslant t \leqslant 0$. Use the development of Section 2.6.2 to obtain an equivalent singular perturbation model. Let $k = 5$ for this problem.

REFERENCES

[2.1] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, 1977

[2.2] M. Malek-Zavarei and M. Jamshidi, "Sensitivity of Linear Time-Delay Systems to Parameter Variations," *IFAC J. Automatica*, Vol. 11, pp. 315-319, 1975

[2.3] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach*, Pergamon Press, 1986

[2.4] J. E. Marshall, *Control of Time-Delay Systems*, Peter Peregrinus, Ltd. (London), 1979.

[2.5] E. I. Jury, *Sampled-Data Control Systems*, Wiley, 1958

[2.6] B. J. Kuo, *Analysis and Synthesis of Sample-Data Control Systems*, Prentice-Hall, Inc., 1963

[2.7] E. I. Jury, "Additions to the Modified z-transform Method," *IRE Wescom Convention Record*, Pt. IV, pp. 136-151 August 21, 1977

[2.8] M. Jamshidi, "A Three-Stage Design of Non-Linear Control Systems with Time-Delay," *Int. J. Contr.*, vol. 21, pp. 753-762, 1975

[2.9] A. E. Bryson, Jr. and Y. C. Ho, *Applied Optimal Control*, Blaisdel Publ. Co. (Waltham, MA), 1969

[2.10] M. Jamshidi, "Sub-optimal Control of Coupled Time-Delay Systems," *Int. J Control,* Vol. 17, pp. 995-1008, 1973

[2.11] P. Sannuti, "Near-Optimum Design of Time-Lag Systems by Singular Perturbation Method," *Proc JACC,* paper No. 20-A, Atlanta, GA, June, 1970

[2.12] K. Inoue, H. Akashi, K. Ogino, and Y. Sawaragi, "Sensitivity Approaches to Optimization of Linear Systems with Time Delay," *IFAC J. Automatica,* Vol. 7, pp. 671-679, 1971

[2.13] M. Jamshidi, "A Near-Optimum Controller for Cold Rolling Mills," *Int. J. Control,* Vol. 15, pp. 1137-1154, 1972

[2.14] M. Jamshidi, "An Integrated Near-Optimum Design of Cold Rolling Mills," CSL Report R-499, University of Illinois, Urbana, IL, January, 1971

[2.15] M. Jamshidi, "An Overview on the Algebraic Matrix Riccati Equation and Related Problems," *J. Large-Scale Systems,* Vol. 1, pp. 167-192, 1980

[2.16] M. Jamshidi and M. Malek-Zavarei, "A Hierarchical Control of Large-Scale Time-Delay Systems," *J. Large-Scale Systems,* Vol. 4, pp. 149-163, 1983. Also, *Proc. IEEE Conference on Decision and Control,* San Diego, CA, December 16-18, 1981, pp. 1303-1304

[2.17] M. Jamshidi and J. E. Merryman, "On the Hierarchical Optimization of Retarded Systems via Costate Prediction," *Proc. Amer. Contr. Conference,* Arlington, VA. June 14-16, 1982

[2.18] M. Jamshidi and C. M. Wang, "Optimal Hierarchical Control of Non-Linear Large-Scale Time-Delay Systems," *Proc. IEEE Intr. Symp. Large-Scale Systems,* Virginia Beach, VA, October 11-13, 1982, pp. 258-262

[2.19] M. Jamshidi, *Large-Scale Systems: Modeling and Control,* Elsevier North-Holland, New York, 1983 (Chapter 6)

Part II ANALYSIS

CHAPTER 3

ANALYSIS OF TIME-DELAY SYSTEMS

3.1 INTRODUCTION

In Chapter 2 we showed that the behavior of any linear continuous-time TD system can be described by the state and output equations (2.3.5) and (2.3.6). Further, we showed that for linear discrete-time TD systems in which the sampling period is fixed and delays are integral multiples of the sampling period the state and output equations will be pure difference equations given in (2.3.9) and (2.3.10). We will not pursue such systems any more. They are special cases of mixed or hybrid systems whose state and output equations are also characterized by (2.3.5) and (2.3.6).

The main problem in the analysis of linear TD systems is solving the state equation (2.3.5). We will denote the solution to (2.3.5) for $t \geqslant t_o$, where t_o is the initial time, and for initial state ϕ and input $\mathbf{u}$ as $\mathbf{x}(t, t_o, \phi, \mathbf{u})$ or sometimes simply as $\mathbf{x}(t)$. The output $\mathbf{y}$ can be determined by the substitution of the state and input in (2.3.6). We will take up the solution of the state equation in this chapter. We will first deal with the case where the input is identically zero and will solve the homogeneous state equation. Then we will provide the complete solution. Adjoint systems will also be introduced in this chapter and their application in the solution of the state equation will be discussed.

3.2 HOMOGENEOUS STATE EQUATION - FUNDAMENTAL MATRIX

When the input is identically zero, the state equation (2.3.5) becomes

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{xi}), \quad t \geqslant t_o \tag{1}$$

Equation (1) is referred to as the *homogeneous* or the *unforced* state equation. The initial state or the initial function for (1) is

$$\mathbf{x}(t) = \phi(t) \quad , \quad t\epsilon[t_o - \Delta_x, t_o] \tag{2}$$

where Δ_x indicates the largest delay. Using the notation introduced in Section 3.1, we indicate the solution to (1) with initial function (2) as $\mathbf{x}(t, t_o, \phi, \mathbf{0})$ where the argument $\mathbf{0}$ indicates zero input. The first question that arises is that of existence and uniqueness of solutions. It has been shown [3.1-3.3] that if $\mathbf{A}(.)$ and $\mathbf{A}_i(.)$, $i=1,2,\ldots,N$, are real continuous matrices and h_{xi} are positive constants, a solution $\mathbf{x}(t, t_o, \phi, \mathbf{0})$ of (1) exists and uniquely depends on the initial state $\phi(.)\epsilon B$ given in (2) where B is the Banach space of real continuous functions defined on the interval $[t_o - \Delta_x, t_o]$. The reader is referred to the above references for the proof of existence and uniqueness of solutions. We will assume in the following discussion that the above conditions are satisfied so that a unique solution always exists.

First we note that the solution of (1) is linear *w.r.t.* the initial function. More precisely, we have the following Theorem.

1. Theorem. Let matrices $\mathbf{A}(.)$ and $\mathbf{A}_i(.)$, $i=1,2,\ldots,N$ in (1) be continuous for $t \geqslant t_o - \Delta_x$ where Δ_x is the largest delay. Then the solution to (1) is linear *w.r.t.* the initial function $\phi(t)$ in (2).

Proof. Let $\mathbf{x}(t, t_o, \phi_1, \mathbf{0})$ and $\mathbf{x}(t, t_o, \phi_2, \mathbf{0})$ be the solutions of (1) corresponding to initial functions ϕ_1 and ϕ_2, respectively. Consider the function

$$\mathbf{x}(t) = c_1\mathbf{x}(t, t_o, \phi_1, \mathbf{0}) + c_2\mathbf{x}(t, t_o, \phi_2, \mathbf{0}) \tag{3}$$

where c_1 and c_2 are two arbitrary scalar constants. It is easily verified that the function $\mathbf{x}(t)$ in (3) satisfies (1). Further, $\mathbf{x}(t)$ in (3) satisfies the initial function

$$\mathbf{x}(t) = c_1\phi_1(t) + c_2\phi_2(t), \quad t\epsilon[t_o - \Delta_x, t_o] \tag{4}$$

Therefore, we have

$$\mathbf{x}(t) \triangleq \mathbf{x}(t, t_o, c_1\phi_1 + c_2\phi_2, \mathbf{0}) = c_1\mathbf{x}(t, t_o, \phi_1, \mathbf{0}) + c_2\mathbf{x}(t, t_o, \phi_2, \mathbf{0}) \tag{5}$$

which established the linearity of the solution *w.r.t.* the initial function. Δ

For the sake of discussion, let us first consider a single delay, that is

$$h_{x1} = h \quad , \quad A_i(t) = 0 \quad , \quad i=2,3,\ldots,N \tag{6}$$

Then the state equation (1) becomes

$$\dot{x}(t) = A(t)x(t) + A_1(t)x(t-h), \quad t \geqslant t_o \tag{7a}$$

Assume that A(.) and A_1(.) are $n \times n$ matrices and that the initial state of (7a) is

$$x(t) = \phi(t) \quad , \quad t_o - h \leqslant t \leqslant t_o \tag{7b}$$

2. Definition. An $n \times n$ matrix $\Phi(t,\tau)$ is called a *kernel matrix* or a *fundamental matrix* of (7) if it satisfies the following conditions:

a) $\dot{\Phi}(t,\tau) = A(t)\Phi(t,\tau) + A_1(t)\Phi(t-h,\tau), \quad t \geqslant t_o$ (8a)

b) $\Phi(t,\tau) = I\delta(t-\tau)$ for $t,\tau \epsilon [t_o-h,t_o]$ (8b)

where I is the identity matrix and δ(.) is the Dirac delta function. Δ

The solution of (7) can be expressed in terms of fundamental matrix $\Phi(t,\tau)$ as given in the following theorem.

3. Theorem. The solution of (7a) with initial function (7b) is

$$x(t) \triangleq x(t,t_o,\phi,0) = \int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau, \quad t \geqslant t_o \tag{9}$$

where the matrix $\Phi(t,\tau)$ satisfies conditions (8a) and (8b).

Proof. We have, from (9) and (8a),

$$\begin{aligned}
\dot{x}(t) &= \int_{t_o-h}^{t_o} \dot{\Phi}(t,\tau)\phi(\tau)d\tau \\
&= \int_{t_o-h}^{t_o} [A(t)\Phi(t,\tau) + A_1(t)\Phi(t-h,\tau)]\phi(\tau)d\tau \\
&= A(t)[\int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau] + A_1(t)[\int_{t_o-h}^{t_o} \Phi(t-h,\ \tau)\phi(\tau)d\tau] \\
&= A(t)x(t) + A_1(t)x(t-h)
\end{aligned} \tag{10}$$

Thus, (9) satisfies (7a). Also, for any $t \epsilon [t_o-h,\ t_o]$, by (8b), (9) becomes

$$\mathbf{x}(t) = \int_{t_o-h}^{t_o} \mathbf{I}\ \delta(t-\tau)\phi(\tau)d\tau = \phi(t) \tag{11}$$

where the sifting property of delta function [3.4] has been used. Thus, $\mathbf{x}(t)$ in (9) also satisfies the initial function (7b) and it is indeed the solution to (7). Δ

Therefore, to solve (7) we need to calculate fundamental matrix $\Phi(t,\tau)$. Calculation of $\Phi(t,\tau)$ from (8) is extremely difficult. An iterative method for the construction of $\Phi(t,\tau)$ is given below.

4. Iterative Method to Construct $\Phi(t,\tau)$. Let us use subscript i to indicate ith iteration. Initially, let

$$\Phi_1(t,\tau) = \begin{cases} \mathbf{I}\ \delta(t-\tau) & \text{for } t,\tau\epsilon[t_o-h,t_o] \\ 0 & \text{for } t > t_o \end{cases} \tag{12}$$

Then let

$$\Phi_2(t,\tau) = \begin{cases} \mathbf{I}\delta(t-\tau) & \text{for } t,\tau\epsilon[t_o-h,\ t_o] \\ \int_{t_o}^{t} [\mathbf{A}(s)\Phi_1(s,\tau) + \mathbf{A}_1(s)\Phi_1(s-h,\ \tau)]ds & \text{for } t > t_o \end{cases} \tag{13}$$

and, in general, for iteration i let

$$\Phi_i(t,\tau) = \begin{cases} \mathbf{I}(\delta(t-\tau) & \text{for } t,\tau\epsilon[t_o-h,\ t_o] \\ \int_{t_o}^{t} [\mathbf{A}(s)\Phi_{i-1}(s,\tau) + \mathbf{A}_1\Phi_{i-1}(s-h,\ \tau)]ds & \text{for } t > t_o \end{cases} \tag{14}$$

It can be shown [3.3] that the sequence $\{\Phi_i(t,\tau)\}$ converges uniformly to fundamental matrix $\Phi(t,\tau)$ in (8). Δ

An alternative characterization of the solution of (7) is provided by the following theorem.

5. Theorem. The solution of (7a) with initial function (7b) is

$$\mathbf{x}(t) \triangleq \mathbf{x}(t,t_o,\phi,0) = \Psi(t,t_o)\phi(t_o) + \int_{t_o-h}^{t_o} \Psi(t,\tau+h)\mathbf{A}_1(\tau+h)\phi(\tau)d\tau \tag{15}$$

where the matrix $\Psi(t,\tau)$ satisfies the following conditions:

a) $\dfrac{\partial}{\partial\tau}\Psi(t,\tau) = -\Psi(t,\tau)A(\tau) - \Psi(t,\tau+h)A_1(\tau+h),\ t_o \leqslant \tau < t-h$ (16a)

b) $\dfrac{\partial}{\partial\tau}\Psi(t,\tau) = -\Psi(t,\tau)A(\tau)\ ,\ t-h \leqslant \tau \leqslant t$ (16b)

c) $\Psi(t,t) = I$ (16c)

d) $\Psi(t,\tau) = 0$ for $\tau > t$ (16d)

The proof of this theorem will be deferred until the presentation of the adjoint state equation in Section 3.4. It is noted that matrix $\Psi(t,\tau)$ in (16) is also called a *kernel matrix* or a *fundamental matrix* of (7).

In the case of multiple delay state equation, *i.e.*, system (1) with initial function (2), Theorem 3 can be extended as follows.

6. Theorem. The solution of (1) with initial function (2) is

$$x(t) \triangleq x(t,t_o,\phi,0) = \int_{t_o-\Delta_x}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau,\ t \geqslant t_o \tag{17}$$

where the fundamental matrix $\Phi(t,\tau)$ is the solution of

$$\dot{\Phi}(t,\tau) = A(t)\Phi(t,\tau) + \sum_{i=1}^{N} A_i(t)\Phi(t-h_{xi},\tau)\ ,\ t \geqslant t_o \tag{18a}$$

with the initial condition

$$\Phi(t,\tau) = I\delta(t-\tau)\ \text{ for }\ t,\tau\epsilon[t_o-\Delta_x,t_o] \tag{18b}$$

where I is the identity matrix and $\delta(.)$ is the Dirac delta function. Δ

The proof of the above theorem is a straightforward extension of the proof of Theorem 3 and will not be presented. Also, $\Phi(t,\tau)$ in (18) can be constructed approximately by an iterative method similar to the single-delay case. (See Problem 3.1.)

An alternate characterization of the solution of (1) can be found by an extension of Theorem 5 as given below.

7. Theorem. The solution of the linear homogeneous multiple-delay state equation (1) where $0 < h_{x1}, < h_{x2} < \cdots < h_{xN} = \Delta_x$ with initial function (2) is

$$\mathbf{x}(t) \triangleq \mathbf{x}(t,t_o,\phi,0) =$$

$$\Psi(t,t_o)\phi(t_o) + \sum_{i=1}^{N} \int_{t_o-h_{xi}}^{t_o} \Psi(t,\tau+h_{xi})\ \mathbf{A}_i(\tau+h_{xi})\phi(\tau)d\tau,\ t \geqslant t_o \tag{19}$$

where fundamental matrix $\Psi(t,\tau)$ satisfies the following conditions:

a) $$\frac{\partial}{\partial\tau}\Psi(t,\tau) = -\Psi(t,\tau)\mathbf{A}(\tau),\ t-h_{x1} \leqslant \tau < t$$

$$= -\Psi(t,\tau)\mathbf{A}(\tau) - \Psi(t,\tau+h_{x1})\mathbf{A}_1(\tau+h_{x1}),\ t-h_{x2} \leqslant \tau < t-h_{x1}$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$= -\Psi(t,\tau)\mathbf{A}(\tau) - \sum_{i=1}^{N} \Psi(t,\tau+h_{xi})\mathbf{A}_i(\tau+h_{xi}),\ \tau < t-\Delta_x \tag{20a}$$

b) $$\Psi(t,t) = \mathbf{I} \tag{20b}$$

c) $$\Psi(t,\tau) = 0 \ \text{ for } \ \tau > t \tag{20c}$$

The proof of this theorem will also be deferred until Section 3.4.

3.3 FORCED STATE EQUATION - COMPLETE SOLUTION

In this section we will solve the linear *TD* state equation with input, *i.e.*, equations (2.3.5) which is repeated here for convenience:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{xi}) + \mathbf{B}(t)\mathbf{u}(t) + \sum_{i=1}^{R} \mathbf{B}_i(t)\mathbf{u}(t-h_{ui}),\ t \geqslant t_o \tag{1a}$$

with initial functions

$$\mathbf{x}(t) = \phi(t),\ t\epsilon[t_o-\Delta_x,t_o] \tag{1b}$$

$$\mathbf{u}(t) = \eta(t),\ t\epsilon[t_o-\Delta_u,t_o] \tag{1c}$$

where Δ_x and Δ_u are, respectively, the largest delays in state and control. We denote the solution $\mathbf{x}(t)$ to (1) more precisely as $\mathbf{x}(t,t_o,\phi,\mathbf{u})$ which indicates the state trajectory for $t \geqslant t_o$, where t_o is the initial time, subject to initial state $\phi(.)$ and control $\mathbf{u}(.)$.

An important property of linear systems is the decomposition property. That is, in a linear system the complete response is the sum of its zero-state response and its zero-input response. Using the above terminology, the decomposition property implies that

$$\mathbf{x}(t,t_o,\phi,\mathbf{u}) = \mathbf{x}(t,t_o,\phi,0) + x(t,t_o,0,\mathbf{u}) \tag{2}$$

where $\mathbf{x}(t,t_o,\phi,0)$ indicates the zero-input (or homogeneous) solution of (1) and $\mathbf{x}(t,t_o,0,\mathbf{u})$ indicates the solution to (1) where the initial state (1b) is identically zero. We will not give a rigorous proof of this fact. It can be found, for example, in reference [3.3]. But in solving (1) we will see that (2) will hold true. Furthermore, both the zero-input response and the zero-state response are linear functions. The linearity of the zero-input response *w.r.t.* initial state was shown in Theorem 3.2.1. For the linearity of the zero-state we have the following theorem.

1. Theorem. The zero-state solution $\mathbf{x}(t,t_o,0,\mathbf{u})$ of (1) is a linear function of the input. More precisely, consider the state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{l=1}^{N} \mathbf{A}_l(t)\mathbf{x}(t-h_{xl}) + \mathbf{f}(t), \ t \geqslant t_o \tag{3a}$$

with initial state

$$\mathbf{x}(t) = 0, \ t \epsilon [t_o-\Delta_x,t_o] \tag{3b}$$

where Δ_x is the largest delay. Let

$$\mathbf{f}(t) = \sum_{p=1}^{k} c_p \mathbf{f}_p(t) \quad , \ t \geqslant t_o \tag{4}$$

where c_p are constant scalars and $\mathbf{f}_p(.)$ are given vector-valued functions. Now if

$$\dot{\mathbf{x}}_p(t) = \mathbf{A}(t)\mathbf{x}_p(t) + \sum_{l=1}^{N} \mathbf{A}_l(t)\mathbf{x}_p(t-h_{xl}) + \mathbf{f}_p(t) \quad , \ t \geqslant t_o \tag{5a}$$

with initial state

$$\mathbf{x}_p(t) = 0, \ t \epsilon [t_o-\Delta_x,t_o] \tag{5b}$$

has a solution $\mathbf{x}_p(t)$, $t \geqslant t_o$ $(p=1,2,\ldots,k)$, then

$$\mathbf{x}(t) = \sum_{p=1}^{k} c_p \mathbf{x}_p(t), \quad t \geq t_o \tag{6}$$

is a solution of (3).

Proof. From (6), (5a) and (4) we have, for $t \geq t_o$,

$$\begin{aligned}
\dot{\mathbf{x}}(t) &= \sum_{p=1}^{k} c_p \dot{\mathbf{x}}_p(t) \\
&= \sum_{p=1}^{k} c_p [\mathbf{A}(t)\mathbf{x}_p(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}_p(t-h_{xi}) + \mathbf{f}_p(t)] \\
&= \mathbf{A}(t)\sum_{p=1}^{k} c_p \mathbf{x}_p(t) + \sum_{i=1}^{N} \mathbf{A}_i(t) \sum_{p=1}^{k} c_p \mathbf{x}_p(t-h_{xi}) + \sum_{p=1}^{k} c_p \mathbf{f}_p(t) \\
&= \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{xi}) + \mathbf{f}(t)
\end{aligned} \tag{7}$$

Thus, $\mathbf{x}(t)$ in (6) satisfies (3a). Also, using (6) and (5b) it is easy to show that initial condition (3b) is satisfied. This proves the theorem. Δ

The above property of linear systems is referred to as the *superposition property.*

The complete solution to (1) can be found from the zero-input solution (3.2.17) or (3.2.19) by using the variation of parameters technique [3.6]. This is detailed in Theorems 2 and 4 below.

2. Theorem. The solution of (1) can be written as

$$\begin{aligned}
\mathbf{x}(t) \triangleq \mathbf{x}(t,t_o,\phi,\mathbf{u}) = \int_{t_o-\Delta_x}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau + \int_{t_o}^{t} \Phi_u(t,\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) \\
+ \sum_{i=1}^{R} \mathbf{B}_i(\tau)\mathbf{u}(\tau-h_{ui})]d\tau, \quad t \geq t_o
\end{aligned} \tag{8}$$

where fundamental matrix $\Phi(t,\tau)$ satisfies (3.2.18) and matrix $\Phi_u(t,\tau)$ satisfies the following conditions:

a) $\dot{\Phi}_u(t,\tau) = \mathbf{A}(t)\Phi_u(t,\tau) + \sum_{i=1}^{N} \mathbf{A}_i(t)\Phi_u(t-h_{xi},\tau), \quad t \geq t_o$ (9a)

b) $\Phi_u(t,t) = \mathbf{I}$ (9b)

c) $\Phi_u(t,\tau) = 0$ for $\tau > t$ (9c)

Proof. Using the decomposition property we can write the solution of (1) as

$$\mathbf{x}(t) = \mathbf{x}(t,t_o,\phi,0) + \mathbf{x}(t,t_o,0,\mathbf{u}) \tag{10}$$

We have already shown that the zero-input solution $\mathbf{x}(t,t_o,\phi,0)$ is given by the first part of (8). (See Theorem 3.2.6.) Let $\mathbf{x}(t,t_o,0,\mathbf{u})$ indicate the second part of (8) by an abuse of notation, and show that it indeed represents the zero-state solution of (1) by showing that it satisfies (1a). We have, using respectively (9b), (9a) and (9c):

$$\begin{aligned}
&\dot{\mathbf{x}}(t,t_o,0,\mathbf{u})\\
&= \int_{t_o}^{t} \dot{\Phi}_u(t,\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \sum_{i=1}^{R} \mathbf{B}_i(\tau)\mathbf{u}(\tau-h_{u_i})]d\tau + \mathbf{B}(t)\mathbf{u}(t) + \sum_{i=1}^{R} \mathbf{B}_i(t)\mathbf{u}(t-h_{u_i})\\
&= \int_{t_o}^{t} \mathbf{A}(t)\Phi_u(t,\tau)[\mathbf{B}(\tau)\mathbf{u}(\tau) + \sum_{i=1}^{R} \mathbf{B}_i(\tau)\mathbf{u}(\tau-h_{u_i})]d\tau +\\
&\int_{t_o}^{t-h_{x_i}} \sum_{i=1}^{N} \mathbf{A}_i(t)\Phi_u(t-h_{x_i},\tau))\ [\mathbf{B}(\tau)\mathbf{u}(\tau) + \sum_{i=1}^{R} \mathbf{B}_i(t)\mathbf{u}(\tau-h_{ui})]d\tau +\\
&\mathbf{B}(t)\mathbf{u}(t) + \sum_{i=1}^{R} \mathbf{B}_i(t)\mathbf{u}(t-h_{u_i})\\
&= \mathbf{A}(t)\mathbf{x}(t,t_o,0,\mathbf{u}) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{x_i}) + \mathbf{B}(t)\mathbf{u}(t) + \sum_{i=1}^{R} \mathbf{B}_i(t)\mathbf{u}(t-h_{u_i})
\end{aligned} \tag{11}$$

Note that by (9c) $\Phi(t-h_{x_i},\tau) = 0$ for $\tau > t - h_{x_i}$. *q.e.d.* Δ

The complete solution of (8) requires the knowledge of $\Phi_u(t,\tau)$. An iterative method similar to that described in Section 3.3.4 for the construction of $\Phi(t,\tau)$ can be used to construct $\Phi_u(t,\tau)$ approximately, as given below.

3. Iterative Method to Construct $\Phi_u(t,\tau)$. To describe the procedure, consider only a single state delay h. The procedure can be easily extended to the case of multiple delays. (See Problem 3.5.) Let subscript i indicate ith iteration. Initially let, for $\tau \geq t_o$,

$$\Phi_{u_1}(t,\tau) = \begin{cases} \mathbf{0} & \text{for} \quad t < \tau \\ \mathbf{I} & \text{for} \quad t \geq \tau \end{cases} \tag{12}$$

Then let

$$\Phi_{u2}(t,\tau) = \begin{cases} 0 & \text{for} \quad t < \tau \\ I + \int\limits_{\tau}^{t} [A(s)\Phi_{u1}(s,\tau) + A_1(s)\Phi_{u1}(s-h,\tau)]\, ds & \text{for} \quad t \geq \tau \end{cases} \tag{13}$$

and, in general, for iteration i let

$$\Phi_{u_i}(t,\tau) = \begin{cases} 0 & \text{for} \quad t < \tau \\ I + \int\limits_{t_o}^{t} [A(s)\Phi_{i-1}(s,\tau) + A_1(s)\Phi_{i-1}(s-h,\tau)]ds & \text{for} \quad t \geq \tau \end{cases} \tag{14}$$

where $\tau \geq t_o$. It can be shown [3.3] that the sequence $\{\Phi_{u_i}(t,\tau)\}$ uniformly converges to $\Phi_u(t,\tau)$. Δ

An alternate characterization of the solution of (1) can be obtained by applying the method of variation of parameters to the zero-input solution provided by Theorem 3.2.7 as follows.

4. Theorem. The solution of (1) can be written as

$$x(t \triangleq x(t,t_o,\phi,u) = \Psi(t,t_o)\phi(t_o) + \sum_{i=1}^{N} \int\limits_{t_o-h_{x_i}}^{t_o} \Psi(t,\tau+h_{x_i})A_i(\tau+h_{x_i})\phi(\tau)d\tau$$
$$+ \int\limits_{t_o}^{t} \Psi(t,\tau)[B(\tau)u(\tau) + \sum_{i=1}^{R} B_i(\tau)u(\tau-h_{u_i})]d\tau, \ t \geq t_o \tag{15}$$

where fundamental matrix $\Psi(t,\tau)$ satisfies (3.2.20a-c). Δ

The proof of the above Theorem is similar to that of Theorem 2 and will not be detailed here.

3.4 ADJOINT STATE EQUATIONS

In this section we will discuss the concept of adjoint state equation and will use it to determine a solution to the linear homogeneous TD state equation. For convenience, let us first concentrate on the case of a single state delay.

1. **Definition.** (Adjoint state equation) [3.7]. Consider the linear homogeneous state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h) \tag{1}$$

The *adjoint* of this state equation is

$$\dot{\mathbf{z}}(t) = -\mathbf{A}'(t)\mathbf{z}(t) - \mathbf{A}'_1(t+h)\mathbf{z}(t+h) \tag{2}$$

where **z** is a column vector, or

$$\dot{\mathbf{z}}(t) = -\mathbf{z}(t)\mathbf{A}(t) - \mathbf{z}(t+h)\mathbf{A}_1(t+h) \tag{3}$$

where **z** is a row vector. Δ

Note that the adjoint of a linear TD state equation involves advanced arguments. Also, it should be immediately noted that the concept of adjoint in TD systems is not reciprocal as in the case of nondelay systems. That is, the adjoint of state equation (2) or (3) is not state equation (1). In the remainder of this section we will use adjoint state equation (3) for (1), *i.e..*, where **z** is a row vector.

2. Theorem. [3.7]. Let $\mathbf{x}(t)$ and $\mathbf{z}(t)$ be arbitrary solutions of (1) and (3), respectively, and define their inner product as

$$(\mathbf{z}(t),\mathbf{x}(t)) = \mathbf{z}(t)\mathbf{x}(t) + \int_t^{t+h} \mathbf{z}(\tau)\mathbf{A}_1(\tau)\mathbf{x}(\tau-h)\,d\tau \tag{4}$$

Then $(\mathbf{z}(t),\mathbf{x}(t))$ is constant.

Proof. We will show that the derivative of the inner product (4) *w.r.t.* time is zero. From (4), we have

$$\frac{d}{dt}(\mathbf{z}(t),\mathbf{x}(t)) = \dot{\mathbf{z}}(t)\mathbf{x}(t) + \mathbf{z}(t)\dot{\mathbf{x}}(t) + \mathbf{z}(t+h)\mathbf{A}_1(t+h)\mathbf{x}(t) - \mathbf{z}(t)\mathbf{A}_1(t)\mathbf{x}(t-h) \tag{5}$$

which by (3) and (1) becomes

$$\frac{d}{dt}(\mathbf{z}(t),\mathbf{x}(t)) = [-\mathbf{z}(t)\mathbf{A}(t) - \mathbf{z}(t+h)\mathbf{A}_1(t+h)]\mathbf{x}(t) + \mathbf{z}(t)[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h)] + \mathbf{z}(t+h)\mathbf{A}_1(t+h)\mathbf{x}(t) - \mathbf{z}(t)\mathbf{A}_1(t)\mathbf{x}(t-h) = 0$$

Δ

We can use the properties of adjoint state equation to prove Theorem 3.2.5.

3. Proof of Theorem 3.2.5. Let $Z(\alpha,t)$ be a matrix solution to (3) (as a function of α) for $\alpha < t$ where $Z(t,t) = I$ and $Z(\alpha,t) = 0$ for $\alpha > t$. That is, let

$$\frac{\partial}{\partial\alpha} Z(\alpha,t) = -Z(\alpha,t)A(\alpha) - Z(\alpha+h,t)A_1(\alpha+h), \quad \alpha < t \tag{6}$$

Consider state equation (1) with t replaced by α and with initial function ϕ, *i.e.*,

$$\frac{d}{d\alpha} x(\alpha) = A(\alpha)x(\alpha) + A_1(\alpha)x(\alpha-h) \quad , \quad \alpha \geqslant t_o \tag{7a}$$

$$x(\alpha) = \phi(\alpha) \quad , \quad t_o-h \leqslant \alpha \leqslant t_o \tag{7b}$$

Multiply both sides of (7a) by $Z(\alpha,t)$ and integrate between t_o and t to obtain

$$\int_{t_o}^{t} Z(\alpha,t) \frac{d}{d\alpha} x(\alpha)d\alpha =$$

$$\int_{t_o}^{t} Z(\alpha,t)A(\alpha)x(\alpha)d\alpha + \int_{t_o}^{t} Z(\alpha,t)A_1(\alpha)x(\alpha-h)d\alpha \tag{8}$$

Integration by part of the left hand side of (8) yields

$$Z(\alpha,t)x(\alpha) \Big|_{t_o}^{t} - \int_{t_o}^{t} [\frac{\partial}{\partial\alpha} Z(\alpha,t)]x(\alpha)d\alpha = \int_{t_o}^{t} Z(\alpha,t)A(\alpha)x(\alpha)d\alpha$$

$$+ \int_{t_o}^{t} Z(\alpha,t)A_1(\alpha)x(\alpha-h)d\alpha \tag{9}$$

Considering the fact that $Z(t,t) = I$ and that $Z(\alpha,t)$ satisfies (6), from (9) we obtain

$$x(t) = Z(t_o,t)\phi(t_o) - \int_{t_o}^{t} Z(\alpha+h,t)A_1(\alpha+h)x(\alpha)d\alpha$$

$$+ \int_{t_o}^{t} Z(\alpha,t)A_1(\alpha)x(\alpha-h)d\alpha \tag{10}$$

Since $Z(\alpha,t) = 0$ for $\alpha > t$, we have $Z(\alpha+h,t) = 0$ for $\alpha > t - h$ and (10) becomes

$$x(t) = Z(t_o,t)\phi(t_o) + \int_{t-h}^{t_o} Z(\alpha+h,t)A_1(\alpha+h)x(\alpha)d\alpha + \int_{t_o-h}^{t-h} Z(\epsilon+h,t)A_1(\epsilon+h)x(\epsilon)d\epsilon$$

$$= Z(t_o,t)\phi(t_o) + \int_{t_o-h}^{t_o} Z(\alpha+h,t)A_1(\alpha+h)\phi(\alpha)d\alpha \tag{11}$$

Now let $\Psi(t,\alpha) = Z(\alpha,t)$. Then $\Psi(t,t) = I$ and $\Psi(t,\alpha) = 0$ for $\alpha > t$ and, from (11) we have

$$\mathbf{x}(t) = \Psi(t,t_o)\phi(t_o) + \int_{t_o-h}^{t_o} \Psi(t,\alpha+h)\mathbf{A}_1(\alpha+h)\phi(\alpha)d\alpha \tag{12}$$

which is the same as (3.2.15). This completes the proof of Theorem 3.2.5. Δ

4. Construction of matrix $\Psi(t,\alpha)$. The matrix $\mathbf{Z}(\alpha,t)$, and therefore fundamental matrix $\Psi(t,\alpha)$, can be constructed backwards in α as follows. For $\alpha = t$, $\mathbf{Z}(t,t) = \mathbf{I}$. For $t - h < \alpha < t$, we have $\mathbf{Z}(\alpha+h,t) = \mathbf{0}$ and

$$\frac{\partial}{\partial\alpha}\mathbf{Z}(\alpha,t) = -\mathbf{Z}(\alpha,t)\mathbf{A}(\alpha) \tag{13}$$

That is, $\mathbf{Z}(\alpha,t)$ in the interval $t - h < \alpha < t$ is determined by a system of homogeneous nondelay matrix differential equations. Call the solution to (13) $\mathbf{Z}^1(\alpha)$, *i.e.*, let

$$\mathbf{Z}(\alpha,t) = \mathbf{Z}^1(\alpha) \quad , \; t - h < \alpha < t \tag{14}$$

Now consider the interval $t - 2h < \alpha < t - h$. We have

$$\frac{\partial}{\partial\alpha}\mathbf{Z}(\alpha,t) = -\mathbf{Z}(\alpha,t)\mathbf{A}(\alpha) - \mathbf{Z}^1(\alpha+h)\mathbf{A}_1(\alpha+h) \tag{15}$$

which is a matrix differential equation with a known forcing function. Call the solution to (15) $\mathbf{Z}^2(\alpha)$, *i.e.*,

$$\mathbf{Z}(\alpha,t) = \mathbf{Z}^2(\alpha) \quad , \; t - 2h < \alpha < t - h \tag{16}$$

Similarly, in the interval $t - 3h < \alpha < t - 2h$ we have

$$\frac{\partial}{\partial\alpha}\mathbf{Z}(\alpha,t) = -\mathbf{Z}(\alpha,t)\mathbf{A}(\alpha) - \mathbf{Z}^2(\alpha+h)\mathbf{A}_1(\alpha+h) \tag{17}$$

We can continue this procedure to obtain $\mathbf{Z}(\alpha,t)$ for $\alpha < t$. Δ

The definition of adjoint state equation can be extended to multiple delay systems.

5. Definition. Consider the linear homogeneous multiple-delay state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_i) \tag{18}$$

The adjoint of this state equation is

$$\dot{\mathbf{z}}(t) = -\mathbf{z}(t)\mathbf{A}(t) - \sum_{i=1}^{N} \mathbf{z}(t+h_i)\mathbf{A}_i(t+h_i) \tag{19}$$

where $\mathbf{z}$ is a row vector.

6. Theorem. Let $\mathbf{x}(t)$ and $\mathbf{z}(t)$ be arbitrary solutions of (18) and (19), respectively, and define their inner product as

$$(\mathbf{z}(t),\mathbf{x}(t)) = \mathbf{z}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \int_{t}^{t+h_i} \mathbf{z}(\tau)\mathbf{A}_i(\tau)\mathbf{x}(\tau-h_i)d\tau \tag{20}$$

Then $(\mathbf{z}(t),\mathbf{x}(t))$ is a constant. Δ

The proof of the above theorem is a straightforward extension of the proof of Theorem 3.3.2 and is left to the reader.

Theorem 3.2.7 can be proved in a manner similar to that used for the proof of Theorem 3.2.5. That is, construct a matrix solution $\mathbf{Z}(\alpha,t)$ for (19); then multiply both sides of (18) (with t replaced by α) by $\mathbf{Z}(\alpha,t)$ and integrate from t_o to t to obtain the solution. Finally let $\Psi(t,\alpha) = \mathbf{Z}(\alpha,t)$.

PROBLEMS

3.1. Show that fundamental matrix $\Phi(t,\tau)$ in (3.2.18) can be constructed by the following iterative method (subscript k indicates kth iteration) [3.3]:

$$\Phi_1(t,\tau) = \begin{cases} \mathbf{I}\delta(t-\tau) & \text{for} \quad t,\tau\epsilon[t_o-\Delta_x,t_o] \\ 0 & \text{for} \quad t > t_o \end{cases}$$

$$\Phi_2(t,\tau) = \begin{cases} \mathbf{I}\delta(t-\tau) & \text{for} \quad t,\tau\epsilon[t_o-\Delta_x,t_o] \\ \displaystyle\int_{t_o}^{t} [\mathbf{A}(s)\Phi_1(s,\tau) + \sum_{i=1}^{N} \mathbf{A}_i(s)\Phi_1(s-h_{x_i},\tau)]ds & \text{for} \quad t > t_o \end{cases}$$

and, in general,

$$\Phi_{k+1}(t,\tau) = \begin{cases} I\delta(t-\tau) & \text{for } t,\tau\epsilon[t_o-\Delta_x,t_o] \\ \int_{t_o}^{t} [A(s)\Phi_k(s,\tau) + \sum_{i=1}^{N} A_i(s)\Phi_k(s-h_{x_i},t)]ds & \text{for } t > t_o \end{cases}$$

3.2. Let $A_1(t) \equiv 0$ in (3.2.7a) and show that the result of Theorem 3.2.3 is consistent with the solution of the linear homogeneous state equation in the nondelay case [3.5].

3.3. Repeat Problem 3.2 for the result of Theorem 3.2.5.

3.4. Prove Theorem 3.3.2 for the case with no control delay. Then use superposition property (Theorem 3.3.1) to prove Theorem 3.3.2 in general.

3.5. Show that matrix $\Phi_u(t,\tau)$ in (3.3.9) can be constructed by the following iterative method (subscript k indicates kth iteration) [3.3]:

$$\Phi_{u_1}(t,\tau) = \begin{cases} 0 & \text{for } t < \tau \\ I & \text{for } t \geqslant \tau \end{cases}$$

$$\Phi_{u2}(t,\tau) = \begin{cases} 0 & \text{for } t < \tau \\ I + \int_{\tau}^{t} [A(s)\Phi_{u1}(s,\tau) + \sum_{i=1}^{N} A_i(s)\Phi_{u_1}(s-h_{x_i},\tau)]ds & \text{for } t \geqslant \tau \end{cases}$$

and, in general,

$$\Phi_{k+1}(t,\tau) = \begin{cases} 0 & \text{for } t < \tau \\ I + \int_{\tau}^{t} [A(s)\Phi_{u_k}(s,\tau) + \sum_{i=1}^{N} A_i(s)\Phi_{u_k}(s-h_{x_i},\tau)]ds & \text{for } t \geqslant \tau \end{cases}$$

3.6. Prove Theorem 3.3.4.

3.7. Consider the homogeneous state equation (3.4.1) and its adjoint (3.4.3) with initial function specified in the interval $[T,T+h]$. Show that the solution to (3.4.3) is

$$z(t) = z(T)\Psi(T,t) + \int_{T}^{T+h} z(\alpha)A_1(\alpha)\Psi(\alpha-h,t)d\alpha \quad \text{for } t \leqslant T$$

where $\Psi(t,\tau)$ is defined in Theorem 3.2.5. *Hint:* Multiply both sides of (3.4.3) by $\Psi(\alpha,t)$ and integrate between t and T.

3.8. Prove Theorems 3.4.6 and 3.2.7.

REFERENCES

[3.1] R. D. Driver, "Existence and stability of solutions of a delay differential system," *Arch. Rational Mech. Anal.*, Vol. 10, pp. 401-426, 1962

[3.2] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, 1977

[3.3] M. N. Oguztoreli, *Time-Lag Control Systems*, Academic Press, 1966

[3.4] L. A. Zadeh and C. A. Desoer, *Linear System Theory*, McGraw-Hill, 1963

[3.5] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach*, Pergamon Press, 1986

[3.6] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955

[3.7] A. Halanay, *Differential Equations - Stability, Oscillation and Time Lags*, Academic Press, 1966

CHAPTER 4

STABILITY OF TIME-DELAY SYSTEMS

4.1 INTRODUCTION

Stability is an important property of any control system, delayed or nondelayed. The stability of TD systems is relatively more difficult because of the delay terms in system model, the characteristic equation is transcendental rather than algebraic in nature. In this Chapter the stability of TD systems described in either time-domain or frequency-domain is discussed. Consider the state equation of a *l.t.i.* homogeneous TD system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i \mathbf{x}(t-h_i) \tag{1a}$$

where $\mathbf{x}$ is $n \times 1$ state vector, $\mathbf{A}_i$ are $n \times n$ constant matrices and h_i, $i=1,\ldots,N$ are the system's constant delays. The initial condition of (1a) is represented by a function

$$\mathbf{x}(\sigma) = \phi(\sigma) \quad , \quad -\Delta \leqslant \sigma \leqslant 0 \tag{1b}$$

where $\Delta = \max_i h_i$.

For TD system (1), we begin with the definition of an equilibrium state.

1. Definition. A state x_e of system (1) is said to be an ***equilibrium state*** if

$$\mathbf{x}(t_o,\phi) = \mathbf{x}_e \implies \mathbf{x}(t;\phi) = \mathbf{x}_e \quad \text{for all } \; t > t_o \tag{2}$$

provided that no input is applied. Δ

In general, system (1) may have more than one equilibrium state, but for our present discussion we can assume that the equilibrium state is $\mathbf{x}_e = 0$, *i.e.*, the origin. This assumption is not restrictive. Other equilibrium points can be similarly used. Now let us consider the stability definitions given below.

2. Definition. The equilibrium state $\mathbf{x}_e = 0$ is ***stable in the sense of Lyapunov***

(abbreviated *i.s.L.*) or simply stable as $t \to \infty$, if, given, any positive numbers t_o and ϵ, there exists a number $\delta > 0$ such that every continuous solution $\mathbf{x}(t)$ of (1) which satisfies

$$\max_{t_o \leq t \leq t_o + \Delta} \left| \mathbf{x}(t) \right| \leq \delta \tag{3}$$

will also satisfy

$$\max_{t_o \leq t < \infty} \left| \mathbf{x}(t) \right| \leq \epsilon \tag{4}$$

Intuitively speaking, the above definition indicates that the solution of (1) is stable if every solution which is initially small will remain small for all t. It should be noted that the number δ in general depends on both t_o and ϵ [4.1]. If a $\delta > 0$ can be found independent of t_o, the solution $\mathbf{x}_e$ of (1) is said to be *uniformly stable w.r.t.* t_o, or more briefly *uniformly stable*. That is, we have the following definition.

3. Definition. The equilibrium state $\mathbf{x}_e = 0$ is *uniformly stable* if, given $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that if $\mathbf{x}(t)$ satisfies (3) for any $t_o \geq 0$, then $\mathbf{x}(t)$ satisfies (4). Δ

Another important concept is asymptotic stability which will be defined next.

4. Definition. The equilibrium state $\mathbf{x}_e = 0$ is *asymptotically stable* if:

(i) it is stable;

(ii) for each $t_o > 0$ there is a $\delta = \delta(t_o)$ such that every solution which satisfies (3) also satisfies

$$\lim_{t \to \infty} \mathbf{x}(t) = 0 \tag{5}$$

An interpretation of the above definition is that all solutions initiating with small states must tend to the origin. However, in TD systems, like many other systems [4.2], initial function variations may in fact be very large. This would lead us to the following definition.

5. Definition. The equilibrium state $\mathbf{x}_e = 0$ is *asymptotically stable in the large* if:

(i) it is stable;

(ii) every solution satisfies (5). Δ

The remainder of this chapter is devoted to two main approaches for checking the stability (*i.s.L.*, asymptotic or uniform) of TD systems described in time-domain or in frequency-domain. Section 4.2 discusses Lyapunov and uniform stability, stability of stochastic TD systems, stability of variable delay systems and stability via the Lyapunov method. An attempt is made to cover the most recent research results in literature. The stability of TD systems described in frequency domain is discussed in Section 4.3.

4.2 STABILITY OF LINEAR TIME-DELAY SYSTEMS VIA TIME-DOMAIN

In this section the stability of linear TD systems described by their state-space representations is discussed. Both stochastic and deterministic systems with time delay are considered. The effects of delay variations with respect to time and the system's state, on the overall stability of TD systems are presented next. The final topic is the application of the Lyapunov method to the stability of TD systems.

The results of this section are based, in part, on the works of Kalman and Bertram [4.3], Nazaroff [4.4], Nazaroff and Hewer [4.5] and Hirai and Satoh [4.6].

4.2.1 The Lyapanov Method

In this section Lyapanov's second method [4.3] is used to find necessary and sufficient conditions for the stability of a class of *l.t.i.* TD systems represented as

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{x}(t-h) \tag{1}$$

This discussion is based on the work of Lee and Dianat [4.7].

Consider a hermitian matrix **P**. It is known that the function $\nu(\mathbf{z}) = \mathbf{z}'\mathbf{P}\mathbf{z}$ of any vector **z** is positive if and only if matrix **P** is positive definite (see Appendix A, Section 9). If vector **z** is defined as

$$\mathbf{z}(t) = \mathbf{x}(t) + \mathbf{Q}(t)*\mathbf{x}(t) \tag{2}$$

where $*$ denotes the convolution operator, the same conclusion for any vector **x** is carried over, provided that the transformation $\mathbf{Q}(t)*\mathbf{x}(t)$ exists. As a preliminary to the main theorem, consider the following lemma.

1. Lemma. Let a scalar function $\nu(.)$ be defined by

$$\nu(\mathbf{x}_t,h) = (\mathbf{x}(t) + \mathbf{Q}(t)\bullet\mathbf{x}(t))'\ \mathbf{P}(\mathbf{x}(t) + \mathbf{Q}(t)\bullet\mathbf{x}(t)) \tag{3}$$

where the $n\times n$ matrix $\mathbf{Q}(t)$ which is continuous and differentiable over $[0,h]$, is a characteristic matrix, $\mathbf{P}$ is hermitian and $\mathbf{x}_t(\sigma) \triangleq \mathbf{x}(t+\sigma)$ for $\sigma \in [-h,0]$. If

$$\mathbf{P}\ (\mathbf{A} + \mathbf{Q}(0)) + (\mathbf{A}+\mathbf{Q}(0))'\mathbf{P} + \mathbf{R} = \mathbf{0} \tag{4}$$

$$\dot{\mathbf{Q}} = (\mathbf{A}+\mathbf{Q}(0))\ \mathbf{Q}(\sigma),\ 0 \leqslant \sigma \leqslant h \tag{5}$$

where $\mathbf{Q}(h) = \mathbf{B}$, and $\mathbf{R}$ is positive definite. Then

$$\dot{\nu}(\mathbf{x}_t,h) = d\nu(\mathbf{x}_t,h)/dt < 0. \tag{6}$$

Proof. Consider the definition of $\nu(\cdot)$ in (3) and rewrite it as

$$\begin{aligned}\nu(\mathbf{x}_t,h) = {} & \mathbf{x}'\mathbf{P}\mathbf{x} + \mathbf{x}'\mathbf{P}\int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma + \left[\int_0^h \mathbf{x}'(t-\sigma)\ \mathbf{Q}(\sigma)d\sigma\right]\mathbf{P}\mathbf{x} \\ & + \int_0^h\int_0^h \mathbf{x}'(t-\eta)\ \mathbf{Q}_1'(\eta)\ \mathbf{P}\mathbf{Q}(\sigma)\ \mathbf{x}\ (t-\sigma)d\eta d\sigma.\end{aligned} \tag{7}$$

Taking the derivative *w.r.t.* t yields

$$\begin{aligned}\dot{\nu}(\mathbf{x}_t,\tau) = {} & \left[\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}(t-h) + d/dt\int_0^h \mathbf{Q}(\sigma)\ \mathbf{x}(t-\sigma)d\sigma\right]' \\ & \times\mathbf{P}\left[\mathbf{x} + \int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma\right] + \left[\mathbf{x} + \int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma\right]' \\ & \times\mathbf{P}\left[\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}(t-h) + d/dt\int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma\right]\end{aligned} \tag{8}$$

Now integrating by parts and simplifying would yield

$$\begin{aligned}\dot{\nu}(\mathbf{x}_t,h) = \mathbf{x}' & \Big[\mathbf{A'P} + \mathbf{Q'}(0)\mathbf{P} + \mathbf{PA} + \mathbf{PQ}(0) \Big] \mathbf{x} \\ & + \mathbf{x}' \Big[\mathbf{PB} - \mathbf{PQ}(h) \Big] \mathbf{x}(t-\sigma) + \mathbf{x}'(t-h) \Big[\mathbf{B'P} - \mathbf{Q'}(h)\,\mathbf{P} \Big] \mathbf{x} \\ & + \mathbf{x}' \int_0^h \Big[\mathbf{A'PQ}(\sigma) + \mathbf{Q'}(0)\,\mathbf{PQ}(\sigma) + \mathbf{P\dot{Q}}(\sigma) \Big] \mathbf{x}(t-\sigma)d\sigma \\ & + \Big[\int_0^h \mathbf{x}'(t-\sigma) \Big[\mathbf{\dot{Q}'}(\sigma)\,\mathbf{P} + \mathbf{Q'}\,(\sigma)\mathbf{PA} + \mathbf{Q'}(\sigma)\,\mathbf{PQ}(0) \Big] d\sigma \Big] \mathbf{x} \\ & + \mathbf{x}(t-h) \int_0^h \Big[\mathbf{B'PQ}(\sigma) - \mathbf{Q'}(\sigma)\,\mathbf{PQ}(\sigma) \Big] \mathbf{x}(t-\sigma)\,d\sigma \\ & + \Big[\int_0^h \mathbf{x}'(t-\sigma) \Big[\mathbf{Q'}(\sigma)\mathbf{PB} - \mathbf{Q'}\,(\sigma)\mathbf{PQ}(\tau) \Big] d\sigma \Big] \mathbf{x}(t-h) \\ & + \int_0^h \int_0^h \mathbf{x}'(t-\eta) \Big[\mathbf{\dot{Q}'}(\eta)\mathbf{PQ}(\sigma) + \mathbf{Q'}(\eta)\mathbf{P\dot{Q}}(\sigma) \Big] \mathbf{x}(t-\sigma)d\eta d\sigma \end{aligned} \tag{9}$$

If (4) and (5) are utilized in (9), the following equation would result:

$$\dot{\nu}(\mathbf{x}_t,h) = \Big[\mathbf{x} + \int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma \Big]' (-\mathbf{R}) \Big[\mathbf{x} + \int_0^h \mathbf{Q}(\sigma)\mathbf{x}(t-\sigma)d\sigma \Big] \tag{10}$$

Noting that **R** is positive definite, it follows that $\dot{\nu}(\cdot) < 0$ and the lemma is proved.

Utilizing the above lemma one can prove the following theorem.

2. Theorem. A sufficient condition for asymptotic stability of (1) is that there exists a positive definite hermitian matrix **P** such that $\mathbf{P}[\,\mathbf{A} + \mathbf{Q}(0)] + [\,\mathbf{A} + \mathbf{Q}(0)]'\,\mathbf{P} + \mathbf{R} = 0$ where **R** is a positive definite hermitian matrix and for $\sigma \in [\,0,h\,]$, $\mathbf{Q}(\sigma)$ satisfies

$$\dot{\mathbf{Q}}(\sigma) = [\,\mathbf{A} + \mathbf{Q}(0)\,]\,\mathbf{Q}(\sigma) \tag{11a}$$

with

$$\mathbf{Q}(h) = \mathbf{B} \text{ and } \mathbf{Q}(\sigma) = \mathbf{0} \text{ elsewhere.} \tag{11b}$$

Proof. By defining a Lyapunov function $\nu(\mathbf{x}_t,h)$ as in (3) for the system (1) and noting that $\mathbf{x}(t) + \mathbf{Q}(t) * \mathbf{x}(t) = 0$ if and only if $\mathbf{x}(t) = 0$, it then follows from Lemma 1 that $\nu(\mathbf{x}_t,h) > 0$ and $\dot{\nu}(\mathbf{x}_t,h) < 0$. This proves the theorem. Δ

To interpret the sufficiency condition (11) more closely, it is required to find a solution for it, *i.e.*,

$$Q(\sigma) = \left[\exp(A + Q(0))\sigma \right] Q(0) \tag{12}$$

evaluating (12) at $\sigma = h$ and noting the first of the two conditions in (11b) the sufficiency condition (11) reduces to the solution of the following nonlinear algebraic matrix equation:

$$\left[\exp(A + Q(0))h \right] Q(0) = B \tag{13}$$

for Q(0). The following theorem provides a necessary and sufficient condition for asymptotic stability of (1).

3. Theorem. Let the TD system be described by (1) and let Q(0) be a nonsingular solution of (13). Then a necessary and sufficient condition for asymptotic stability is that

$$\left| \alpha_i \right| = \left| \lambda_i \{ BQ^{-1}(0) \} \right| < 1 \tag{14}$$

for all i where α_i are distinct.

Proof. It is clear from (13) that

$$\exp\left[A + Q(0) \right] h = BQ^{-1}(0) \tag{15}$$

Then introducing $\alpha_i = \exp(\beta_i)$ where $\beta_i = \lambda_i \{A + Q(0)\}$, it follows that $|\alpha_i| < 1$ implies Re $\{ \beta_i \} < 0$. Therefore, the real parts of the eigenvalues of A + Q(0) are negative. Thus, the sufficiency part of the theorem follows immediately by Theorem 2.

For the necessity part of the theorem, assume that the system is asymptotically stable, then it follows that the roots of the characteristic equation

$$\det\left[sI - A - B\exp(-hs) \right] = 0 \tag{16}$$

all have negative real parts. Substituting (13) into (16) results in

$$\det\left[sI - A - \exp\left[A + Q(0) - sI \right] hQ(0) \right] = 0 \tag{17}$$

Expanding the exponential term in (17) in a power series yields

$$det\left[\mathbf{D} + \mathbf{Q}(0) - \left[\mathbf{I} - \mathbf{D}h + (\mathbf{D}h)^2/2! - (\mathbf{D}h)^3/3! + \cdots \right] \mathbf{Q}(0) \right]$$

$$= det\left[\mathbf{D} \left[\mathbf{I} + \mathbf{Q}(0)h - (\mathbf{D}h^2)\, \mathbf{Q}(0)/2! + (\mathbf{D}^2h^3)\, \mathbf{Q}(0)/3! - \cdots \right] \right]$$

$$= det(\mathbf{D}) det\left[\mathbf{I} + \mathbf{Q}\,(0)h - (\mathbf{D}h^2)\, \mathbf{Q}\,(0)/2! + \cdots \right] = 0 \tag{18}$$

where $\mathbf{D} \triangleq s\mathbf{I} - \mathbf{A} - \mathbf{Q}(0)$. Now since $\mathbf{I} + \mathbf{Q}(0)h - (\mathbf{D}h^2)\,\mathbf{Q}(0)/2! + \cdots$ converges to an entire function, then the eigenvalues $\lambda_i\{\, \mathbf{A} + \mathbf{Q}(0)\, \}$ are contained in the spectrum β_j of the characteristic equation $det(s\mathbf{I} - \mathbf{A} - \mathbf{B}\, exp(-hs)\,)$ for all i and some j. Since the system (1) is asymptotically stable, then $Re\{\lambda_i(\mathbf{A} + \mathbf{Q}(0))\} < 0$ for all i. Then by a theorem due to Kalman and Bertram [4.3] for every positive definite hermitian matrix $\mathbf{R}$, there exists a positive definite hermitian matrix $\mathbf{P}$ such that $\mathbf{P}\ (\mathbf{A} + \mathbf{Q}(0)) + (\mathbf{A} + \mathbf{Q}(0))^* \,\mathbf{P} = -\mathbf{R}$. Now by the hypothesis of this theorem, $\mathbf{Q}(\sigma)$ for $\sigma \in [0,h]$ exists and $\mathbf{Q}(\sigma)$ satisfies the differential equation $\dot{\mathbf{Q}}(\sigma) = (\,\mathbf{A} + \mathbf{Q}(0))\ \mathbf{Q}(\sigma)$ with boundary condition $\mathbf{B} = (exp\ (\mathbf{A} + \mathbf{Q}(0))h\mathbf{Q}(0)$. This proves the theorem.

4. Example. Consider a second-order system discussed by Lee and Dianat [4.7]:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & -3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1.66 & -0.697 \\ 0.93 & -0.330 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix} \tag{19}$$

It is desired to check whether it is asymptotically stable.

Following the development of Theorem 2 one needs to solve the algebraic matrix equation (13) rewritten as

$$\mathbf{Q}(0) = exp\left[-(\mathbf{A} + \mathbf{Q}(0)) \right]\mathbf{B} \tag{20}$$

This equation can be solved by numerical technique such as Newton's method. The solution to (20) after a few iterations turns out to be

$$\mathbf{Q}(0) = \begin{bmatrix} 1.99 & 0 \\ 0 & 1 \end{bmatrix} \approx \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \tag{21}$$

which can be used to solve equation (11) to result

$$Q(\sigma) = \begin{bmatrix} 6e^{-\sigma} - 4e^{-2\sigma} & -3e^{-\sigma} + 3e^{-2\sigma} \\ 4e^{-\sigma} - 4e^{-2\sigma} & -2e^{-\sigma} + 3e^{-2\sigma} \end{bmatrix} . \tag{22}$$

Now using $\mathbf{R} = diag\ (6,6)$, one can solve $\mathbf{P}$ from (4) to obtain

$$\mathbf{P} = \begin{bmatrix} 11 & -7 \\ -7 & 6 \end{bmatrix} \tag{23}$$

which is positive definite. Thus by Theorem 2, the system is asymptotically stable with the following Lyapunov function $\nu(\mathbf{x}(t))$ and its derivative $\dot{\nu}(\mathbf{x}(t))$;

$$\nu(\mathbf{x}(t)) = \left[\mathbf{x}(t) + \mathbf{Q}(t)*\mathbf{x}(t)\right]*\mathbf{P}\left[\mathbf{x}(t) + \mathbf{Q}(t)*\mathbf{x}(t)\right] \tag{24}$$

$$\dot{\nu}(\mathbf{x}(t)) = -6\left[\mathbf{x}(t) + \mathbf{Q}(t)*\mathbf{x}(t)\right]*\left[\mathbf{x}(t) + \mathbf{Q}(t)*\mathbf{x}(t)\right] \tag{25}$$

4.2.2 Uniform Asymptotic Stability

Uniform asymptotic stability was defined in Section 4.1. Here the stability of linear deterministic TD systems with multiple delays will be of concern. Consider the TD system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \sum_{i=1}^{j} \mathbf{A}_i\mathbf{x}(t-h_i)\ ,\ t_o = 0 \tag{26}$$

where $\mathbf{x}$ is an $n\times 1$ vector, $\mathbf{A}$, $\mathbf{A}_i, i=1,\ldots,j$ are constant $n\times n$ matrices and $\Delta = \max_i h_i$. The problem is to find conditions under which system (26) is uniformly asymptotically stable (UAS). The characteristic equation of (26), as it was seen in detail in Section 3.4, is transcendental and checking that its roots are in the *l.h.p* is not, in general, an easy undertaking. There are "small delay" Theorems where easy conditions can be checked for stability [4.8]. These Theorems, although easy to apply, are not appropriate for most TD system whose delays are not necessarily small. In order to present the conditions for UAS of TD system consider first the following Lemma [4.9,4.10].

5. Lemma. Consider the partitioned matrix

$$\mathbf{F} = \begin{bmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{D}' & \mathbf{E} \end{bmatrix} \tag{27}$$

where $\mathbf{C} = \mathbf{C}'$, $\mathbf{E} = \mathbf{E}'$ and dimensions of $\mathbf{C}$, $\mathbf{E}$ and $\mathbf{D}$ are $q \times q$, $r \times r$ and $q \times r$, respectively. Then,

I) $\mathbf{F}$ is nonnegative definite if and only if either

$$I_a : \begin{cases} \mathbf{E} \geqslant 0 & (28a) \\ \mathbf{D} = \mathbf{LE} & (28b) \\ \mathbf{C} - \mathbf{LEL}' \geqslant 0 & (28c) \end{cases}$$

or

$$I_b : \begin{cases} \mathbf{C} \geqslant 0 & (29a) \\ \mathbf{D} = \mathbf{CM} & (29b) \\ \mathbf{E} - \mathbf{M}'\mathbf{C}M \geqslant 0 & (29c) \end{cases}$$

hold, where $\mathbf{L}$ and $\mathbf{M}$ are some (not necessarily unique) matrices of dimension $q \times r$. Matrices $\mathbf{L}$ and $\mathbf{M}$ exist if and only if, respectively,

$$rank\ [\mathbf{D}'\mathbf{E}] = rank\ \mathbf{E} \tag{30a}$$

$$rank\ [\mathbf{CD}] = rank\ \mathbf{C} \tag{30b}$$

II) $\mathbf{F}$ is positive definite if and only if either

$$II_a : \begin{cases} \mathbf{E} > 0 & (31a) \\ \mathbf{C} - \mathbf{DE}^{-1}\mathbf{D}' > 0 & (31b) \end{cases}$$

or

$$II_b : \begin{cases} \mathbf{C} > 0 & (32a) \\ E - \mathbf{D}'\,\mathbf{C}^{-1}\mathbf{D} > 0 & (32b) \end{cases}$$

holds

Proof. I) Clearly $\mathbf{E} \geqslant 0$ is necessary. Partitioning $\mathbf{x}' = [\mathbf{x}'_1\ \mathbf{x}'_2]$ and in accordance to the partitioning of $\mathbf{F}$, the quadratic term $\mathbf{x}'\mathbf{F}\mathbf{x}$ is given by

$$\mathbf{x}'\mathbf{F}\mathbf{x} = \mathbf{x}'_1\mathbf{C}\mathbf{x}_1 + 2\ \mathbf{x}'_1\mathbf{D}\mathbf{x}_2 + \mathbf{x}'_2\mathbf{E}\mathbf{x}_2 \tag{33}$$

Let $\mathbf{x}_2$ be such that $\mathbf{E}\mathbf{x}_2 = \mathbf{0}$. If $\mathbf{D}\mathbf{x}_2 \neq 0$, let $\mathbf{x}_1 = -\ \alpha\ \mathbf{D}\mathbf{x}_2$, $\alpha > 0$. Then

$$\mathbf{x}'\mathbf{F}\mathbf{x} = \alpha^2\ \mathbf{x}'_2\mathbf{D}'\mathbf{C}\mathbf{D}\ \mathbf{x}_2 - 2\ \alpha\ \mathbf{x}'_2\mathbf{D}'\mathbf{D}\ \mathbf{x}_2 \tag{34}$$

which is negative for a sufficiently small $\alpha > 0$. Therefore it is necessary that $\mathbf{D}\mathbf{x}_2 = \mathbf{0}$ for all $\mathbf{x}_2$ such that $\mathbf{E}\ \mathbf{x}_2 = \mathbf{0}$. This means that the rows of $\mathbf{D}$ must be a linear combination of those of $\mathbf{E}$, *i.e.*

$$\mathbf{D} = \mathbf{L}\mathbf{E} \tag{35}$$

for some nonunique $\mathbf{L}$.

Since $\mathbf{E} \geqslant 0$, the quadratic form (33) has, for any $\mathbf{x}_1$, a minimum over $\mathbf{x}_2$. Thus differentiating (33) *w.r.t.* $\mathbf{x}_2$ one has

$$\partial(\mathbf{x}'\ \mathbf{F}\mathbf{x})/\partial\mathbf{x}_2 = 2\ \mathbf{D}'\ \mathbf{x}_1 + 2\ \mathbf{E}\ \mathbf{x}_2 = 2\ \mathbf{E}\ \mathbf{L}'\ \mathbf{x}_1 + 2\ \mathbf{E}\ \mathbf{x}_2 = \mathbf{0} \tag{36}$$

whence

$$\mathbf{E}\ \mathbf{L}'\ \mathbf{x}_1 = -\ \mathbf{E}\ \mathbf{x}_2\ . \tag{37}$$

Using (35) and (37) in (33), one finds that the minimum of $\mathbf{x}'\mathbf{F}\ \mathbf{x}$ over $\mathbf{x}_2$ and for any $\mathbf{x}_1$ is

$$\underset{\mathbf{x}_2}{Min}\ \mathbf{x}'\ \mathbf{F}\ \mathbf{x} = \mathbf{x}'_1(\mathbf{C} - \mathbf{L}\mathbf{E}\mathbf{L}')\mathbf{x}_1'. \tag{38}$$

Thus (28c) is necessary.

The conditions I_a in (28) are thus necessary for $\mathbf{F} \geqslant \mathbf{0}$, and since together they imply that the minimum of $\mathbf{x}'\mathbf{F}\mathbf{x}$ over $\mathbf{x}_2$ for any $\mathbf{x}_1$ is nonnegative, they are also sufficient. One notes that the condition $\mathbf{C} > \mathbf{0}$ of I_b in (29a), which is also necessary, is implied by (28a) and (28c):

$$\mathbf{C} \geqslant \mathbf{L}\mathbf{E}\mathbf{L}' \geqslant \mathbf{0} \tag{39}$$

The condition $\mathbf{D} = \mathbf{C}\mathbf{M}$ of (29b) is also implied in (28c):

$$x'_1Cx_1 - x'_1LEL'x_1 = x'_1Cx_1 - x'_1DL'x_1 \geqslant 0 \tag{40}$$

shows that x'_1D must vanish for all x_1 such that $x'_1C = 0$. In other words, the columns of D must be a linear combination of those of C. Clearly, conditions I_b in (29) can be derived as those of I_a in (28) by beginning with C. The equivalence between conditions (30a), (30b) and (28b), (29b) is easily seen.

II) Conditions II_a in (31) and II_b in (32) are proved along the same lines as those in I_a *and* I_b. For further insight the reader may consult References [4.9] and [4.10].

Moreover, the reader is reminded that a matrix A is said to be stable if all its eigenvalues are in the left-half s-plane, *i.e.*, Re $\{\lambda(A)\} < 0$ [4.2]. Now consider the following theorem.

6. Theorem. Let A be a stable matrix. If matrices Q and R which are symmetric and are associated with the Lyapunov equation

$$A'Q + QA + (j+1)\, R = 0 \tag{41}$$

satisfy

$$R > 0 \tag{42a}$$

$$R - \sum_{i=1}^{j} QA_iR^{-1}A'_iQ > 0 \tag{42b}$$

then system (26) is UAS. Note that this scheme is independent of the values of delays.

Proof. Define the Lyapunov functional

$$\nu(x(t),t) = x'(t)\; Qx(t) + \sum_{i=1}^{j} \int_{t-h_i}^{t} x'(\sigma)\; Rx(\sigma)\; d\sigma\,. \tag{43}$$

Then, sufficient conditions for the TD system to be UAS for all t and $x \neq 0$ are

$$d_1 \,\|\, x \,\| \leqslant \nu(x,t) \leqslant d_2 \,\|\, x \,\|\,,\; d_1, d_2 > 0 \tag{44a}$$

$$\dot{\nu}\,(x,t) \leqslant -d_2 \,\|\, x \,\|^2 < 0\,,\; d_2 > 0 \tag{44b}$$

where $\nu(0,0) = 0$.

Since A is stable and $\mathbf{R} > 0$, then matrix Q in (41) is also positive definite which implies that $\nu(\mathbf{x},t)$ in (43) satisfies condition (44a). In order to show (44b), differentiate (43) *w.r.t.* time and evaluate it along the trajectories $(\mathbf{x}(t),t)$ to obtain

$$\dot{\nu}(\mathbf{x}(t),t) = \mathbf{x}'(t)\ (\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A})\ \mathbf{x}(t) + \sum_{i=1}^{j} \mathbf{x}'(t-h_i)\ \mathbf{A}_i'\mathbf{Q}\mathbf{x}(t)$$

$$+ \sum_{i=1}^{j} \mathbf{x}'(t)\ \mathbf{Q}\mathbf{A}_i\mathbf{x}(t-h_i) - \sum_{i=1}^{j} \mathbf{x}'(t)\ \mathbf{R}\mathbf{x}(t-h_i) \tag{45}$$

which can be rewritten as

$$\dot{\nu}(\mathbf{x}(t),t) = - \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-h_1) \\ \\ \vdots \\ \mathbf{x}(t-h_j) \end{bmatrix}' \begin{bmatrix} \mathbf{R} & -\mathbf{Q}\mathbf{A}_1 & \cdots -\mathbf{Q}\mathbf{A}_j \\ -\mathbf{A}'_1\mathbf{Q} & \mathbf{R} & 0 \\ & 0 & \\ \vdots & \vdots & \vdots \\ -\mathbf{A}'_j\mathbf{Q} & 0 & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-h_1) \\ \\ \vdots \\ \mathbf{x}(t-h_j) \end{bmatrix} \tag{46}$$

Now, the hypothesis (42b) and Lemma 1 would imply that $\dot{\nu}(\mathbf{x}(t),t) \leqslant -d_2 \,||\, \mathbf{x} \,||^2$, or condition (44b). This proves the theorem.

7. Example. Consider the following two-delay *l.t.i.* TD system.

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \mathbf{x}(t-1) + \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}(t-2) \tag{47}$$

It is desired to see whether this system is UAS.

Matrix A has both its eigenvalues at -1, hence it is stable and $h_1 = 1$ and $j=1$. In order to test whether (47) is UAS one needs to see whether a pair of symmetric matrices (Q,R) can be obtained which satisfy (42). Let $\mathbf{R} = \mathbf{I}_2$ which can be used in (41) to find a matrix Q from

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{Q} + \mathbf{Q} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 0 \tag{48}$$

This leads to

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{49}$$

In order to see whether (42) is satisfied, one needs to compute

$$\mathbf{R} - \mathbf{Q}\mathbf{A}_1\mathbf{R}^{-1}\mathbf{A}'_1\mathbf{Q} = \begin{bmatrix} 3/4 & 0 \\ 0 & 3/4 \end{bmatrix} \tag{50}$$

which is positive definite, leading to the conclusion that the system is UAS.

8. Example. For the TD system

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix} \mathbf{x}(t-h) \tag{51}$$

it is desired to find a region on the b-a plane such that the system is UAS.

Here, let $\mathbf{R} = \frac{1}{2}\,\mathbf{I}_2$ and find $\mathbf{Q}$ from (41) as

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \tag{52}$$

The second condition in (42) can be calculated as

$$\mathbf{R} - \mathbf{Q}\mathbf{A}_1\mathbf{R}^{-1}\mathbf{A}'_1\mathbf{Q} = \begin{bmatrix} (1-a^2)/2 & 0 \\ 0 & (1-b^2)/2 \end{bmatrix} \tag{53}$$

which would be positive definite if $-1 < a < 1$ and $-1 < b < 1$ resulting in a square region shown in Figure 1 (independent of the value of delay h).

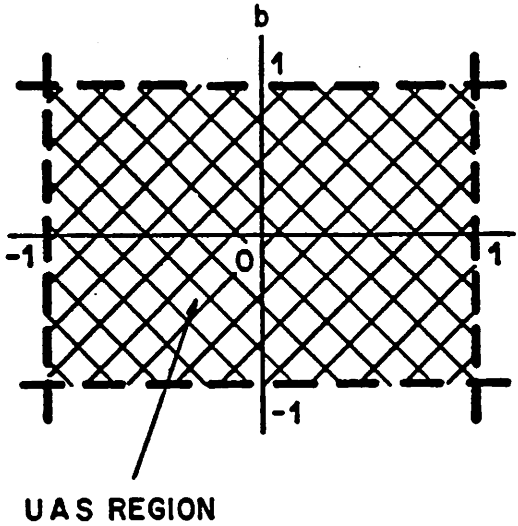

Figure 1. Uniformly Asymptotically Stable Region for the System of Example 4.2.8

4.2.3 Stability of Stochastic Time-Delay Systems

Consider a linear stochastic TD system described by the following differential-difference Ito equation [4.4]:

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{B}\mathbf{x}(t-h)dt + \left[\sum_{j=1}^{n} \mathbf{C}_j d\omega_j\right]\mathbf{x}(t) + \left[\sum_{j=1}^{n} \mathbf{D}_j dw_j\right]\mathbf{x}(t-h) \tag{54}$$

where $\mathbf{x}$ is $n \times 1$ state vector, initial state function $\mathbf{x}_t(\alpha)$, $-h \leqslant \alpha \leqslant 0$ and ω_j is a scalar Wiener process with

$$E\left\{\omega_j(t) - \omega_j(s)\right\}\left\{\omega_j(t) - \omega_j(s)\right\}' = \sigma_{ij} \mid t-s \mid \tag{55}$$

and $\mathbf{x}(0)$ is assumed to be independent of

$$\left\{\omega_j(t) - \omega_j(0)\right\} \text{ for } t \geqslant 0,\ j=1,2,\ldots,n.$$

Kushner [4.11] has considered a scalar case Ito differential difference equation (54) and has given sufficient conditions for its asymptotic stability with probability one. The results of Kushner [4.11] and Kleinman [4.12] are used to establish the asymptotic stability of (54) here.

Consider the following two Lemmas first.

9. Lemma. Let $V(\mathbf{x})$ be a continuous, positive-definite real-valued function on C and $\mathbf{x}(t)$ satisfy (54). Suppose that a bounded open set U exists such that

$$\mathbf{x}_0 = \mathbf{x} \in X \equiv \left\{ \mathbf{x} : V(\mathbf{x}) < r \right\} \cap U \tag{56}$$

and $sup_{t > s \geqslant 0} V(\mathbf{x}_s) < \mathrm{r}$ imply $\mathbf{x}_s \in X$ for all $0 \leqslant s < t$. Let $V(\mathbf{x})$ be an element of the domain of $K(.)$. If $K\,V(\mathbf{x}) = -k(\mathbf{x})$ where $k(\mathbf{x})$ is positive definite in X, then $\lim_{t \to \infty} \mathbf{x}(t) = 0$ with probability one. This lemma is a special case of Theorem 6.2 of Kushner [4.11].

10. Lemma. Let $\mathbf{x}(t)$ satisfy (54), with $\mathbf{B} = \mathbf{D}_j = 0, j = 1, 2, \ldots, n,$ then $\lim_{t \to \infty} \mathbf{x}(t) = 0$ with probability one if and only if for any $\mathbf{R} > 0$, $V(\mathbf{x}) = \mathbf{x}'(\mathrm{t})\mathbf{Q}\mathbf{x}(\mathrm{t})$ is a stochastic Lyapunov function where $\mathbf{Q} > 0$ satisfies

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \sum_{i,j=1}^{n} \mathbf{C}'_i\mathbf{Q}\mathbf{C}_j\sigma_{ij} + \mathbf{R} = \mathbf{0}\,. \tag{57}$$

This lemma for stochastic delay-free systems has been proved in [4.12]. Now consider the following theorem.

11. Theorem. Let A be a stable matrix. Let matrices Q and R satisfy

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \sum_{i,j=1}^{n} \mathbf{C}'_i\mathbf{Q}\mathbf{C}_j\sigma_{ij} + 4\mathbf{R} = \mathbf{0} \tag{58}$$

with additional conditions $\mathbf{R} > 0$ and

$$2\mathbf{R} - \sum_{i,j=1}^{n} \mathbf{D}'_i\mathbf{Q}\mathbf{D}_j\sigma_{ij} - \left[\mathbf{Q}\mathbf{B} + \sum_{i,j=1}^{n} \mathbf{C}'_i\mathbf{Q}\mathbf{D}_j\sigma_{ij} \right]' \times (2\mathbf{R})^{-1} \left[\mathbf{Q}\mathbf{B} + \sum_{i,j=1}^{n} \mathbf{C}'_i\mathbf{Q}\mathbf{D}_j\sigma_{ij} \right] > 0 \tag{59}$$

then $\lim_{t\to\infty}\mathbf{x}(t) = 0$ with probability one.

Proof. Define the following Lyapunov functional,

$$V(\mathbf{x}) = \mathbf{x}'(0)\mathbf{Q}\mathbf{x}(0) + \int_{-\tau}^{0} \mathbf{x}'(s)\mathbf{R}\mathbf{x}(s)\, ds \tag{60}$$

Since $\mathbf{A}$ is stable and $\mathbf{R} > 0$ it follows that $\mathbf{Q} > 0$. Let $\mathbf{x}(t) = \mathbf{x} \in \mathbf{C}$ and choose $r < \infty$.

One can always make U large enough such that condition (56) in Lemma 9 is satisfied. Let $X = \{\mathbf{x} : v(\mathbf{x}) < r\}' \cap U$ and note that $V(\mathbf{x})$ is of class C^2 with respect to $\mathbf{x}(0)$. Then by a corollary of [4.11, Theorem 5.3] $V(\mathbf{x})$ is an element of the domain of $K(.)$ and

$$\begin{aligned} K\,V(\mathbf{x}) = {} & \left[\mathbf{x}'(-h)\mathbf{B}' + \mathbf{x}'(0)\mathbf{A}' \right] \mathbf{Q}\mathbf{x}(0) \\ & + \tfrac{1}{2}\, tr \left\{ \left(\sum_{j=1}^{n} \mathbf{C}_j, \sum_{j=1}^{n} \mathbf{D}_j \right) \left[\mathbf{x}'(0)\ \mathbf{x}'(-h) \right]' \sigma_{ij} \left[\mathbf{x}'(0)\mathbf{x}'(-h) \right] \right. \\ & \times \left. \left(\sum^{n} \mathbf{C}_i, \sum_{i=1}^{n} \mathbf{D}_i \right)' Q \right\} + \mathbf{x}'(0)\mathbf{R}\mathbf{x}(0) - \mathbf{x}'(-h)\mathbf{R}\mathbf{x}(-h) \end{aligned} \tag{61}$$

or

$$\begin{aligned} K\,V(\mathbf{x}) = \tfrac{1}{2} \Bigg\{ & \mathbf{x}'(0) \left(\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \sum_{i,j=1}^{n} \mathbf{C}_i'\mathbf{Q}\mathbf{C}_j\sigma_{ij} \right) \mathbf{x}(0) \\ & + \mathbf{x}'(0) \left(\mathbf{Q}\mathbf{B} + \sum_{i,j=1}^{n} \mathbf{C}_i'\mathbf{Q}\mathbf{D}_j\sigma_{ij} \right) \mathbf{x}(-h) \\ & + \mathbf{x}'(-h) \left(\mathbf{B}'\mathbf{Q} + \sum_{i,j=1}^{n} \mathbf{D}_i'\mathbf{Q}\mathbf{C}_j\sigma_{ij} \right) \mathbf{x}(0) \\ & + \mathbf{x}'(-h) \left(\sum_{i,j=1}^{n} \mathbf{D}_i'\mathbf{Q}\mathbf{D}_j\sigma_{ij} \right) \mathbf{x}(-h) + 2\mathbf{x}'(0)\mathbf{R}\mathbf{x}(0) \\ & - 2\mathbf{x}'(-h)\mathbf{R}\mathbf{x}(-h) \Bigg\} \end{aligned} \tag{62}$$

or

$K\,V(\mathbf{x}) =$

$$- \tfrac{1}{2} \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{x}(-h) \end{bmatrix}' \begin{bmatrix} 2\mathbf{R} & -\mathbf{Q}\mathbf{B} - \sum_{i,j=1}^{n} \mathbf{C}'_i \mathbf{Q} \mathbf{D}_j \sigma_{ij} \\ -\mathbf{B}'\mathbf{Q} - \sum_{i,j=1}^{n} \mathbf{D}'_i \mathbf{Q} \mathbf{C}_j \sigma_{ij} & 2\mathbf{R} - \sum_{i,j=1}^{n} \mathbf{D}'_i \mathbf{Q} \mathbf{D}_j \sigma_{ij} \end{bmatrix} \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{x}(-h) \end{bmatrix} \tag{63}$$

By the hypothesis and Lemma 4.2.5, $K\,V(\mathbf{x}) = -k(\mathbf{x}) < 0$. Therefore by Lemma 9, $\lim_{t\to\infty}\mathbf{x}(t) = 0$ with probability one.

12. Example. Consider the following scalar stochastic TD system

$$dx(t) = -2x(t)\,dt + 1/2\,x(t-1)\,dt + \tfrac{1}{2}\,x(t)\,d\omega + \tfrac{1}{2}\,x(t-1)\,d\omega \tag{64}$$

where $n=1$, $A = -2$, $B=C_1=D_1=1/2$. Let $R=1 > 0$ and assume $\sigma_{11} = 1$. If condition (59) can be shown to hold, then the system (64) would be asymptotically stable. Using values of R, A and C in the Lyapunov equation (58), Q can be determined from

$$-4Q + 1/4\,Q + 4 = 0 \tag{65}$$

which yields $Q = 16/15 = 1.067$. Evaluating the left-hand side of condition (59) one gets

$$2R - D'_1 Q D_1 \sigma_{11} - \left[QB + C'_1 Q D_1 \sigma_{11} \right]' (2R)^{-1} \left[QB + C'_1 Q D_1 \sigma_{11} \right]$$
$$= 2 - 1/4Q - 9/32Q^2 = 2 - 4/15 - 144/450 = 318/225 > 0$$

Therefore, system (64) is asymptotically stable.

4.2.4 Stability of Time-Delay Systems with Variable Delays

One of the well-known characteristics of a time-varying TD system is that it is not necessarily stable (unstable) if its time-invariant version is stable (unstable) *i.e.*, one in which the delay is constant. Bhatt and Hsu [4.13] have shown this characteristic for constant delay TD systems. Although a great deal of stability studies for TD system have been reported [4.4,4.5,4.8,4.13-4.15] and the subject of variable delays has been of great interest to many researches [4.16-4.20] for some time, relatively few studies exist on the effects of the variations of the delay itself on the overall system stability [4.9-

4.21]. The effect of the delay variations on system stability is considered in this section. The thrust of the present discussion is based on the results of Hirai and Satoh [4.6].

Consider the linear autonomous scalar TD system

$$\dot{x}(t) = -ax(t) - bx(t-h(.)) \tag{66}$$

where a and b are real constants, not necessarily positive, and $h(\cdot)$ is the variable delay. The following two examples illustrate the effects of the variations of $h(\cdot)$ with respect to time t and state variable $x(t)$, respectively.

13. Example. Let the time delay be represented by

$$h(t) = t - kT_0,\ kT_0 < t \leqslant (k+1)\ T_0,\ T_0 > 0 \tag{67}$$

for $k=1, 2, \cdots$. This variation is shown in Figure 2. It is clear that $0 \leqslant h(t) \leqslant T_0$.

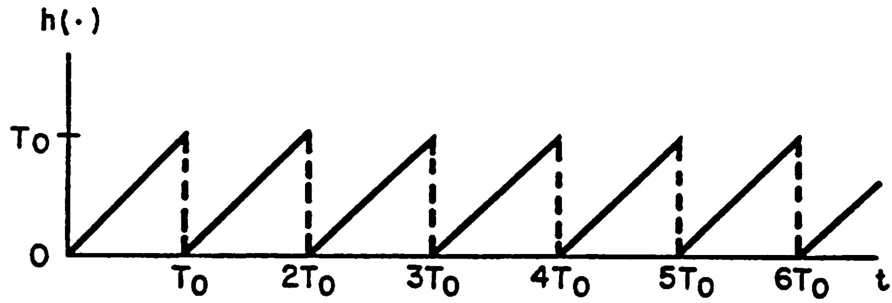

Figure 2. Variation of Delay for Example 4.2.13

Using this definition of delay (66) can be rewritten as

$$\dot{x}(t) = -ax(t) - b\,x(kT_0),\ \ kT_0 < t \leqslant (k+1)\ T_0\,. \tag{68}$$

The solution of this equation at $t = (k+1)\ T_0$ satisfies

$$x\left[(k+1)\ T_0 \right] = \left\{ (1+b/a)\ e^{-aT_0} - b/a \right\}^k x(kT_0),\ for\ a \neq 0 \tag{69a}$$

$$= (1-bT_0)\ x(kT_0),\ for\ a = 0 \tag{69b}$$

The necessary and sufficient condition for asymptotic stability of this system, *i.e.*, $\lim_{k\to\infty} x\ ((k+1)\ T_0) = 0$, is

$$\left| (1+b/a)\ e^{-aT_0} - b/a \right| < 1,\ a \neq 0 \tag{70a}$$

$$\left| 1 - bT_0 \right| < 1, \qquad a = 0 \tag{70b}$$

For $T_0 = 1$, the stability region (70) in the $b - a$ plane is shown in Figure 3.

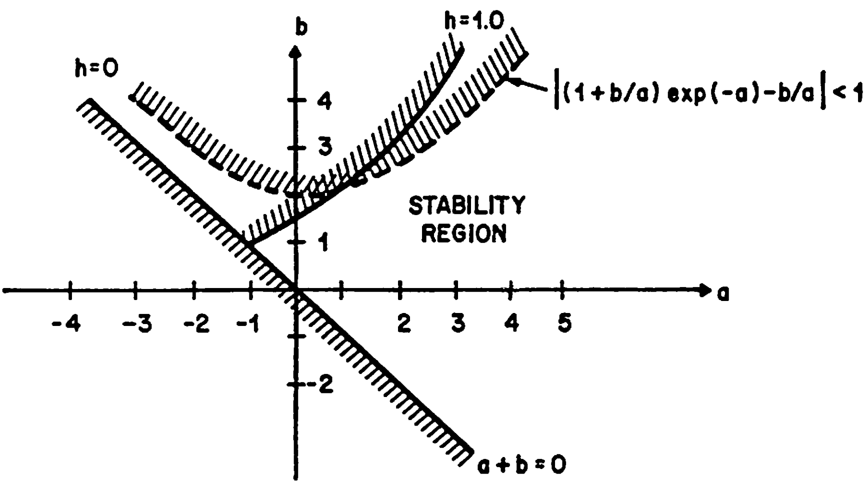

Figure 3. Stability Region for Variable-Delay Example 4.2.13.

It is clear from this figure that the system with time-varying delay $0 \leqslant h(t) \leqslant 1$ is not always stable even if the system with constant time delay which satisfies $0 \leqslant h \leqslant 1$ is stable.

14. Example. Consider the stability of system (66) for the delay expressed by

$$h\ (x) = a + b\ sgn(x) \tag{71}$$

which depends on the state variable x. One industrial system where the delay is a function of the state is a multi-stand cold rolling mill [4.22] whose near-optimum control is discussed in Section 9.2. In (71), a and b are constant parameters. Assume that the solution of (66) with delay (71) is of the form

$$x(t) = A(t)\,(\sin(\omega t + \phi) + u_0) \tag{72}$$

where u_0 is constant, ϕ is a phase angle and $A(t)$ is assumed to vary slowly compared with ω.

Substituting (71) and (72) into (66) and making the approximation $A(t-h(x)) \approx A(t)$, one has

$$\begin{aligned} A(t)u_0\,(a+b) + \dot{A}(t)\,u_0 + (\dot{A}(t) + a\,A(t) + b\,A(t)\cos\omega h)\sin\psi \\ + (A(t)\omega - bA(t)\sin\omega h)\cos\psi = 0 \end{aligned} \tag{73}$$

where $\psi = \omega t + \phi$. Expanding the left-hand side of (73) in Fourier series and neglecting the higher harmonics, the following relations result:

$$\dot{A}(t) = -\left\{ (a+b)\,\pi\,u_0 - \mathrm{b}\cos\theta\,(\cos\omega h_2 - \cos\omega h_1) \right\} A(t)/\pi u_0 \tag{74a}$$

$$\begin{aligned} \dot{A}(t) = -\Big\{ & 2\,a\,\pi + b\pi(\cos\omega h_2 + \cos\omega h_1) \\ & - 2b\,(\theta - \cos\theta)\,(\cos\omega h_2 - \cos\omega h_1) \Big\} A(t)/2 \end{aligned} \tag{74b}$$

$$\begin{aligned} 0 = \pi\,\omega + b\Big\{ & \Big[\sin\omega h_2 - \sin\omega h_1 \Big] (\theta + \cos\theta) \\ & - \pi \Big[(\sin\omega h_1 + \sin\omega h_2)/2 \Big] \Big\} \end{aligned} \tag{74c}$$

where $\theta = \sin^{-1}(u_0)$, $0 \leqslant \theta \leqslant \pi/2$, $h_1 = a + b$ and $h_2 = a - b$. Since at the stability boundary $\dot{A}(t) = 0$, the desired region of stability can be obtained by setting the left-hand sides of (74a) and (74b) equal to zero and solving (74) for the three unknowns a, b and u_0 while keeping ω as a parameter. Figure 4 shows the stability region for $a = 1.5$ and $b = 1.0$.

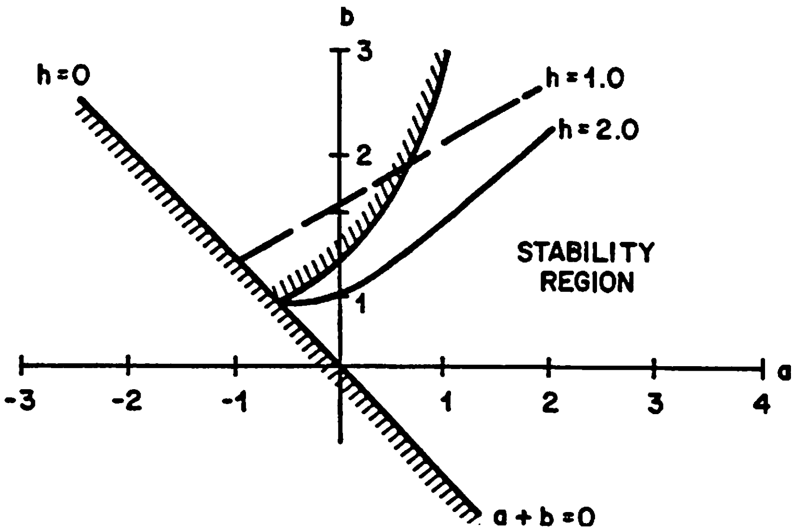

Figure 4. Stability Region for the TD System of Example 4.2.14

The boundaries for constant $h = 0$, 1.0 and 2.0 are also shown. Since the system's stability region with variable time delay $h(t) \leqslant 2$ contains the constant time-delay ($h \leqslant 2$) case's stability region, it is enough to investigate system's stability by replacing $h(x)$ with its maximum values.

It is noted that in Example 14 due to the variation of delay $h(x)$, the system is actually nonlinear and one should investigate the effects of its initial function on the stability. Let the solution of (66) with delay in (71) be represented by $x(t, \eta(\sigma))$, where $\eta(\sigma)$, $-h_{\max} \leqslant \sigma \leqslant 0$ is the initial function for $x(t)$ and $h_{\max}$ is maximum of $h(t)$. Noting that $x(t, \eta_1(\sigma)) = f\, x(t, \eta_2(\sigma))$, $f > 0$ where $\eta_1(\sigma) = f\, \eta_2(\sigma)$ *i.e.*, a linear relationship between the solution of $x(t)$ and the initial function has been established. Therefore, the initial function has no influence on the stability of the scalar TD system (66).

4.3 STABILITY OF TIME-DELAY SYSTEMS VIA FREQUENCY DOMAIN

Perhaps the largest effort in the stability of TD system has been made in frequency domain which is applicable by and large to SISO systems. Consider an nth order m-delay differential-difference equation

$$\sum_{k=0}^{n} \sum_{\ell=0}^{m} a_{k\ell}\, x^{(k)}(t-h_\ell) = u(t) \tag{1}$$

where $x^{(i)}$ represents ith derivative, h_ℓ is the ℓth time-delay, $u(t)$ is the system's single input and $a_{k\ell}$ are constant coefficients. It is, moreover, easy to note that the stability of systems described by (1), as also seen in Section 4.2, is somewhat more difficult.

From the mathematical point of view, stability of linear differential-difference equations has been completely treated. Razumikhin [4.23] and Krasovskii [4.24] applied Lyapunov's Second method to these systems, while Bellman and Cooke [4.1] used Laplace transformation to develop rigorous theorems for stability. These theorems, although mathematically exact, are very difficult to apply to check stability. Some easier criteria for time-domain representation of TD system were presented in the previous section.

The frequency-domain methods, discussed in this section, are geared around the roots of the following characteristics equation

$$R\ (s, e^{-s}) = \sum_{k=0}^{n} \sum_{\ell=0}^{m} a_{k\ell}\, s^k\, e^{-h_\ell} = 0 \tag{2}$$

which need to be in the *l.h.p.* for the system stability. There are three primary categories of frequency-domain techniques used to check the stability of TD system. These are (i) *analytical*, (ii) *graphical* and (iii) *grapho-analytical* schemes. Since (2) is not an algebraic equation, standard analytic approaches such as Routh-Hurwitz criterion are not appropriate here. Various analytical criteria, most of which are based on the "principle of argument" [4.25], and deduced from the well known residue theorem of Cauchy have been proposed. Sherman [4.26], Hayes [4.27] and Wright [4.28] have suggested analytical criteria for special TD system. Perhaps the most general analytic method is based on the works of Pontryagin [4.29,4.30] and Pyatniskii [4.31,4.32] which give criteria to check the roots of the characteristic equation (2). The approach of Pontryagin [4.29,4.30], as presented by Ho and Chan [4.25], is considered in Section 4.3.1. Graphical schemes, utilizing classical methods such as Nyquist, Bode and root locus, are discussed in Section 4.3.2. A typical grapho-analytical scheme, known as "generalized phase comparison" [4.33,4,34] is presented in Section 4.3.3.

4.3.1 Pontryagin's Stability Method

Pontryagin [4.29,4.30] has proved a number of theorems on the location and types of the roots of transcendental equations which can be exploited to provide relatively more convenient criteria for the stability of TD systems. Assume that all the delays h_ℓ, $\ell=0,\ldots,m$ are real and positive integers. Let us rewrite (2) such that all the exponents are positive, *i.e.*, multiplying both sides by a sufficiently high power of e^s:

$$P\,(s,e^s) = \sum_{k=0}^{n}\sum_{\ell=0}^{m} b_{k\ell}\, s^k\, e^{h_\ell} = 0\,. \tag{3}$$

Let N and M denote the highest powers of s and e^s. respectively; then the term $b_{NM}s^N e^{h_M}$ in (3) is called the "principle term" of the transcendental polynomial. It should be noted that not all polynomials have a principle term. With the above points made, let us now consider the first of three analytic criteria due to Pontryagin [4.29,4.30].

1. Criterion. If the characteristic equation $P\,(s,e^s) = 0$ of a TD system has no principal term, then the system is unstable for all $b_{k\ell}$, $k=0,1,\ldots,n$ and $\ell=0,1,\ldots,m$.

Clearly, the existence of the principle term is only a necessary condition for stability and not a sufficient one. In order to develop the stability criterion, let $s=j\omega$ in (3) to obtain

$$\begin{aligned} P\,(j\omega,e^{j\omega}) &= E(\omega) + jF(\omega) \\ &= E(\omega,\, sin\omega,\, cos\omega) + jF(\omega,\, sin\omega,\, cos\omega)\,. \end{aligned} \tag{4}$$

Now consider the second criterion given below.

2. Criterion. A TD system with a principal term is stable if and only if one of the following three conditions is satisfied:

a) All the roots of $E(\omega)=0$ and $F(\omega)=0$ are real and simple, and the inequality

$$E(\omega)\,F'(\omega) - E'(\omega)\,F(\omega) > 0 \tag{5}$$

is satisfied for at least one value of ω.

b) All the roots of $E(\omega)=0$ are real, and for each root $\omega=\omega_e$, the inequality

$$E'(\omega_e)\, F(\omega_e) < 0 \tag{6}$$

is satisfied.

c) All the roots of $F(\omega)=0$ are real, and, for each root $\omega=\omega_f$, the inequality

$$E(\omega_f)\, F'(\omega_f) > 0 \tag{7}$$

is satisfied. In above relations the prime denotes differentiation *w.r.t.* ω.

In order to exploit the above criterion, another result of Pontryagin [4.30, Theorem 3] can be utilized to find the number of real roots of $E(.)=0$ and $F(.)=0$. This constitutes the third criterion given below.

3. Criterion. For the equations $E\ (\omega,\ sin\omega,\ cos\omega)=0$ and $F\ (\omega,\ sin\omega,\ cos\omega)=0$ to have only real roots, it is necessary and sufficient that, in the interval

$$-2k\pi + \eta \leqslant \omega \leqslant 2k\pi + \eta \tag{8}$$

it has exactly $4kN + M$ real roots starting with a sufficiently large k; η is an appropriate constant such that the coefficient of terms of highest degree in ω does not vanish at $\omega=\eta$. The following two examples illustrate these stability criteria.

4. Example. Consider a first-order system with a time-delay feedback:

$$G(s) = ke^{-hs}/(s+a) \tag{9}$$

where h is the delay (transport lag). The system's characteristic equation is

$$s + a + ke^{-hs} = 0 \tag{10a}$$

or

$$(s+a)\, e^{hs} + k = 0 \tag{10b}$$

If the delay h is not an integer, the equation can be normalized by setting $s' = hs$, leading to,

$$(s' + a') e^{s'} + k' = 0 \tag{11}$$

where $a' = ha$ and $k' = hk$. One can, henceforth, omit the primes while remembering that the equation has been normalized. The normalized transcendental equation is, then,

$$P(s, e^s) = (s+a) e^s + k = 0 \tag{12}$$

It is noted that $M=N=1$ and $P(s, e^s)$ has a principle term indicating that the necessary condition for stability is satisfied through Criterion 1. In order to determine the sufficient condition for stability, let $s=j\omega$ in (12) to obtain

$$P(j\omega, e^{j\omega}) = (j\omega+a) e^{j\omega} + k \tag{13a}$$

$$E(\omega) = a \cos\omega - \omega \sin\omega + k \tag{13b}$$

$$F(\omega) = \omega \cos\omega + a \sin\omega \tag{13c}$$

By Criterion 2 all roots of either $E(\omega)$ or $F(\omega)$ must be real and simple for the system to be stable. In this example it is more convenient to study the roots of $F(\omega)$ instead of $E(\omega)$. Thus, one needs to find the roots of

$$F(\omega) = \omega \cos\omega + \mathrm{a} \sin\omega = 0 \tag{14a}$$

or

$$\tan\omega = -\omega/\mathrm{a} \tag{14b}$$

Let us consider first the case when $a=0$. Under this situation one must have $\omega \cos\omega=0$ indicating that one has either $\omega=0$ or $\omega=(2m+1)\ \pi/2$, $m=0, \pm 1, \cdots$. For $a=0$ and $\omega=0$, as clear from (13b), it leads to $k=0$ which is not physically meaningful. For $\omega = (2m + 1)\ \pi/2$, it is clear that the term $F(\omega) = 0$ would have $4k$ roots in the range $-2k\pi \leqslant \omega \leqslant 2k\pi$ contradicting Criterion 3 which requires $4kN + M = 4k + 1$ roots for the zeroes of $F(\omega) = 0$ to be all real. Thus one can proceed with the two cases where $a > 0$ or $a < 0$. Under such condition one can plot both sides of (14b) as functions of ω and examine the points of intersection of the two. In Figure 1 the plots of $\tan\omega$ and $-\omega/a$ are shown for both positive and negative values of a. Since both $\tan\omega$ and $-\omega/a$ are odd functions it suffices to show the plots for $0 < \omega \leqslant 2k\pi$. Consider the range $-2\pi \leqslant \omega \leqslant 2\pi$. It is required to have $4kN + M = 4 + 1 = 5$ roots (Criterion 3), or in the range $0 \leqslant \omega \leqslant 2\pi$ we need 3 roots. From Figure 1 it is clear that for $a < 0$ there are only two roots within the range of $0 \leqslant \omega \leqslant 2\pi$. However, for $a > 0$ there are three roots which would satisfy the first part of Criterion 2(c). In order to consider the second part of this criterion condition (7) must also be satisfied. It is noted that

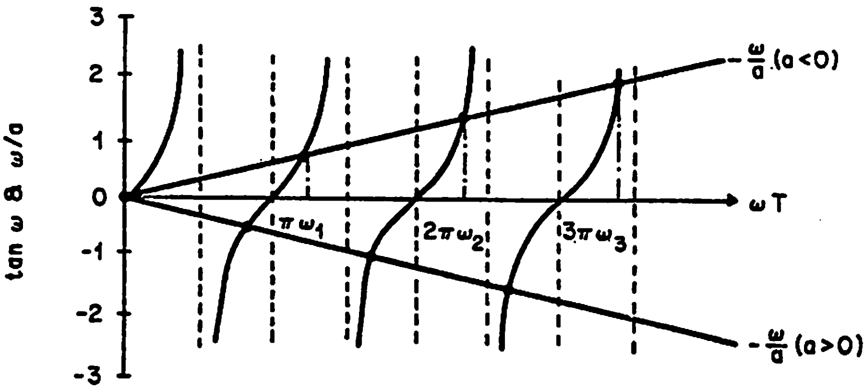

Figure 1. Solution of Equation $F(\omega) = 0$ for Example 4.3.4 with $a \neq 0$

$$F'(\omega) = \cos\omega - \omega \sin\omega + a \cos\omega \tag{15a}$$

which can be rewritten, after considering (14b), as

$$F'(\omega) = (\frac{1}{a} \cos\omega)\ (\omega^2 + a^2 + a) \tag{15b}$$

The above expression will be positive only if $\cos\omega > 0$, since $a > 0$ and ω is real. Similarly, by utilizing (14b) the term $E(\omega)$ in (13b) can be reduced to

$$E(\omega) = (\frac{1}{a} \cos\omega)\ (a^2 + \omega^2) + k \tag{16a}$$

leading to the fact that the sign of $E(\omega)\ F'(\omega)$ is that of

$$H(\omega) = (\frac{1}{a} \cos\omega)^2\ (a^2 + \omega^2) + \frac{k}{a} \cos\omega. \tag{16b}$$

Using (14b) and some trigonometric identities, (16b) reduces to

$$H(\omega) = 1 + \frac{k}{a} \cos\omega. \tag{16c}$$

Let (as shown in Figure 1) the roots of $F(\omega) = 0$ be denoted by $\omega_1 < \omega_2 < \cdots < \omega_i \cdots$ and $\omega_{-i} = -\omega_i$ for $i=1,2,\cdots$. The remaining problem is to check the sign of $H(\omega)$ for all ω_i, $i=0, \pm 1, \cdots$. Noting that $\cos\omega_i$ is positive for odd i and $\omega_i > 0$, it remains to see under what conditions on k and a the term $H(\omega)$ is positive for $\omega = \omega_i$, i being even. Solving for $\cos\omega$ from (14b) on has $\cos\omega = \pm a (\omega^2 + a^2)^{-1/2}$ and after keeping the negative value of $\cos\omega$, $H(\omega)$ becomes

$$H(\omega) = 1 - k (\omega^2 + a^2)^{-1/2} , \quad \omega = \omega_i, \; i \text{ even}. \tag{17}$$

From this expression it is clear that $0 < H(\omega) < 1$ for all $-\infty \leqslant \omega \leqslant \infty$ and $k > 0$. Thus, the system (9) is stable as long as $a > 0$ and $k > 0$, *i.e.*, the first quadrant of the $k - a$ plane. Δ

The above example illustrates the fact that investigating the stability of even a first-order TD system by analytic methods such as Pontryagin's transcendental polynomial theorems [4.29,4.30] is a lengthy proposition. In fact as the order of the system and/or the number of plant parameters, such as k and a, increase the analytical work gets more and more complicated. In [4.25], Ho and Chan have presented the stability analysis for the system $G(s) = ke^{-hs}/(s(s^2 + as + b))$ with no prior restrictions on plant parameters k, a and b.

A number of authors have examined the stability of linear TD system by incorporating a Smith Controller [4.33] which utilizes a model of the plant [4.34-4.41]. The simulation of a Smith Controller on analog and/or digital computers has received attention only in recent years due to its fast response characteristics. However, the analytic treatment of stability of TD systems through a Smith Controller remains too involved for use in any practical TD system.

In the remaining part of this section an analytic criterion for a class of linear TD system, due to Thowsen [4.42] will be discussed.

Let us assume that the characteristic equation (2) can be rewritten in the form

$$R(s,\tau) = \sum_{i=0}^{m} \rho_i(s)\, e^{-i\tau s} \tag{18}$$

where the delay $h_i \triangleq i\tau$, for a constant $\tau > 0$ and $\rho_i(s)$ is a polynomial in the complex variable s of order less than or equal to n, the system's order. The asymptotic stability of (18) can be determined from the roots of the following nontranscendental characteristic equation:

$$\sum_{i=0}^{m} \rho_i(s)\ (1 - Ts)^{2i}\ (1 + Ts)^{2m-2i} = 0 \tag{19}$$

where T is a nonnegative number and the polynomial on the left side of (19) is of order $2m+n$. The relation between imaginary roots of (18) and (19) can be established by the following theorem.

5. Theorem. The value $s = j\omega$ ($\omega \geqslant 0$) is an imaginary root of (18) for some $h \geqslant 0$ if and only if $s = j\omega$ is also a root of (19) for some nonnegative number T.

Proof. Let $s = j\omega$ be an imaginary root of (18) for some $h = h_o \geqslant 0$, *i.e.*,

$$\sum_{i=0}^{m} \rho_i(j\omega) e^{-j\omega h d}\ 0 \tag{20a}$$

The term $exp(-j\omega h_o)$ represents a point on the unit circle which can be described by $[(1-j\omega T)\ /\ (1+j\omega T)]^2$ from some $T \epsilon [0,\infty)$. Taking this into account, (20a) can be rewritten as

$$\sum_{i=0}^{m} \rho_i(j\omega)\ (1-j\omega T)^{2i}\ /\ ((1 + j\omega T))^{2i} = 0 \tag{20b}$$

Now if (20b) is multiplied by $(1 + j\omega T)^{2m}$, one obtains

$$\sum_{i=0}^{m} \rho_i(j\omega)\ (1 - j\omega T)^{2i}\ (1 + j\omega T)^{2m-2i} = 0 \tag{20c}$$

which corresponds to (19). The converse can be easily proved. Consider the following lemma.

6. Lemma. If $s = j\omega$ is an imaginary root of (18) for $h = h_0$, then $s = j\omega$ is also an imaginary root of (18) for $h = h_0 + 2\pi\ell/\omega$> for $\ell = 0, \pm\ 1, \dots$.

The proof follows directly from $R(j\omega,h) = R(j\omega,h + 2\pi/\omega)$. The following example illustrates how this theorem can be utilized to check the stability of a TD system.

7. Example. Consider a second-order TD system considered elsewhere by Goesiwski and Olbrot [4.43] and Thowsen [4.42]

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x}(t-\tau) \tag{21a}$$

The characteristic polynomial of this system corresponding to (18) is

$$s^2 + s + se^{-s\tau} + 1 = 0 \tag{21b}$$

when $\tau = 0$, the above equation leads to $s^2 + 2s + 1 = (s + 1)^2 = 0$ and the auxiliary equation (19) will become

$$T^2s^4 + 2T(T + 1)\, s^3 + (T^2 + 1)\, s^2 + 2\,(T + 1)\, s + 1 = 0 \tag{21c}$$

Using standard Routh-Hurwitz stability test for this polynomial reveals that for $T = 1$ the system has a pair of imaginary roots at $s = \pm j$. Since $T = 1$ and $\omega = 1$, it follows from Lemma 6 that $\tau = \pi + 2\pi\ell$, ℓ=0, $\pm$ 1, $\cdots$. Hence, system (12a) is asymptotically stable for all $\tau\epsilon[0,\pi)$. This conclusion, as also mentioned by Thowsen [4.42], modifies the result of Gosiewski and Olbrot [4.43], which concluded that the system is stable for $0< \tau < 0.085$.

Before we leave the discussions on analytic frequency-domain methods, it is worth mentioning another scheme due to Brierley *et. al.* [4.44] who have presented a necessary and sufficient condition for asymptotic stability "independent of delay" for the system

$$\dot{\mathbf{x}}(t) = \sum_{\ell=0}^{r} \mathbf{A}_\ell\, \mathbf{x}(t-\ell\tau) \tag{22}$$

where $\mathbf{A}_\ell$ are $n\times n$ matrices and $\tau > 0$ is the delay. Their main result, which was also suggested by Kamen [4.45], is that the necessary and sufficient condition for asymptotic stability independent of delay is that a complex matrix Lyapunov equation

$$\mathbf{A}^*(z)\; \mathbf{P}(z) + \mathbf{P}(z)\; \mathbf{A}(z) + \mathbf{Q}(z) = 0 \tag{23}$$

has a positive-definite hermitian solution $\mathbf{P}(z)$ where $\mathbf{Q}(z)$ is also a positive-definite hermitian matrix. In (23), $z = exp(j\omega)$ and complex matrix $A(z)$ is given by:

$$A(e^{-hs}) = \sum_{\ell=0}^{r} A_\ell \, e^{-\ell \tau s} \, . \tag{24}$$

4.3.2 Graphical Methods

The main graphical schemes for the stability of TD system are the classical approaches such as the Nyquist criterion, Bode diagrams and the Root Locus method. The applications of these methods to TD system are also rather complicated even for the simplest delay-free plant and a single constant delay [4.41]. In this section two graphical approaches involving Nyquist and Bode diagrams will be considered which may be convenient for checking the stability of SISO TD system.

4.3.2.1 Modified Nyquist Approach

The method described here is based on the modified Nyquist scheme of Satche' [4.46] which has also been treated by Marshall [4.41]. Consider a delay-free system with transfer function *G(s)* in cascade with a transport delay and a negative unity feedback. The characteristic equation of this system is

$$1 + G(s)e^{-hs} = 0 \tag{25}$$

whose roots must all lie in the *l.h.p.* for stability. In terms of the Nyquist stability criterion the polar plot of $G(s)e^{-hs}$ on the Ge^{-hs} -plane must encircle the point $(-1, j0)$ in the counter-clockwise direction as many times as the number of open-loop poles $G(s)e^{-hs}$. In most cases, there are no open-loop poles for the system which indicates that the corresponding polar plot should not encircle $(-1, j0)$ point. The polar plot represents a mapping of the $+j\omega$ axis onto the $G(s)e^{-hs}$ plane. However, a complete Nyquist diagram includes the mapping of the entire right-half s-plane, which is encompassed by a closed path known as the *Nyquist Path*. It is this mapping process from the s-plane to the $G(s)e^{-hs}$ plane which would make the application of the original Nyquist Stability Criterion very difficult for TD systems.

The main outcome of Stache''s method [4.46] is to rewrite the characteristic equation as

$$[G(s)]^{-1} = -e^{-sh} \tag{26}$$

and sketch the Nyquist plots for $[G(s)]^{-1}$ and $-exp(-sh)$ and check whether the mappings of these two terms overlap. If the two polar plots do not intersect, there is no purely imaginary roots $\pm j\omega_1$ at any value of $\omega = \omega_1$. It is noted that polar plot of the delay term $exp(-j\omega h) = cos\omega h - j\ sin\omega h$ is a unit circle regardless of the value of h. Similarly, the polar plot of $K\ exp(-sh)$ is also a circle with radius K. The following example illustrates the graphical scheme of Satche′ [4.46].

8. Example. Consider a simple first-order TD system with characteristic equation

$$1 + G(s)\,e^{-sh} = 1 + Ke^{-sh}/(1 + h_p s) = 0 \tag{27}$$

it is desired to find the region in the $K - h_p$ plane for which the system is stable. Following the reformulation of TD systems (27), one has

$$(1 + j\omega h_p) = -Ke^{-j\omega h} \tag{28}$$

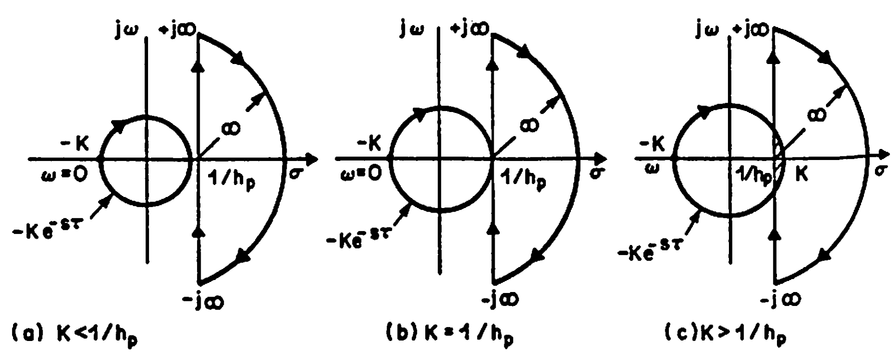

Figure 2. Three Possible Stability Conditions for the TD System of Example 4.3.8

for the two sides of the Nyquist diagram. The Nyquist diagram of the left-hand side of (28) is a straight line passing through the point $(1/h_p, 0)$ while the diagram of the right-half side is a circle with radius K and center at (0,0). Figure 2 shows three possible situations for two parameters K and h_p. For the case $K < 1/h_p$ the two regions, *i.e.*, the circle and the D-shaped infinite contour, do not overlap. Thus there is no root of the characteristic equation (27) in the *r.h.p.* This means that the system is stable for any value of delay h. As the value of gain K is increased, the delay's circle may have a tangent point with the D-contour, or it may overlap the contour for $K > 1/h_p$ as shown in Figure 2(c). These two cases (*b* and *c* in Figure 2) may correspond to the oscillatory and unstable situations. The desired region of stability in the $K-h_p$ plane is shown in Figure 3. It is interesting to note that should $1/h_p < 0$ and $K < |1/h_p|$ then the two Nyquist diagrams would be in such a way that the infinite D-contour would be completely overlapping the circle as shown in Figure 4. Here again there are roots of the characteristic equation which lie in the *r.h.p.*, *i.e.*, unstable system. This result is of course expected since the open-loop pole of the delay-free plants is already in the *r.h.p.* and the system is inherently unstable. This situation is, however, unusual and is only useful for the unintersecting cases.

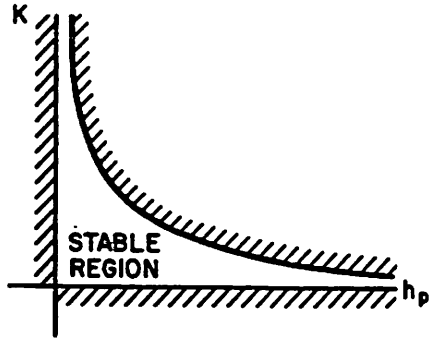

Figure 3. Stability Region for Example 4.3.8

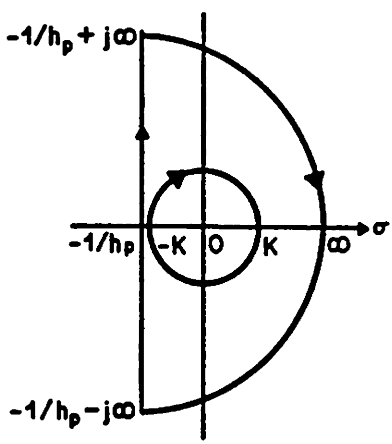

Figure 4. A Complete Overlap Situation for $h_p < 0$ in Example 4.3.8

It is noted that in the above example, the stability of the TD system considered did not depend on the delay itself. In sequel, another example, due to Marshall [4.41], is used to illustrate the effects of delay h on the stability and to show that under restricted choice of h, the TD system can still be stable even though the two contours (circle and D-type) intersect.

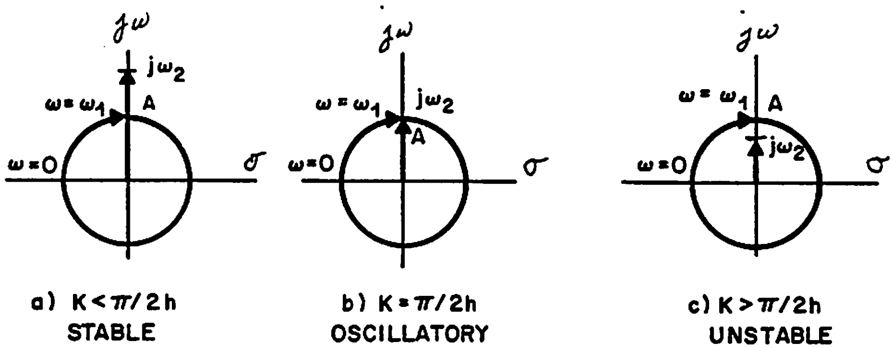

Figure 5. Stability Conditions for the TD System of Example 4.3.9

9. Example. Consider a TD system with the overall open-loop transfer function Ke^{-sh}/s. The two Nyquist diagrams are obtained by polar plots of $Kexp\ (-j\omega h)$ and $j\omega$. Figure 5a shows the two plots which intersect each other at point A. The values of frequencies on the circle and the $j\omega$-plot are $\omega_1 = \pi/2h$ and $\omega_2 = K$, respectively. If $K < \pi/2h$, it would mean that $\omega_2 < \omega_1$ or the circle plot would intersect the $j\omega$-axis after the $j\omega$-plot has passed through the circle. This situation corresponds to a stable system if one checked the stability of Pontryagin's [4.30] method discussed in Section 4.3.1. If $K = \pi/2h$ then the two frequencies are the same as the two plots meet at point A (see Figure 5(b)) and the characteristic equation would have a root on the $j\omega$-axis. Finally, for $K > \pi/2h$ the circular locus would reach point A before the $j\omega$-plot, *i.e.*, $\omega_2 > \omega_1$ and the system would become unstable.

4.3.2.2 A Frequency Response Approach

In this Section the stability of linear TD systems under the variations of delay and open-loop gain by superimposing the polar (Nyquist) plot on the Bode diagram is considered. The essence of our discussion here is based on the result of N-Nagy and Al-Tikriti [4.47].

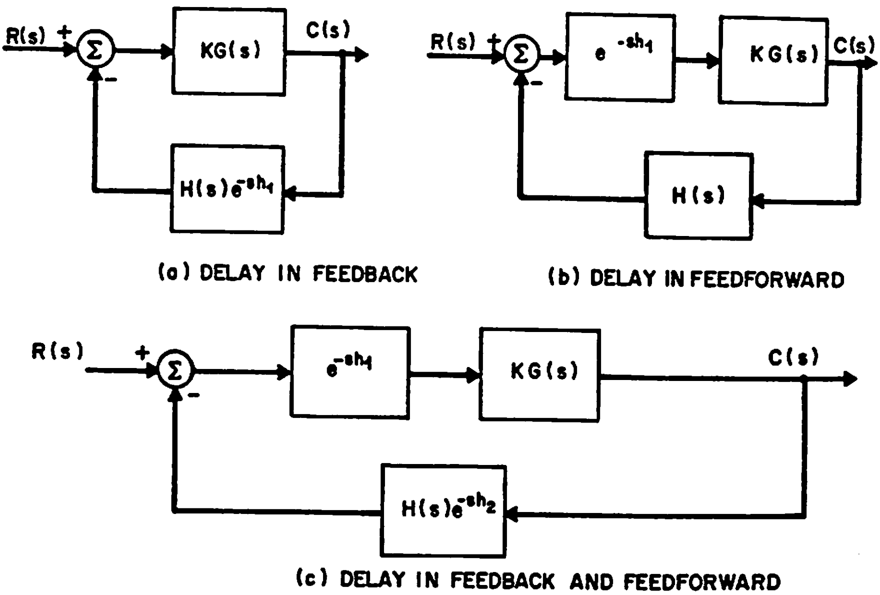

Figure 6. The Class of TD Systems Considered in Section 4.3.2.2

Consider the block diagram of the class of three types of delays in linear TD systems shown in Figure 6. As seen, the delay can be in either feedback or feed-forward or in both. The stability of these TD systems can be affected by both the delay and the gain of the delay-free plant. Consider the closed-loop transfer function of the most general of the three cases of Figure 6:

$$\frac{C(s)}{R(s)} = \frac{Ke^{-sh_1}G(s)}{1+KG(s)\,H(s)\,e^{-s(h_1+h_2)}} \tag{29}$$

Let the phase and magnitude of the open-loop transfer function $KG(s)H(s)$ be denoted by $\phi(\omega)$ in degrees and $M(\omega)$ in dbs, respectively. Moreover, let the overall time delay be $h = h_1 + h_2$. The open-loop transfer function in frequency domain can be represented by

$$KG(j\omega)H(j\omega)e^{-j\omega h} = M(\omega) \measuredangle\ \phi(\omega) - \omega h(180^\circ/\pi) \tag{30}$$

The Bode diagram for this system is shown in Figure 7. Shown is a straight line representing $-180^\circ + \omega h(180^\circ/\pi)$. Let the point at which $\phi(\omega)$ and $\omega h(180^\circ/\pi) - 180^\circ$ intersect be represented by C at frequency $\omega = \omega_1$. At this point the phase angle of the TD system, defined by (30), is given by

$$\phi(\omega_1) - \omega_1 h(180^\circ/\pi) = -180^\circ \ . \tag{31}$$

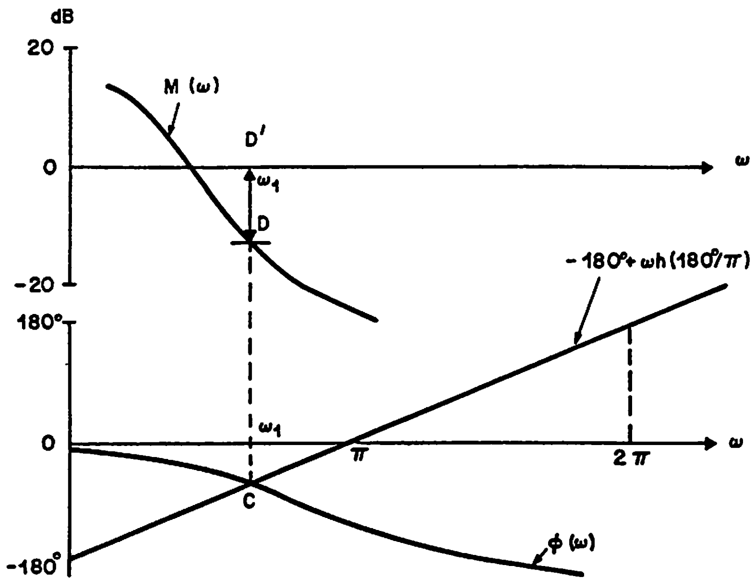

Figure 7. Bode Diagram for the TD System of Figure 6c

In other words, at the point of intersection of phase plot $\phi(\omega)$ of delay-free open-loop transfer function $KG(j\omega)\ H(j\omega)$ and the straight line segment $-180^\circ + \omega h(180^\circ/\pi)$ the phase of TD system is -180°. Then, the stability criterion, within the context of Nyquist, is simply that the TD system is stable if $M(\omega)$ at frequency of intersection point C, *i.e.*, $\omega = \omega_1$ is less than 0 db, see line segment DD' in Figure 7. Clearly, if the delay h is fixed, but the gain K is increased the magnitude plot $M(\omega)$ moves upward causing eventual system instability. On the other hand, if K is fixed and the delay h is increased, the line $-180^\circ + \omega h(180^\circ/\pi)$ would become steeper causing the point of intersection C to be shifted to the left, hence causing instability also. Figure 8 shows a family of line segments $-180^\circ + \omega h(180^\circ/\pi)$ for various values of h. This family of lines is called "Time Delay Chart" by N-Nagy and Al-Tikriti [4.47]. By superimposing the delay-free transfer function's magnitude and phase angle plots on time delay charts one, as it is demonstrated below, can find the desired delay h or gain K or both for system stability.

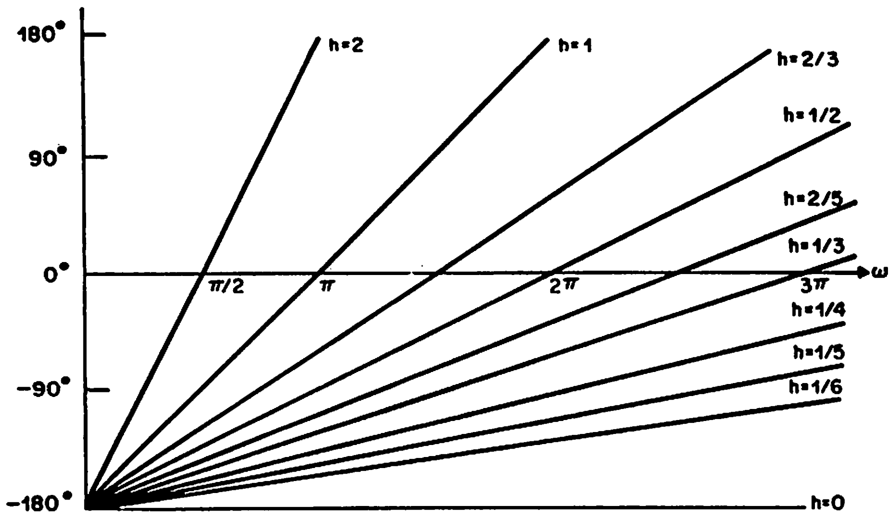

Figure 8. Time Delay Charts for Linear TD Systems

10. Example. Consider a three-dimensional SISO TD system described by

$$KG(s)H(s)e^{-sh} = Ke^{-sh}/\left[s(s+2)(s+5)\right] . \tag{32}$$

In this example it is desired to make the following investigations: (i) Find the maximum time delay for $K=10$ and system still remaining stable, (ii) For $h = 0.1$ what is the maximum possible value of K? and (iii) If the gain K is reduced by a factor of 5, what is the maximum time delay which still preserves stability? This problem can be solved rather easily by using the "Time-Delay Charts" of Figure 8. The Bode diagram of delay-free part of system (32) superimposed on the time-delay charts is shown in Figure 9. For $K = 10$ the magnitude is 0 db at $\omega=1$ (point B) which after referring the point down to reach the phase diagram $\phi(\omega)$ at A, the corresponding maximum allowable delay is $h = 0.4$ as shown. In order to encounter the second question, the $h = 0.1$ line is considered. Its intersection with $\phi(\omega)$ is point C at $\omega = 2.35$ *rad/sec*. The corresponding point D on $M(\omega)$ at $K = 10$ indicates that the magnitude plot is -9.45 dbs below the 0 db line. This would indicate that the gain K can be increased to $2.97 \times 10 = 29.7$ without destabilizing the system. The final question can be answered similar to the first one. Draw the magnitude plot corresponding to $K = 2$ and bring down its 0 db intersection (point F) to reach point E at -107.23 degrees. The corresponding value of delay is $h_{max} = 4.4$. Δ

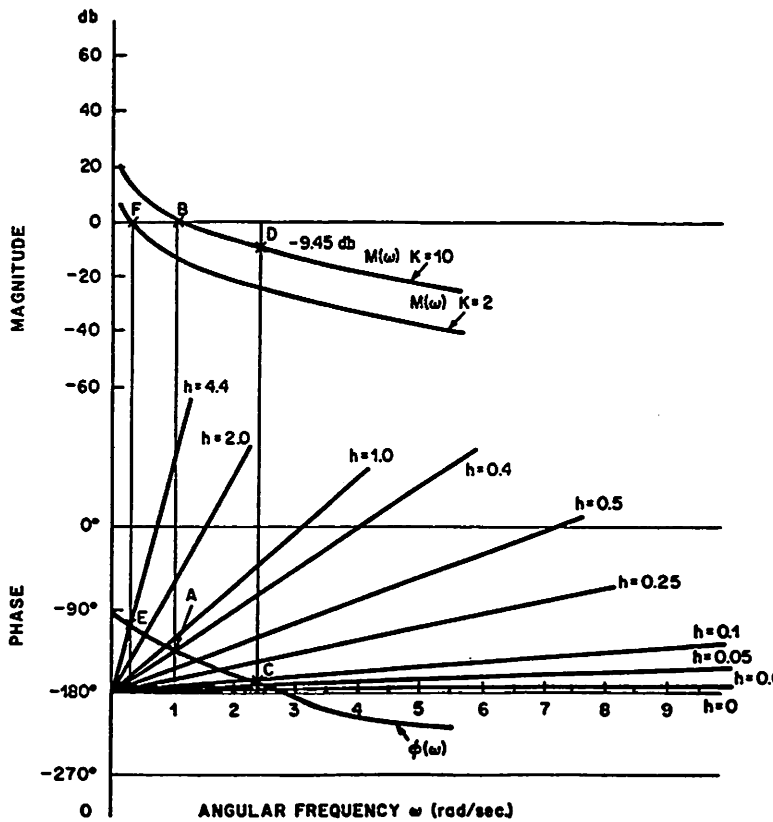

Figure 9. Frequency-Domain Solution of Example 4.3.10

There are many other graphical techniques which have been applied to check the stability of TD systems [4.48 - 4.59]. The notable schemes are the root locus [4.49, 4.58 - 4.59] and the D-partition [4.52, 4.53] schemes are applicable to linear SISO TD-systems. Certain other techniques such as Popov [4.49] and phase plane [4.50, 4.51, 4.56] are for nonlinear TD-systems. Some of these techniques will be briefly considered in the next section. However, by far almost all of these techniques require so much computational time that they would be cumbersome for any large or practical TD systems.

4.3.3 Grapho-Analytical Methods

In the previous two sections some analytical and graphical methods for the stability of TD systems were discussed. As it was mentioned earlier the stability of TD systems in general and its treatment in frequency-domain in particular, have received considerable attention in literature [4.27-4.89]. A short survey of TD-systems stability techniques has been presented by Mukherjee and Dasgupta [4.88] in which they argue that the most convenient method is "*dual locus*" schemes such as the one by Mukherjee [4.57] and N-Nagy and Al-Tikriti [4.47]. The latter approach was presented in the last section and the former scheme, termed "phase comparison" is discussed here. Both the transport delay [4.57] and distributed delay [4.89] will be considered.

4.3.3.1 Systems with Transport Delay

Consider the characteristic equation of a SISO TD system with unity feedback

$$1 + KG(s)\ exp(-sh) = 0 \tag{33}$$

then the following criterion sets forth the "phase comparison" method for determining the stability of this system.

11. Criterion. The linear TD system (33) is stable if the following two conditions are satisfied:

(i) the characteristic equation yields zero phase margin at a lower frequency.

(ii) the gain margin at this frequency should be unity.

It is noted that once the system's characteristic equation is decomposed into an equality relation similar to the previous section's approach, one can determine that phase margin and gain margin are zero and one, respectively. Depending on the individual problem's merit, the nature of decomposition varies. Furthermore, this criterion is equally applicable to discrete-time TD system. It is well known that the sample/hold units can be described by exponentials of the sampling time. That is, let the characteristic equation of a discrete-time TD system be described by

$$1 + \left[(1-e^{-sT})/s \right] \left[KN(s)/D(s) \right] exp(-sh) = 0 \qquad (34)$$

where T is the sampling period, and $N(s)$ and $D(s)$ are the delay-free plant's numerator and denominator polynominals, respectively. Let us decompose the characteristic equation (34) to

$$(e^{-sT}-1)K = s\, D(s)e^{sh}/N(s) \qquad (35)$$

which would become much simpler to handle. The locus of left-hand side of (35) is a circle. When the transfer functions are represented by the z-transform, it is often suitable to rewrite the characteristic equation such that the adjustable parameter of the system remains on one side while the fixed parameter is kept on the other side. The transformation $z = exp(sT) = cos\omega T + j\ sin\omega T$ can be used to evaluate the phase margin. Here again, time-delay charts can be used to try to satisfy the conditions of Criterion 11. The following two examples illustrate the stability criterion for continuous and discrete-time systems with time delay.

12. Example. Consider a sixth order unity delay TD system described by the following characteristic equation:

$$s(s+2)(s+3)(s+4)(2s+1)(3s+1) + 354(s+1)(a_1s+a_2)e^{-Ts} = 0 \qquad (36)$$

it is desired to find the values of parameters a_1 and a_2 such that the system is stable.

Let us rewrite the characteristic equation (36) as

$$-(s+a_2/a_1) = \left[s(s+2)(s+3)(s+4)(2s+1)(3s+1)e^{Ts} \right] / \left[354a_1(s+1) \right] \qquad (37)$$

Now let $\alpha = a_2/a_1$, $\omega = 1$ and $T = 1$. Then the phase angles of both sides of (37) are related by

$$arg(-\alpha-j) = arg(j) + arg(2+j) + arg(3+j) + arg(4+j)$$
$$+ arg(1+j2) + arg(1+j3) + arg(e^{j}) - arg(1+j) \tag{38}$$

The right-hand side of (38) is -63.51°. Equating both sides of (38) leads to

$$\alpha = cot(63.51°) = tan(26.49°) = 0.49836 \tag{39}$$

Let $\alpha = a_2/a_1 = 0.5$ at $\omega = 1$ and equate the magnitudes of both sides of (37) for $s=j\omega=j$ to obtain a_1, *i.e.*,

$$1.118034 = 206.15/500.63\ a_1 \tag{40}$$

or $a_1 \cong 0.37$ and $a_2 = 0.185$. These are the desired values of parameters for stability of (36).

13. Example. Consider a second-order discrete-time TD system described by

$$1 + (1-e^{-sT})/s \cdot (Ke^{-hs})/(s(s+2)) = 0. \tag{41}$$

It is desired to check its stability.

Rearranging the above equation such that the parameter K remains on one side, we obtain

$$(e^{-sT}-1) = s^2\ (s+2)\ e^{hs}/K \tag{42}$$

For $s = j\omega$, (42) becomes

$$(e^{-j\omega T}-1) = (j\omega)^2(j\omega+2)e^{j\omega h}/K \tag{43}$$

In order to apply the "phase comparison" stability criterion it is clear that the arguments (phase angles) of the terms of types (exp $(-j\omega T)-1$) and exp $(j\omega\tau)$ are necessary. Figure 10 shows time-delay charts for the above two terms. Using these charts at $\omega=1$ rad/sec, $T=1$ sec. and $h=0.5$, one obtains

$$arg(LHS) = arg(e^{-j}-1) = -120.34° \tag{44}$$

while

$$arg(RHS) = 180^\circ + arg(2+j) + arg(e^{j0.5})$$

$$= -124.75^\circ \tag{45}$$

At $\omega = 1.1$, $T_s = 1$ and $h=0.5$ the argument of both sides of (42) are

$$arg(LHS) = arg(e^{-j1.1}-1) = -123.8^\circ \tag{46}$$

and

$$arg(RHS) = 180^\circ + arg(2+e^{j1.1}) = -119.67 \tag{47}$$

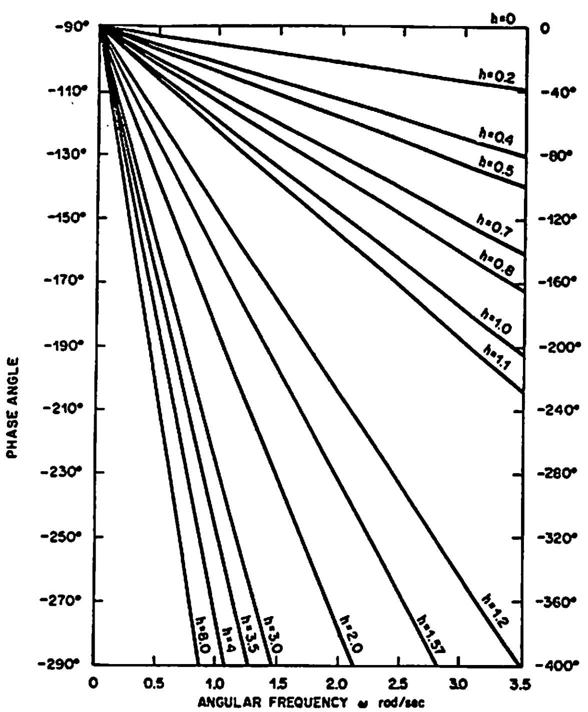

Figure 10. Time-Delay Charts for the Phase Comparison Technique

For $\omega=1.05$ the two sides arguments turn out to be -122.41° and -122.2° which can be considered as equal for practical purposes. Thus the first condition of Criterion 11 is satisfied. Using the second condition (gain or magnitude) one can find the value of K, *i.e.*,

$$\left| e^{-j1.05} - 1 \right| = \left(\left| (j1.05)^2 \right| \cdot \left| 2 + j1.05 \right| \left| 1.0 \right| \right) / K \tag{48}$$

or $K = 2.484$. Hence the system of (42) is stable for this value of K.

4.3.3.2 Systems with Distributed Delay

In this section the use of Criterion 11 is demonstrated for systems with distributed delays. This discussion is based on network of Mukherjee [4.89].

Consider a first-order TD system with proportional feedback whose open-loop transfer function is

$$G(s) = exp\left(-\sqrt{sh} \right)/s \tag{49}$$

with the characteristic equation

$$1 + K\, exp\left(-\sqrt{sh} \right)/s = 0 \ . \tag{50}$$

The application of the stability Criterion 11 requires equating phase and magnitude of both sides of (50) when $s = j\omega$. For this purpose, the term $\sqrt{s} = \sqrt{j\omega}$ is approximated as $(1 + j)\sqrt{\omega/2}$. Using this the phase condition leads to

$$180^\circ + 180^\circ \sqrt{\omega h/2} = -90^\circ \tag{51}$$

while the magnitude condition gives

$$K = \omega\, exp\sqrt{\omega h/2} \tag{52}$$

The above two relations lead to two equalities involving K, ω and h, *i.e.*,

$$\omega h = 4.93,\ Kh = 23.7 \tag{53}$$

which provide a locus of K versus h with ω acting as a parameter. This locus is shown in Figure 11.

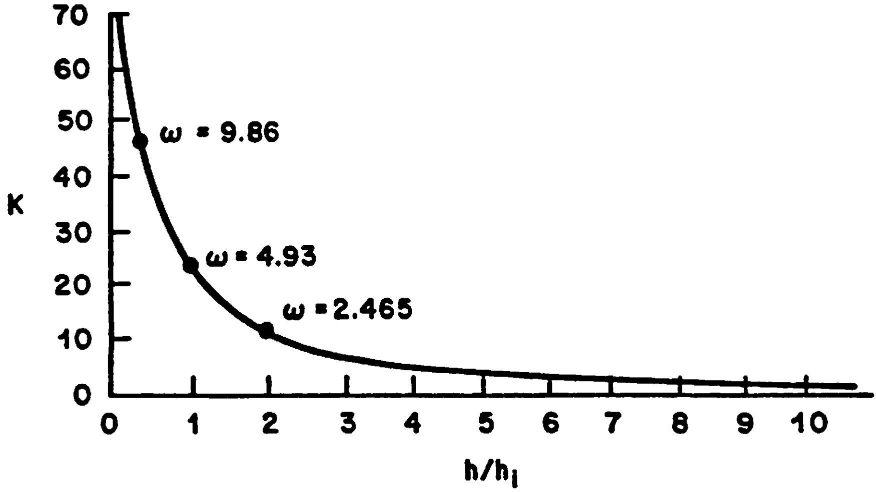

Figure 11. Stability Boundary of a First-Order System with Distributed Delay and Proportional Feedback

Next let us consider the same system with a proportional plus derivative control, *i.e.*, the characteristic equation (50) will become

$$1 + K(1+sh_d)\ exp\left[-\sqrt{sh}\right]/s = 0 \tag{54}$$

where h_d is the time constant associated with the derivation part of the controller. Normalizing (54) with respect to $p = sh$ one obtains,

$$exp\left[\sqrt{p}\right] + Kh\ (1+ph_d/h)/p = 0 \tag{55}$$

Now, applying the stability Criterion 11 to (55) yields

$$tan^{-1}\ (ph_d/h) + 90^\circ = \left[180^\circ\sqrt{p/2}\right]/\pi \tag{56}$$

and

$$\rho \ exp\left(\sqrt{\rho/2}\right)/\sqrt{1+p^2h_d^2/h^2} = Kh \tag{57}$$

The stability boundary of this case can be similarly drawn by plotting Kh^2 versus h_d/h.

Now let us consider the same system with a proportional plus integral. The characteristic equation of the system is

$$1 + K \ exp\left(-\sqrt{sh}\right)/\left(s(1+sh_i)\right) = 0 \tag{58}$$

Once again through normalization $s = \rho/h_I$ one can simplify (58) to get the following phase and magnitude conditions:

$$-tan^{-1}(\text{p})-90^{\circ} = 180^{\circ} + 180^{\circ} \sqrt{(p/2)(h/h_i)}/\pi \tag{59}$$

$$Kh_I = p\sqrt{1+p^2} \ exp(p/2)(h/h_i) \tag{60}$$

The stability boundary, plotting Kh_I versus h/h_I is similar to that of Figure 11.

The above cases illustrate one point, namely, the stability criterion of systems with transport delay can be extended to systems with distributed delay.

PROBLEMS

4.1 Consider the TD system

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -5 & 3 \\ -2 & -1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 2.5 & -0.7 \\ 0.9 & -0.1 \end{bmatrix} \mathbf{x}(t-2)$$

Check its stability by using the Lyapunov method.

4.2 For the TD system

$$\dot{x}(t) = a\,x(t) + b\,x(t-1)$$

find a region on plane $b-a$ such that the system is not stable in the sense of Lyapunov.

4.3 Consider a *2nd* order TD system,

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -2 & 0 \\ 1 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{x}(t-2)$$

Is this system uniformly asymptotically stable?

4.4 Repeat Problem 4.1 for uniform asymptotic stability

4.5 For the system

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} a & 0 \\ 1 & b \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}(t-2)$$

find conditions for uniform asymptotic stability.

4.6 A TD system is described by the following open-loop transfer function:

$$G(s) = Ke^{-hs}/(s+2)$$

Following the discussions of Example 4.3.4, check its stability.

4.7 Following the Discussions of Example 4.3.7 Determine the Stability of the TD System in Problem 4.3.

4.8 Repeat Problem 4.7 for the TD System,

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x}(t-h)$$

4.9 Repeat Example 4.3.8 for a TD system with the following characteristic equation:

$$1 + K\, e^{-hs}/(s+2) = 0$$

4.10 Use the modified Nyquist method and Example 4.3.9 to check the stability of the system

$$G\,(s) = K\, e^{-hs}/(s+1)$$

4.11 Repeat Example 4.3.9 using the Pontryagin's Stability Method of Section 4.3.1.

4.12 Consider a Stochastic TD system described by

$$d\, x(t) = -x(t)\, dt + 2\, x(t-1)\, dt + 2\, x(t)\, d\omega + 2\, x(t-1)\, d\omega$$

Use the method of Example 4.2.11 to check its stability.

4.13 For a TD system described by

$$\dot{x}\,(t) = -2\, x(t) - 2\, x(t-h(t))$$

where $h(t) = t - kT$, $k = 0,1,\cdots$ and T is a sampling time, follow the outcome of Example 4.2.12 to check its stability.

4.14 Repeat Example 4.3.10 for a TD system described by

$$KG(s)H(s)e^{-hs} = Ke^{-hs}/(s(s+1)(s+4)).$$

4.15 A fifth-order system is described by the following characteristic equation:

$$s(s+1)(s+2)(s+3)(s+4) + 20(s+5)(2s+5)\, e^{-Ts} = 0$$

Use the scheme of Section 4.3.3.1 to check its stability.

4.16 Repeat Example 4.3.13 for the following TD system,

$$1 + (1-e^{-sT})/s \cdot (K\, e^{-hs})/(s\,(s+5)) = 0$$

4.17 Demonstrate the use of Criterion 4.3.11 for the TD system described by the following transfer function:

$$G(s) = exp\left(-\sqrt{sh} \right) /(s+1)$$

(Hint: See Section 4.3.3.2.)

REFERENCES

[4.1] R. Bellman and K. L. Cooke, *Differential-Difference Equations* (New York: Academic Press, 1963)

[4.2] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach*, Pergamon Press, 1986

[4.3] R. E. Kalman and J. E. Bertram, "Control System Analysis and Design via the Second Method of Lyapunov," *J. Basic Engr.*, Vol. , pp. 371-391, 1960

[4.4] G. J. Nazaroff, "Stability of Linear Stochastic Differential Delay Systems," *IEEE Trans. Aut. Control*, Vol. AC-18, pp. 672-673, 1973

[4.5] G. J. Nazaroff and G. A. Hewer, "Stabilization of Linear, Autonomous Differential Delay Systems," *IEEE Trans. Aut. Control*, Vol. AC-18, pp. 673-674, 1973, see also G. J. Nazaroff, "Stability and Stabilization of Linear Differential Delay Systems," *IEEE Trans. Aut. Control*, Vol. AC-18, pp. 317-318, 1973

[4.6] K. Hirai and Y. Satoh, "Stability of a System with Variable Time Delay," *IEEE Trans. Aut. Control.*, Vol. AC-25, pp. 552-554, 1980

[4.7] T. N. Lee and S. Dianat, "Stability of Time-Delay Systems," *IEEE Trans. Aut. Control*, Vol. AC-26, No. 4, pp. 951-953, 1981

[4.8] A. Halanay, *Differential Equations: Stability, Oscillations and Time Lags*, Academic Press, 1966

[4.9] E. Kreindler and A. Jameson, "Conditions for Nonnegativeness of Partitioned Matrices," *IEEE Trans. Aut. Control*, Vol. AC-17, pp. 147-148, 1972

[4.10] A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse," *SIAM J. Appl. Math.*, Vol. 17, pp. 434-440, 1969

[4.11] H. J. Kushner, "On the Stability of Processes Defined by Stochastic Difference-Differential Equations," *J. Differential Equations.*, Vol. 4, pp. 424-443, 1968

[4.12] D. L. Kleinman, "On the Stability of Linear Stochastic Systems," *IEEE Trans. Aut. Control*, Vol. AC-14, pp. 429-430, 1969

[4.13] S. J. Bhatt and C. S. Hsu, "Stability Criteria for Second Order Dynamical Systems with Time Lag," *Trans. ASME J. Appl. Mech.* pp. 113-118, 1966

[4.14] R. S. Razumikhin, "The Application of Lyapunov's Method to Problem in the Stability of Systems with Delay," *Automat. Remote Contr.* Vol. 21, pp. 515-520, 1960

[4.15] J. A. York, "Asymptotic Stability for One-dimensional Differential Delay Equations," *J. Differential Equations*, Vol. 7, pp. 189-202, 1970

[4.16] L. A. Zadeh, "Operational Analysis of Variable-Delay System," *Proc. IRE*, Vol. 40, pp. 564-568, 1952

[4.17] S. G. Margolis and J. J. O'Donnel, "Rigorous Treatments of Variable Time Delays," *IEEE Trans. Electron. Computers*, Vol. EC-12, pp. 307-309, 1963

[4.18] J. P. Seddon and R. A. Johnson, "The Simulation of Variable Delay," *IEEE Trans. Computers*, Vol. C-17, pp. 89-94, 1968

[4.19] R. A. Johnson, "Functional Equations, Approximations, and Dynamic Response of Systems with Variable Time Delay," *IEEE Trans. AC.*, Vol. AC-17, pp. 398-401, 1972

[4.20] Y. Satoh, "A Study on Stability and Stabilization of a Time or State-Variable Time Lag System," M. Eng. Thesis (in Japanese), Dept. Syst. Eng., Kobe Univ., Kobe, Japan, 1978

[4.21] M. Ikeda and T. Ashida, "Stabilization of Linear Systems with Time-Varying Delay," *IEEE Trans. Aut. Control*, Vol. AC-24, pp. 369-370, 1979

[4.22] M. Jamshidi, "An Integrated Near-Optimum Control for Cold Rolling Mills," *CSL Report R-499* Univ. Illinois, Urbana, IL. 1971

[4.23] B. S. Razumikhin, "Stability in First Approximation of Systems with Lag," *Priklad Mat. i. Mekh.*, Vol. 22, pp. 155-166, 1958

[4.24] N. N. Krasovskii, *Stability of Motion*, (Stanford, CA: Stanford University Press, 1963) Chapters 6,7, Stanford, California, 1963

[4.25] K. W. Ho and W. K. Chan, "Stability of Third-Ordr Control Systems with Time Lag Using Pontryagin's Method for Zero Determination," *Proc IEE*, Vol. 116, No. 12, pp. 2063-2068, 1969

[4.26] S. Sherman, "A Note on Stability Calculations and Time Lag," *Quart. Appl. Math*. Vol. 5, pp. 92-97, 1947

[4.27] N. D. Hayes, "Roots of the Transcendental Equation Associated with a Certain Difference-Differential Equation," Lond. Math. Soc., Vol. 25, pp. 226-232, 1950

[4.28] E. M. Wright, "Stability Criteria and the Real Roots of a Transcendental Equation," *SIAM*, Vol. 9, pp. 136-148, 1961

[4.29] L. S. Pontryagin, "On the Zeros of Some Elementary Transcendental Functions," IZV. Akad. Nauk. SSSR. Ser. Mat., Vol. 6, pp. 115-134, 1942 (in Russian). See also *Amer. Math. Soc. Transl.*, Vol. 2-1, pp. 95-110, 1955

[4.30] L. S. Pontryagin, "On the Zeros of Some Elementary Transcendental Functions-Supplement," *Amer. Math. Soc. Transl.*, Vol. 2-8, pp. 19-20, 1958

[4.31] E. S. Pyatniskii, "On the Structural Stability of Single-Loop Control Systems in the Presence of Lag," *Automation and Remote Control*, Vol. 23, pp. 787-797, 1962

[4.32] E. S. Pyatniskii, "The Upper Bound on the Degree of Stability for Time-Lag Systems," *Automation and Remote Control*, Vol. 24, pp. 434-439, 1963

[4.33] O. J. Smith, *Feedback Control Systems* (New York: McGraw-Hill Book Co., 1958)

[4.34] J. E. Marshall, "On the Digital Control of Time-Delay Systems," Ph.D. Dissertation, University of Bath, Bath, England, 1972

[4.35] O. J. Gray and P. W. Hunt, "State-Feedback Controller for Systems with Dead Time," *Electronics Lett.*, Vol. 7, Nos. 5-6, pp. 131-133, 1971

[4.36] O. J. Gray and P. W. Hunt, "State-Feedback Controller with Linear Predictor for Systems with Dead Time," *Electronics Lett.*, Vol. 7, No. 12, pp. 335-337, 1971

[4.37] B. Garland and J. E. Marshall, "Sensitivity Considerations of Smith's Method for Time-Delay Systems," *Electronic Lett., Vol. 10, pp. 308-309, 1974*

[4.38] *J. E. Marshall, "Extensions of O. J. Smith's Method to Digital and Other Systems," Int. J. Control*, Vol. 19, No. 5, pp. 933-939, 1974

[4.39] B. Garland and J. E. Marshall, "Application of the Sensitivity Points Method to a Linear Predictor Control System," *Int. J Control*, Vol. 21, No. 4, pp. 681-688, 1975

[4.40] A. C. Ioannides, G. J. Rogers and V. Latham, "Stability Limits of Smith Controller in Simple Systems Containing a Time, Delay," *Int. J. Control*, Vol. 29, pp. 557-563, 1979

[4.41] J. E. Marshall, *Control of Time-Delay Systems* (Stevenage, England: Peter Peregrinus, Ltd., 1979) IEE Engr. Series 10

[4.42] A. Thowsen, "An Analytic Stability Test for a Class of Time-Delay Systems," *IEEE Trans. Aut. Control*, Vol. AC-26, pp. 735-736, 1981

[4.43] A. Gosiewski and A. W. Olbrot, "The Effects of Feedback Delays on the Performance of Multivariable Linear Control Systems," *IEEE Trans. Aut. Control*, Vol. AC-25, pp. 729-735, 1980

[4.44] S. D. Brierley, J. N. Chiasson, E. B. Lee, and S. H. Zak, "On Stability Independent of Delay for Linear Systems," *IEEE Trans. Aut. Control*, Vol. AC-27, pp. 252-254, 1982

[4.45] E. W. Kamen, "On the Relationship Between Zero Criteria for Two-Variable Polynomials and Asymptotic Stability of Delay Differential Equations," *IEEE Trans. Aut. Control*, Vol. AC-25, No. 5, pp. 983-948, 1980

[4.46] M. J. Satche', "Graphical Method for Time-Delay Analysis," *J. Appl. Mech.*, Vol. 16, p. 419, 1949

[4.47] F. L. N. Nagy and M. N. Al-Tikriti, "Stability Criterion of Linear Control Systems with Time Delay," *Meas. Control*, Vol 3, pp. 86-87, 1970

[4.48] Y. Chu, "Feedback Control Systems with Dead Time-Lag or Distributed-Lag by Root-Locus Method," *AIEE Trans. Appl. Ind.*, Vol. 71, pp. 291-296, November, 1952

[4.49] G. Tagata, "Absolute Stability of Non-Linear Control Systems with Lag," *Proc. Inst. Elec. Eng.*, Vol. 115, pp. 882-883, May, 1968

[4.50] F. El-Geldawi, "On the Phase-Plane Analysis of Relay Control System with Dead Time," *Int. J. Control*, Vol. 11, No. 1, pp. 123-131, 1970

[4.51] F. El-Geldawi, "Stabilization of a Relay System Containing Three-Position Element and Dead Time," *Int. J. Control*, Vol. 11, No. 2, pp.267-275, 1970

[4.52] S. G. Loo, "Stability of Linear Stationary Systems with Time-Delay," *Int. J. Control*, Vol. 9, No. 1, pp. 103-109, 1969

[4.53] J. Mckay, "The D-Partition Method Applied to Systems with Dead Time and Distributed Lag," *Meas. Contr.*, Vol. 3, pp. 293-296, 1970

[4.54] S. K. Mukherjee, "Stability Analysis of Retarded Linear Control Systems," *Proc. Int. Conf. on Systems and Control*, Coimbatore, India, pp. J1.1-1.11, September, 1973

[4.55] D. R. Westlund and P. N. Nikiforuk, "Relative Stability from the Dual-Locus Diagram," *IEEE Trans. Aut. Control*, Vol. AC-10, pp. 103-105, 1965

[4.56] S. Dasgupta, P. K. Malik, and A Chakravarty, "Certain Aspects of Effect of Time Delay In Space Vehicle Attitude Controller," *J. Inst. Eng.* (India), Vol. 57, pp. 163-167, February, 1977

[4.57] S. K. Mukherjee, "Stability Analysis of Linear-Continuous and Discrete-Data Control Systems with Delay," *IEEE Trans. Aut. Control*, Vol. AC-24, No. 3, pp. 505-507, 1979

[4.58] S. E. Williamson, "Accurate Root Locus Plotting Including the Effects of Pure Time Delay," *Proc. IEE*, Vol. 116, pp. 1269-1271, 1969

[4.59] I. H. Suh and Z. Bien, "A Root-Locus Technique for Linear Systems with Delay," *IEEE Trans. Aut. Control*, Vol. AC-27, No. 1, pp. 205-208, 1982

[4.60] J. S. Karmarker and D. D. Siljak, "Stability Analysis of Systems with Time-Delay," *Proc. Inst. Elec. Eng.*, Vol. 117, No. 7, pp. 1421-1424, 1970

[4.61] L. Eisenberg, "Stability of Linear Systems with Transport-Lags," *IEEE Trans. Aut. Control*, Vol. 11, pp. 247-254, April, 1966

[4.62] Y. Y. K. Chen, K. W. Han, and G. J. Thaler, "Analysis of Non-Linear Control Systems with Transport-Lags," *IEEE Trans. Ind. Gen. Appl.*, Vol. 7, pp. 576-579, September/October, 1971

[4.63] D. D. Siljak, "Sampled-Data Control System with Transport-Lag," *AIEE Appl. Ind.*, Vol. 57, pp. 247-252, 1961

[4.64] R. J. Mehrotra, "Liapunov Analysis of Systems with Transport-Lag," M. Tech. Dissertation, Elec. Eng. Dep., Banaras Hindu Univ., Varanasi, India, 1977

[4.65] K. W. Ho and W. K. Chen, "Stability of Third Order Control Systems with Time-lag Using Pontryagin's Method for Zero Determination," *Proc. Inst. Elec. Eng.*, Vol. 116, No. 12, pp. 2065-2068, 1969

[4.66] E. A. Kingsmith, "Stability Analysis of Linear Sampled-Data Time-Delay Feedback Systems," *Int. J. Control*, Vol. 14, pp. 1-26, January, 1971

[4.67] S. K. Mukherjee and S. Agrawal, "On Certain Aspects of Linear Retarded Control Systems," in *Proc. National Systems Conf.*, Coimbatore, India, pp.L 1.1-1.8, June, 1977

[4.68] S. K. Mukherjee, "Design and Stability of Sampled-Data Control Systems Using Dual-Locus Technique," *J. Inst. Electron. Telecommun. Eng. (India)*, Vol. 24, No. 1, pp. 14-16, 1978

[4.69] F. L. N. Nagy and M. N. A. Tikriti, "Stability Criteria of Linear Control Systems with Time-Delay," *Trans. Inst. Meas. Contr.*, Vol. 3, pp. 86-91, March, 1970

[4.70] M. Habermayer, "Analysis of Linear Control Systems with Dead-Time in the Time Domain," *Periodica Polytechnica (Poland, English Language)*, Vol. 18, pp. 271-284, March, 1974

[4.71] S. Dasgupta and B. Majumdar, "Applicability of Pade Approximation to a Time-Lag Systems," *J. Inst. Eng. India*, Vol. EL-49, pp. 156-162, March, 1969

[4.72] S. K. Mukherjee and S. Dasgupta, "Comments on the Stability Analysis of Linear and Sampled-Data Time-Delay Systems," *Int. J. Control*, 1978, Vol. 28, No. 3 pp. 489-492, 1978

[4.73] S. K. Mukherjee, "Design and Analysis of non-Linear Retarted Control Systems," *J. Inst. Electron. Telecommun. Eng. (India)*, Vol. 24, No. 9, pp. 374-376, 1978

[4.74] E. A. Kingsmith, "Stability Analysis of Linear Continuous Time-Delay Feedback Systems," *Int. J. Control*, Vol. 13, No. 4, pp. 633-645, 1971

[4.75] T. Nagraja, "Generation of Improved Liapunov Function," Ph.D. Dissertation, Elec. Eng. Dep., Banaras Hindu Univ., Varanasi, India, 1976

[4.76] A. Singh and D. H. McEwan, "The Control of Process Having Appreciable Transport-Lag," *IEEE Trans. Ind. Electron. Contr. Instrum.*, Vol IECI-22, pp. 396-402, 1975

[4.77] J. I. Soliman and H. Kwon, "A Frequency Stability Criterion for a Class of Non-Linear Sampled Data System with a Pure Lag," *Int. J. Control*, Vol. 7, pp. 579-590, 1968

[4.78] E. I. Jury and B. W. Lee, "On the Absolute Stability of Non-Linear Sampled-Data Systems," *IEE Trans. Aut. Control.*, Vol. AC-9, pp. 51-61, 1964

[4.79] S. E. Williamson, "Accurate Root-Locus Plotting Including the Effects of Pure Time-Delay," *Proc. Inst. Elec. Eng.*, Vol. 116, No. 7, pp. 1269-1271, 1969

[4.80] S. K. Mukherjee, "Linear Retarded Adaptive Control System Design via Dual-Locus Technique," *Proc. National Systems Conf.*, Ludhiana, India, part I, pp. 275-279, September, 1978

[4.81] S. K. Mukherjee, S. Agrawal, and S. Dasgupta, "On Certain Aspects of Linear Retarded Multivariable Feedback Control System," *J. Inst. Electron. Telecommun. Eng. (India)*, Vol, 24, No. 9, pp. 361-366, 1978

[4.82] C. F. Chen and Y. T. Tsay, "A Generalized Frequency Stability Criterion for MIMO Lumped and Distributed Parameter Feedback Systems," *Int. J. Control.*, Vol. 23, No. 3, 341-348, 1976

[4.83] S. J. Bhatt and C. S. Hsu, "Stability Criteria for Second-Order Dynamical Systems with Time-Lag," *Trans. ASME*, Vol. E-33, pp. 113-118, 1966

[4.84] C. S. Hsu and S. J. Bhatt, "Stability Chart for Second-Order Dynamical Systems with Time-Lag," *Trans. ASME*, Vol. E-33, pp. 119-124, 1966

[4.85] Y. Y. Chen, K. W. Han and C. J. Thaler, "Analysis of Non-Linear Control Systems with Transport Lag," *IEEE Trans. Ind. Gen. Appl.*, Vol. IGA-7, pp. 576-579, 1971

[4.86] M. Barszcz and A. W. Olbrot, "Stability Criterion for a Linear Differential-Difference System," *IEEE Trans. Aut. Control*, Vol. AC-24, pp. 368-369, 1979

[4.87] F. L. N. Nagy and M. N. Al-Tikriti, "Stability Determination of Linear Control Systems with Distributed-Lag," *Inst. Measur. Contr.*, Vol. 5, No. 6, pp. 238-242, 1972

[4.88] S. K. Mukherjee and S. Dasgupta, "The Analysis of Control Systems with Transport Lag," *IEEE Trans. Ind. Electron. and Control Instr.*, Vol. IECI-26, No. 26, pp. 116-119, 1979

[4.89] S. K. Mukherjee, "The Analysis of Control Systems with Distributed Lag," *IEEE Trans. Ind. Electron. Control. Instr.*, Vol. IECI-27, No. 3, pp. 207-210, 1980

CHAPTER 5

CONTROLLABILITY AND OBSERVABILITY OF LINEAR TIME-DELAY SYSTEMS

5.1 INTRODUCTION

Controllability and observability are two fundamental structural attributes of any control system. They deal, respectively, with the relationship between the input and the state and between the state and the output of the system. More specifically, system controllability addresses the following question: does a control $\mathbf{u}$ always exist which can transfer the initial state $\mathbf{x}_o$ of the system to any desired state $\mathbf{x}_1$ in a finite time? System observability addresses the following question: can the initial state $\mathbf{x}_o$ of the system be always identified by observing the output y and the input $\mathbf{u}$ over a finite time? Thus, in a TD system characterized by

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{xi}) + \mathbf{B}(t)\mathbf{u}(t) \tag{1}$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{C}_i(t)\mathbf{x}(t-h_{xi}) + \mathbf{D}(t)\mathbf{u}(t) \tag{2}$$

one expects that matrices $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{A}_2, \ldots, \mathbf{A}_N$ and $\mathbf{B}$ characterize controllability and matrices $\mathbf{C}$, $\mathbf{C}_1$, $\mathbf{C}_2, \ldots, \mathbf{C}_N$ and $\mathbf{D}$ characterize observability.

As in the case of nondelay systems [5.1], the exact definitions of controllability and observability in TD systems depend on the nature of the problem under consideration. In this chapter we will first present definitions of controllability of TD systems. We will then present criteria for the controllability of linear TD systems. Similar sections on observability will follow.

5.2 DEFINITIONS OF CONTROLLABILITY

Consider the linear TD system represented by the state equation (5.1.1) where $\mathbf{x}(t)\epsilon R^n$, $\mathbf{u}(t)\epsilon R^r$, matrices $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{A}_2, \ldots, \mathbf{A}_N$ and $\mathbf{B}$ are continuous functions of time and h_{xi}, $i=1,2,\ldots,N$ are positive constants. If $\mathbf{u}$ is measurable and bounded on every finite time interval, it will be called an *admissible* control. Let B be the

Banach space of real continuous functions defined on the interval $[t_o - \Delta, t_o]$ where Δ is the largest time delay in the system, *i.e.*,

$$\Delta = \max_i h_{xi} \tag{1}$$

Then a solution of (5.1.1) exists for $t > t_o$ and uniquely depends on the initial state $\phi(.)\epsilon B$. (See Section 3.2.) Let K be a normed vector space of functions defined on the interval $[t_o - \Delta, t_o]$. The following definitions are essentially due to Weiss [5.2].

1. Definition. System (5.1.1) is *controllable to a function* $\alpha(.)\epsilon K$ *w.r.t.* the space of initial functions B if for any given initial function $\phi \epsilon B$ there exists a finite time $t_1 > t_o$ and an admissible control $\mathbf{u}(t)$, $t\epsilon[t_o, t_1+\Delta]$ such that

$$\mathbf{x}(t,t_o,\phi,\mathbf{u}) = \alpha(t-t_1+t_o-\Delta) \quad , \quad t\epsilon[t_1, t_1 + \Delta] \tag{2}$$

where Δ is defined in (1).

2. Definition. If system (5.1.1) is controllable to *all* functions in K, then it is said to be *controllable to the space* K.

3. Definition. If $\alpha(.) \equiv 0$ in Definition 1, then the system is said to be *controllable to the origin,* or *fixed-time completely controllable.*

4. Definition. If t_1 is constant, the corresponding type of controllability is said to be *uniform.* Δ

For linear nondelay systems, controllability to the origin is equivalent to controllability to any function [5.3]. This, however, is not in general true in the case of linear TD systems.

Morse [5.21] used ring models for retarded delay-differential systems and provided an alternative definition as well as the corresponding criteria for controllability of such systems. We will not, however, detail that definition here.

5.3 CONTROLLABILITY CRITERIA

In this section we will present controllability criteria for the following system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h) + \mathbf{B}(t)\mathbf{u}(t) \tag{1}$$

where $\mathbf{x}(t)\epsilon R^n$, $\mathbf{u}(t)\epsilon R^r$, matrices $\mathbf{A}$, $\mathbf{A}_1$ and $\mathbf{B}$ are continuous functions of time and h is a positive constant. The assumption of single delay is for convenience. The results can often be easily generalized to the case of multiple delays.

Similar to the case of linear nondelay systems, the controllability criteria for linear TD systems are in the form of rank conditions on certain matrices. In most cases these criteria provide sufficient conditions for controllability. Additional constraints on the system would be needed to make such criteria necessary as well.

1. Theorem [5.2,5.4]. System (1) is controllable to the origin if there exists a finite time $t_1 > t_0$ such that

i) the matrix

$$\mathbf{Q}_c(t_1) = \int_{t_0}^{t_1} \Psi(t_1,\tau)\mathbf{B}(\tau)\mathbf{B}'(\tau)\Psi'(t_1,\tau)d\tau \tag{2}$$

has rank n where $\Psi(t,\tau)$ is a fundamental matrix of (1) with $\mathbf{u}\equiv 0$ satisfying conditions (3.2.16 a-d).

ii) for any given initial function $\phi\epsilon B$ and for some admissible $\mathbf{u}(t)$, $t\epsilon[t_o,t_1]$ such that $\mathbf{x}(t_1,t_o,\phi,\mathbf{u}_{[t_o,t_1]}) = 0$, the equation

$$\mathbf{A}_1(t)\mathbf{x}(t-h,t_0,\phi,\mathbf{u}_{[t_o,t_1]}) + \mathbf{B}(t)\mathbf{u}(t) = 0 \tag{3}$$

has an admissible solution $\mathbf{u}(.)$ on the interval (t_1, t_1+h).

Proof. In Section 3.3 we showed that the complete solution to (1) can be written as

$$\mathbf{x}(t) \triangleq \mathbf{x}(t,t_o,\phi,\mathbf{u}) = \mathbf{x}(t,t_o,\phi,0) + \int_{t_o}^{t} \Psi(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau, \quad t \geqslant t_o \tag{4}$$

where fundamental matrix $\Psi(t,\tau)$ satisfies (3.2.16 a-d). We are looking for a control $\mathbf{u}(t)$, $t\epsilon[t_o,t_1]$, which results in $\mathbf{x}(t_1,t_0,\phi,\mathbf{u}) = 0$. Consider a control

$$\hat{\mathbf{u}}(t) = \mathbf{B}'(t)\ \Psi'(t_1,t)\mathbf{K}, \qquad t_o \leqslant t \leqslant t_1 \tag{5}$$

where $\mathbf{K}$ is an arbitrary n-dimensional column vector. Thus from (4) we have

$$\mathbf{x}(t_1,t_0,\phi,\hat{\mathbf{u}}) = \mathbf{x}(t_1,t_0,\phi,0) + \int_{t_o}^{t_1} \Psi(t_1,\tau)\mathbf{B}(\tau)\mathbf{B}'(\tau)\Psi'(t_1,\tau)\mathbf{K}d\tau \tag{6}$$

Now if condition (2) is satisfied, we can let

$$\mathbf{K} = -\left[\int_{t_o}^{t_1} \Psi\ (t_1,\tau)\mathbf{B}(\tau)\mathbf{B}'(\tau)\Psi'(t_1,\tau)d\tau\right]^{-1} \mathbf{x}(t_1,t_0,\phi,0) \tag{7}$$

which, by (6), yields $\mathbf{x}(t_1,t_0,\phi,\hat{\mathbf{u}}) = \mathbf{0}$, where

$$\hat{\mathbf{u}}(t) = -\mathbf{B}'(t)\Psi'(t_1,t)\left[\int_{t_o}^{t_1} \Psi(t_1,\tau)\mathbf{B}(\tau)\mathbf{B}'(\tau)\Psi'(t_1,\tau)d\tau\right]^{-1} \mathbf{x}(t_1,t_o,\phi,0), \tag{8}$$

$$t_o \leqslant t \leqslant t_1$$

Now if (3) holds, then over the interval (t_1,t_1+h) (1) becomes

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t), \quad \mathbf{x}(t_1) = \mathbf{0} \tag{9}$$

and by uniqueness theorem of ordinary differential equations [5.5] it follows that $\mathbf{x}(t) = \mathbf{0}$ for $t\epsilon[t_1,t_1+h]$. Δ

With certain constraints on system (1), the sufficient conditions of Theorem 1 become also necessary for controllability to the origin. First consider the following definition.

2. Definition Let B be the Banach space of real continuous functions defined on the interval $[t_o-h,t_o]$. System (1) is said to be *pointwise complete* if its homogeneous solution (*i.e.*, solution when $\mathbf{u}\equiv\mathbf{0}$) corresponding to all possible initial functions $\phi \in B$ spans R^n at each point in time. Δ

Weiss [5.2] conjectured that system (1) where $\mathbf{A}$ and $\mathbf{A}_1$ are constant matrices and $\mathbf{B} = \mathbf{0}$, is pointwise complete. Popov has shown [5.6] that this conjecture is false for $n>2$ and that for $n>2$ such a system is pointwise complete under either one of the following conditions:

$i)$ $A_1 = ab'$ where a and b are arbitrary n − dimensional column vectors (10)

$ii)$ $\rho(A_1) = n$ (11)

The importance of pointwise completeness becomes apparent from the following theorem.

3. Theorem [5.2]. If system (1) is pointwise complete, then conditions of Theorem 1 are necessary as well as sufficient for controllability to the origin.

Proof. We will prove that conditions (i) and (ii) of Theorem 1 are necessary for controllability to the origin of system (1). If system (1) is controllable to the origin, then, by definition, for any $\phi \in B$ there exists a finite time $t_1 > t_o$ and an admissible control $u(.)$ on the interval $[t_o, t_1+h]$ such that $x(t,t_o,\phi,u) = 0$ for all $t \epsilon [t_1, t_1+h]$. Thus condition (ii) of Theorem (1) must hold. Now we will show, by contradiction, that condition (i) of Theorem 1 is also necessary.

Consider an initial function $\phi \epsilon B$ and suppose that there exists $t_1 > t_0$ and a control $u(t)$, $t \epsilon [t_o, t_1]$ such that $x(t_1, t_o, \phi, u) = 0$ but condition (i) of Theorem 1 does not hold. This implies that a nonzero vector $\lambda \epsilon R^n$ exists such that

$$\lambda' \left[\int_{t_o}^{t_1} \Psi(t_1,\tau) B(\tau) B'(\tau) \Psi'(t_1,\tau) d\tau \right] \lambda = 0 \tag{12}$$

or

$$\int_{t_o}^{t_1} \| \lambda' \Psi(t_1,\tau) B(\tau) \|_2^2 d\tau = 0 \tag{13}$$

which yields

$$\lambda' \Psi(t_1,\tau) B(\tau) = 0 , \quad t_o \leqslant \tau \leqslant t_1 \tag{14}$$

Thus from (4) and by hypothesis we have

$$\lambda' x(t_1,t_o,\phi,u) = \lambda' x(t_1,t_o,\phi,0) + \int_{t_o}^{t_1} \lambda' \Psi(t,\tau) B(\tau) u(\tau) d\tau \tag{15}$$

which by (14) implies

$$\lambda' \mathbf{x}(t_1,t_o,\phi,0) = 0 \tag{16}$$

Now since the system is pointwise complete, initial function $\phi \epsilon B$ can be chosen such that $\mathbf{x}(t_1,t_o,\phi,0) = \lambda$. This, by (16), would imply $\lambda'\lambda = 0$, contradicting the assumption that $\lambda \neq 0$. *q.e.d.* Δ

The following theorem due to Weiss [5.2] is a generalization of Theorems 1 and 3 to the case of controllability to a function. Its proof is essentially similar to those of Theorems 1 and 3.

4. Theorem. Let system (1) be pointwise complete and define the operator

$$L(.) = \left[\frac{d}{dt} - \mathbf{A}(t) \right] (.) \tag{17}$$

Then (1) is controllable to function $\alpha(.) \in K$ if and only if a finite time $t_1 > t_0$ exists such that condition (i) of Theorem 1 is satisfied and for any given initial function $\phi \in B$ and for some admissible control $\mathbf{u}_{[t_0,t_1]}$ such that $\mathbf{x}(t_1,t_0,\phi,\mathbf{u}[t_0,t_1]) = \alpha(t_0-h)$, an admissible solution $\mathbf{u}(.)$ on the interval $[t_1,t_1+h]$ exists to the equation

$$\mathbf{A}_1(t)\mathbf{x}(t-h,t_o,\phi,\mathbf{u}_{[t_o,t_1]}) + \mathbf{B}(t)\mathbf{u}(t) = (L\alpha)(t-t_1+t_o-h) \qquad \Delta \tag{18}$$

Application of Theorem 1 or 3 to verify the controllability of linear TD systems is difficult due to the requirement of the knowledge of a fundamental matrix of the system. This has prompted many authors to develop algebraic controllability criteria in terms of system matrices [5.7-5.10, 5.12]. Such criteria often present sufficient but not necessary conditions for controllability. In the sequel we will discuss several criteria for controllability to the origin.

For the case of *l.t.i.* single-delay systems, algebraic controllability criteria has been developed by Kirillova and Churakova [5.8]. Consider the matrix

$$\mathbf{Q} = \left[\mathbf{Q}_1^1\mathbf{B},\mathbf{Q}_1^2\mathbf{B},\mathbf{Q}_2^2\mathbf{B},\mathbf{Q}_1^3\mathbf{B},\mathbf{Q}_2^3\mathbf{B},\mathbf{Q}_3^3\mathbf{B},\ldots,\mathbf{Q}_1^n\mathbf{B},\mathbf{Q}_2^n\mathbf{B},\ldots,\mathbf{Q}_n^n\mathbf{B} \right] \tag{19a}$$

where

$$\mathbf{Q}_1^1 = \mathbf{I},\ \mathbf{Q}_j^k = \mathbf{0} \text{ for } j = 0 \text{ or } j > k \tag{19b}$$

and

$$Q_j^{r+1} = AQ_j^r + A_1Q_{j-1}^r \tag{19c}$$

The following theorem characterizes controllability to the origin of *l.t.i.* single-delay systems.

5. Theorem [5.8]. System (1) where **A**, A_1 and **B** are constant matrices is controllable to the origin if the matrix **Q** defined by (19) has rank n. Further, this condition is also necessary if the system is pointwise complete. Δ

A more general theorem dealing with linear time-varying TD systems will be presented and proved shortly. (See Theorem 7 and Problem 5.6.)

6. Example. Consider the system

$$\dot{x}(t) = Ax(t) + A_1x(t-h) + Bu(t) \tag{20}$$

where

$$A = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 2 \\ 0 & 1 & -3 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{21}$$

Note that this system is pointwise complete since condition (10) holds for $a' = [-2,1,-4]$ and $b' = [0,0,1]$. It can be easily verified from (19b) and (19c) that we have

$$\begin{aligned} &Q_1^1 = I, \quad Q_1^2 = A, \quad Q_1^3 = A^2, \\ &Q_2^2 = A_1, \quad Q_2^3 = AA_1 + A_1A, \quad Q_3^3 = A_1^2 \end{aligned} \tag{22}$$

Thus from (19a) we have

$$Q = \left[\, Q_1^1B,\ Q_1^2B,\ Q_2^2B,\ Q_1^3B,\ Q_2^3B,\ Q_3^3B \,\right] = \begin{bmatrix} 0 & -1 & -2 & 3 & 10 & 8 \\ 0 & 2 & 1 & -7 & -13 & -4 \\ 1 & -3 & -4 & 11 & 25 & 16 \end{bmatrix} \tag{23}$$

which has rank $n=3$. Therefore system (20) with coefficient matrices (21) is controllable to the origin. Δ

Buckalo [5.10] developed sufficient conditions for controllability to the origin of system (1) similar to those developed for the controllability of linear time-varying nondelay systems [5.11]. We will present his result here without proof.

7. Theorem [5.10]. System (1) is controllable to the origin if a finite time $t_1 > t_0$ exists such that

i) the matrix

$$\mathbf{Q}(t) = \left[\mathbf{P}_0(t), \mathbf{P}_1(t), \ldots, \mathbf{P}_{n-1}(t) \right] \tag{24a}$$

where

$$\mathbf{P}_0(t) = \mathbf{B}(t) \text{ and } \mathbf{P}_{k+1}(t) = \dot{\mathbf{P}}_k(t) - \mathbf{A}(t)\mathbf{P}_k(t),$$
$$k = 0,1,2,\ldots,n-2 \tag{24b}$$

has rank n for some $t \epsilon [t_1 - h, t_1]$, and

ii) The equation

$$\mathbf{A}_1(t)\mathbf{x}(t-h, t_0, \phi, \mathbf{u}) + \mathbf{B}(t)\mathbf{u}(t) = \mathbf{0} \tag{25}$$

has an admissible solution $\mathbf{u}(.)$ on the interval $[t_1, t_1 + h]$. Δ

The results of the above theorem also hold for systems with multiple state delays, *i.e.*, system (5.1.1) with initial function $\phi(t)$ defined for $t \epsilon [t_0 - h_m, t_0]$ where h_m is the maximum state delay [5.10].

Weiss [5.12] developed algebraic sufficient conditions for controllability to the origin of linear single-delay systems which reduce to algebraic sufficient conditions for nondelay systems. Further, his results reduce to sufficiency conditions of Kirillova and Churakova [5.8] in the time-invariant case and include, as a special case, the result of Buckalo [5.10]. Consider the linear single-delay system (1) and define the matrix

$$\mathbf{Q}(t) = \left[\mathbf{Q}_1^1(t),\ \mathbf{Q}_1^2(t), \ldots, \mathbf{Q}_1^n(t),\ \mathbf{Q}_2^2(t-h), \ldots, \mathbf{Q}_2^n(t-h), \ldots, \mathbf{Q}_n^n(t-(n-1)h) \right] \tag{26a}$$

where

$$Q_1^1(t) = B(t) \tag{26b}$$

$$Q_j^{k+1}(t) = \dot{Q}_j^k(t) - A(t+(j-1)h)Q_j^k(t) - A_1(t+(j-1)h)Q_{j-1}^k(t),$$

$$j = 1,2,\ldots,k+1;\ k = 1,2,\ldots,n-1 \tag{26c}$$

$$Q_j^k(t) = 0 \ \text{ for } j = 0 \text{ or } j > k \tag{26d}$$

8. Theorem [5.12]. System (1) where A and A_1 are $n-2$ times and **B** is $n-1$ times continuously differentiable for $t \geqslant t_o$ is controllable to the origin if a finite time $t_1 > t_0$ exists such that $\rho(Q(t_1)) = n$ where Q(.) is defined in (26).

NOTE: Although Weiss [5.12] does not require it, condition (3) must also be satisfied to assure that $x(t) = 0$ for $t \epsilon [t_1, t_1+h]$.

Proof. We will prove that $\rho(Q(t_1)) = n$ implies that matrix $Q_c(t_1)$ in (2) has rank n. Assume that $\rho(Q_c(t_1)) < n$ for any $t_1 > t_0$. Then a nonzero vector $\lambda \in R^n$ exists such that

$$\lambda' \, \Psi(t_1,\tau)B(\tau) = 0 \ \text{ for } t_0 \leqslant \tau < t_1 \tag{27}$$

(See the proof of Theorem 3.) In particular we have

$$\lambda' \, \Psi(t_1,\tau)B(\tau) = 0 \ \text{ for } \ \tau \epsilon [t_1-(k+1)h, t_1-kh], \quad k = 0,1,2,\ldots,n \tag{28}$$

Differentiating (28) $n-1$ times *w.r.t.* τ when $k=0$ and using (3.2.16) and (26) yields

$$\lambda' \, [\Psi(t_1,\tau)Q_1^i(\tau) + \Psi(t_1,\tau+h)Q_2^i(\tau)] = 0, \quad \tau\epsilon[t_1-h,t_1], \quad i=1,2,\ldots,n \tag{29}$$

For $\tau = t_1^-$, using (3.2.16c) and (3.2.16d), (29) becomes

$$\lambda' Q_1^i(t_1) = 0, \quad i=1,2,\ldots,n \tag{30}$$

and for $\tau = (t_1-h)^+$ using (3.2.16d), (29) becomes

$$\lambda' \Psi(t_1,t_1-h)Q_1^i(t_1-h) = 0, \quad i=1,2,\ldots,n \tag{31}$$

Differentiating (28) $n-1$ times *w.r.t.* τ when $k=1$, and using (3.2.16b) and (26) yields

$$\lambda'[\Psi(t_1,\tau)Q_1^i(\tau)+\Psi(t_1,\tau+h)Q_2^i(\tau) + \Psi(t_1,\tau+2h)Q_3^i(\tau)] = 0,$$
$$\tau\epsilon[t_1-2h,t_1-h], \quad i=1,2,\ldots,n \tag{32}$$

For $\tau = (t_1-h)^-$, using (31) and (3.2.16c), (32) becomes

$$\lambda'Q_2^i(t_1-h) = 0, \; i=1,2,\ldots,n \tag{33}$$

and for $\tau = (t_1-2h)^+$, using (3.2.16d), (32) becomes

$$\lambda'[\Psi(t_1,t_1-2h)Q_1^i(t_1-2h) + \Psi(t_1,t_1-h)Q_2^i(t_1-2h)] = 0, \; i=1,2,\ldots,n \tag{34}$$

Repeating this procedure for $k=2,3,\ldots,n$ results

$$\lambda'Q_j^k(t_1-(j-1)h) = 0, \; j=1,2,\ldots,k; \; k=1,2,\ldots,n \tag{35}$$

Since by hypothesis $\lambda\neq 0$, from (35) we conclude that

$$\rho\left\{ Q_j^k(t_1-(j-1)h) \right\} < n, \, j=1,2,\ldots,k; \; k=1,2,\ldots,n \tag{36}$$

which implies that matrix $Q(t)$ defined in (26) does not have full rank. *q.e.d.*

5.4 DEFINITIONS OF OBSERVABILITY

Consider the linear multiple-delay system represented by the state and output equations (5.1.1) and (5.1.2). The concept of observability is concerned with the following problem. Given the above system, its input $u(\cdot)$ and output $y(\cdot)$ over a finite interval $[t_o,t_1]$, determine the initial function $\phi(t)$, $t\epsilon[t_o-\Delta,t_o]$ where Δ is the maximum state delay defined by (5.2.1). As in the case of controllability, the exact definition of observability depends on the nature of the problem. The following two definitions are essentially due to Delfour and Mitter [5.14].

1. Definition. The above system is *strongly observable* in $[t_o,t_1]$ if for all initial functions $\phi(\cdot)\epsilon B$, $x(t)$ in the interval $[t-\Delta,t]$, $t\epsilon[t_o,t_1]$, can be uniquely determined from the knowledge of the input (control) $u(\cdot)$ and the output (observation) $y(\cdot)$ over $[t_o,t_1]$. Δ

The main problem is the recovery of the initial function $\phi(\cdot)$ in $[t_o-\Delta,t_o]$ from the observation. Thus we also have the following definition.

2. Definition. The above system is *observable* in $[t_o,t_1]$ if for all $\phi(\cdot)\epsilon B$, the initial function $\phi(t)$ in $[t_o-\Delta,t_o]$ can be uniquely determined from the knowledge of the control $\mathbf{u}(\cdot)$ and observation $\mathbf{y}(.)$ over $[t_o,t_1]$. Δ

Since the zero-state response can be calculated directly, the problem of system observability can be addressed when the control $\mathbf{u}(.)$ is identically zero. That is, given the system and its zero-input response over the finite interval $[t_o,t_1]$, find the initial function $\phi(.)$ in $[t_o-\Delta,t_o]$. Thus we can assume, with no loss of generality, that $\mathbf{u}(.) \equiv \mathbf{0}$ and study the observability of the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{x_i}) \quad , t \geqslant t_o \tag{1a}$$

$$\mathbf{x}(t) = \phi(t) \quad , t_o-\Delta \leqslant t \leqslant t_o \tag{1b}$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) \quad , t \geqslant t_o \tag{1c}$$

This implies that only the matrices, $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{A}_2$, . . . , $\mathbf{A}_N$; and $\mathbf{C}$ in the system representation (5.1.1) and (5.1.2) will be involved in the characterization of system observability.

In Definitions 1 and 2, B is the Banach space of real continuous functions defined on the interval $[t_o-\Delta,t_o]$. Olbrot [5.15] noted that sometimes not all the information contained in the initial function is necessary to solve the state equation of the system. (See Problem 5.8.) He proposed the following definition of observability.

3. Definition. Let F denote a class of initial functions for the system defined by (1a) and (1c). The system is said to be *F-observable* in $[t_o,t_1]$ if and only if for any initial function from F, $\mathbf{x}(.)$ in the interval $[t_o,t_1]$ can be determined from the observation $\mathbf{y}(.)$ over $[t_0,t_1]$. Δ

In practical applications it is often desirable to estimate not the initial function but a final trajectory value. This is a motivation for the following definition.

4. Definition [5.15]. System (1a) and (1c) is *finally F-observable* in $[t_o,t_1]$ if and only if for any initial function from F, $\mathbf{x}(t_1)$ can be determined from the observation $\mathbf{y}(.)$ over $[t_o,t_1]$.

Variations and extensions of the above definitions (such as *initial observability, final observability, infinite-time observability, spectral observability,* *essential observability* and *hyperobservability)* have also been presented [5.22], but will not be detailed here. Also Morse [5.21] used ring models for retarded delay-differential systems to provide an alternative definition for observability of such systems.

5.5 OBSERVABILITY CRITERIA

The criteria for observability, similar to those for controllability, are in the form of rank conditions of certain matrices. Also, in most cases, these criteria provide only sufficient conditions for observability and additional constraints on the system would be needed to make them necessary as well.

Consider the system represented by (5.4.1) in which $x(t)\epsilon R^n$, $y(t)\epsilon R^m$, matrices $A, A_1, A_2, \ldots, A_N$ are continuous functions of time and h_{x_i}, $i=1,2,\ldots,N$ are positive constraints. The following theorems are due to Delfour and Mitter [5.14].

1. Theorem. System (5.4.1) is observable in $[t_o,t_1]$ if the matrix

$$Q_o(t_1) = \int_{t_o}^{t_1} \Psi'(t,t_o)C'(t)C(t)\Psi(t,t_0)\, dt \tag{1}$$

has rank n where $\Psi(t,t_o)$ is a fundamental matrix of (5.4.1a) satisfying conditions (3.2.16a-d).

2. Theorem. If system (5.4.1) is pointwise complete, then the condition of Theorem 1 is necessary as well as sufficient for observability. Δ

The proofs of the above theorems are essentially similar to those of Theorems 5.3.1 and 5.3.3 and will not be given.

Application of Theorem 1 or 2 to verify observability of linear TD systems is difficult because they depend on the knowledge of a fundamental matrix of the system. Unlike controllability, the development of algebraic conditions for the investigation of the observability of linear TD systems has not received much attention. Sendaula [5.16] developed an algebraic criterion, similar to the one developed for controllability by Weiss [5.12], for the observability of linear systems with single time-varying delays. However, his results were disputed later [5.15, 5.17]. Some results are available for the case of stationary linear TD systems. Bhat and Koivo [5.18] used spectral decomposition of *l.t.i.*, single delay systems to decompose the state space into a finite-dimensional and a complementary part. They developed necessary and sufficient conditions for the observability of the finite-dimensional part of such systems. Malek-Zavarei [5.19] developed conditions for the observability of multiple-delay systems. Olbrot [5.15, 5.20] developed algebraic necessary and sufficient conditions for L^p-observability of stationary linear TD systems.

The following two theorems are from Reference [5.19].

3. Theorem. System (5.4.1) is observable in $[t_0,t_1]$ if and only if the observability Grammian

$$G(\alpha) = \int_{t_0}^{t_f} \int_{t-\Delta}^{t} \bar{\Phi}'(\sigma,\alpha)C'(\sigma) \int_{t_0-\Delta}^{t_0} C(\sigma)\bar{\Phi}(\sigma,\beta)d\beta d\sigma dt \tag{2}$$

where

$$\bar{\Phi}(t,\alpha) = \Phi(t,\alpha)\delta(\alpha-t_0) + \Phi(t,\alpha),\ \alpha\epsilon[t_0-\Delta,t_0],\ t\epsilon[t_0,t_f] \tag{3}$$

and $\Phi(t,\alpha)$ is the fundamental matrix of (1), has rank n for all $\alpha\epsilon[t_o-\Delta,t_0]$. Δ

Due to the difficulty in computing $\bar{\Phi}(t,\zeta)$, checking the rank of $G(\alpha)$ in (2) for observability is, in general, extremely difficult. In the next theorem an algebraic sufficient condition will be established for observability of system (5.4.1). The condition involves only the matrices $A(t)$, $A_i(t)$, $i = 1,2,\ldots,N$ and $C(t)$.

Define the observability matrix as

$$\mathbf{P}(t) = \left[\mathbf{P}_1^1,\mathbf{P}_1^2,\ldots,\mathbf{P}_1^n,\mathbf{P}_{2,J_1}^2,\mathbf{P}_{2,J_1}^3,\ldots,\mathbf{P}_{2,J_1}^n,\mathbf{P}_{3,J_2}^3,\mathbf{P}_{3,J_2}^4,\ldots,\mathbf{P}_{3,J_2}^n,\ldots,\mathbf{P}_{n,J_{n-1}}^n\right] \tag{4}$$

where dependence on t is dropped inside the square brackets for convenience. The sets J_k are defined as

$$J_k = \{l_1,l_2,\ldots,l_k\},\ l_i = 1,2,\ldots,N$$
$$i = 1,2,\ldots,k,\ k = 1,2,\ldots,n-1 \tag{5}$$

Also

$$\mathbf{P}_1^1(t) = \mathbf{C}'(t) \tag{6a}$$

$$\mathbf{P}_1^{j+1}(t) = \mathbf{A}'(t)\mathbf{P}_1^j(t) + \dot{\mathbf{P}}_1^j(t),\ j = 1,2,\ldots,n-1 \tag{6b}$$

$$\mathbf{P}_{2,J_1}^{j+1}(t) = \mathbf{A}'(t-h_{l_1})\mathbf{P}_{2,J_1}^j(t) + \mathbf{A}'_{l_1}(t)\mathbf{P}_1^j(t) + \dot{\mathbf{P}}_{2,J_1}(t),$$
$$j = 1,2,\ldots,n-1 \tag{6c}$$

$$\mathbf{P}_{k+1,J_k}^{j+1} = \mathbf{A}'\left(t - \sum_{l=1}^{k} h_{l_l}\right)\mathbf{P}_{k+1,J_k}^j(t) + \mathbf{A}'_{l_k}\left(t - \sum_{l=1}^{k-1} h_{l_l}\right)\mathbf{P}_{k,J_{k-1}}^j(t)$$
$$+ \dot{\mathbf{P}}_{k+1,J_k}^j(t),\ k = 2,3,\ldots,n-1,\ j \geqslant k \tag{6d}$$

$$\mathbf{P}_k^j(t) = 0 \quad \text{for} \quad k > j \tag{6e}$$

$$P_{k,J}^j = 0 \quad \text{for} \quad k > j \text{ and } \textit{any set } J \tag{6f}$$

$$\mathbf{P}_{k,J}^j = 0 \text{ if } J \textit{ is an empty set.} \tag{6g}$$

4. Theorem. If matrices $\mathbf{A}_i(t)$ and $i = 1,2,\ldots,N$ are $n-2$ times, and matrix $\mathbf{C}(t)$ is $n-1$ times continuously differentiable on $[t_0,t_f]$, then system (5.4.1) is observable in $[t_0,t_f]$ with $t_f > t_o + \Delta$ if the observability matrix $\mathbf{P}(\sigma)$ in (4) has rank n for some $\sigma\epsilon[t_0,t_f]$.

5. Special Cases

1. For stationary linear systems with multiple delays and $\mathbf{A} = \mathbf{0}$, the observability matrix (4) becomes

$$\mathbf{P} = \left[\mathbf{P}_1^1,\mathbf{P}_{2,J_1}^2,\mathbf{P}_{3,J_2}^3,\ldots,\mathbf{P}_{n,J_{n-1}}^n\right] \tag{7}$$

where, from relations (6),

$$\mathbf{P}_1^1 = \mathbf{C}', \quad \mathbf{P}_{2,J_1}^2 = \mathbf{A}'_{l_1}\mathbf{P}_1^1, \mathbf{P}_{k+1,J_k}^{k+1} = \mathbf{A}'_{l_k}\mathbf{P}_{k,J_{k-1}}^k, \quad k = 2,3,\ldots,n-1. \tag{8}$$

Thus, by (7)

$$\mathbf{P} = \Big[I,\mathbf{A}'_1,\ldots,\mathbf{A}'_m,\mathbf{A}'^2_1,\mathbf{A}'_1\mathbf{A}'_2,\ldots,\mathbf{A}'_1\mathbf{A}'_m,\mathbf{A}'_2\mathbf{A}_1,\mathbf{A}'^2_2,\ldots,$$
$$\mathbf{A}'_2\mathbf{A}'_m,\ldots,\mathbf{A}'^2_m,\ldots,\mathbf{A}'^{n-1}_m\Big]\mathbf{C}' \tag{9}$$

2. For the single delay case $(m=1)$ and $\mathbf{A}=\mathbf{0}$, (9) reduces to

$$\mathbf{P} = \left[\mathbf{I},\mathbf{A}'_1,\mathbf{A}'^2_1,\ldots,\mathbf{A}'^{n-1}_1\right]\mathbf{C}' \tag{10}$$

3. For the case of single delay ($m=1$), the sets J_k in (5) become

$$J_K = \underset{(k\ times)}{\{1,1,\ldots,1\}}, \quad k = 1,2,\ldots,n-1 \tag{11}$$

4. For linear time-varying nondelay systems where $A_i(t) = 0$, $i = 1,2,\ldots,m$, all the sets J_k, $k = 1,2,\ldots,n-1$ will be empty and by (6e-f) the observability matrix (4) becomes

$$\mathbf{P}(t) = \left[\mathbf{P}_1(t),\mathbf{P}_2(t),\ldots,\mathbf{P}_n(t)\right] \tag{12}$$

where, from (6a) and (6b)

$$\mathbf{P}_1(t) = \mathbf{C}'(t), \tag{13}$$

$$\mathbf{P}_{j+1}(t) = \mathbf{A}'(t)\mathbf{P}_j(t) + \dot{\mathbf{P}}_j(t),\ j = 1,2,\ldots,n-1 \tag{14}$$

Thus if $\mathbf{A}(t)$ is $n-2$ times and $\mathbf{C}(t)$ is $n-1$ times continuously differentiable on $[t_0,t_f]$, by Theorem 4 the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t),\ t_0 \leqslant t \leqslant t_f,\ \mathbf{x}(t)\epsilon R^n \tag{15}$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) \tag{16}$$

will be observable in $[t_0,t_f]$ if matrix $\mathbf{P}(t)$ in (12) has rank n for some $t\epsilon[t_0,t_f]$. This is identical to the result of Silverman and Meadows [5.11].

6. Example. Consider the system

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-1)\\ \mathbf{y}(t) &= \mathbf{C}(t)\mathbf{x}(t)\end{aligned} \tag{17}$$

where

$$\mathbf{A}(t) = \begin{bmatrix} 0 & 1 \\ -2 & -t \end{bmatrix},\ \mathbf{A}_1(t) = \begin{bmatrix} 0 & 0 \\ -t & -1 \end{bmatrix},\ \mathbf{C}(t) = [1\ 0]\,. \tag{18}$$

We have $n = 2$, $m = p = 1$ and the observability matrix (4) is

$$\mathbf{P}(t) = [\mathbf{C}',\mathbf{A}'(t)\mathbf{C}',\mathbf{A}'_1(t)\mathbf{C}'] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{19}$$

which has full rank. Therefore the system is observable for $t \geqslant 0$.

7. Example. Consider system (17) where

$$\mathbf{A}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -3 & -1 \end{bmatrix}, \mathbf{A}_1(t) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & -2 & 3 \end{bmatrix}, \mathbf{C}(t) = [\,1 \;\; 0 \;\; 0\,] \tag{20}$$

We have $n = 3$, $m = p = 1$ and the observability matrix (4) becomes

$$\mathbf{P} = [\mathbf{I},\mathbf{A},\mathbf{A}'^2,\mathbf{A}'_1,\mathbf{A}'\mathbf{A}'_1\mathbf{A}',\mathbf{A}'^2_1]\mathbf{C}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \tag{21}$$

which has full rank. Therefore the system is observable for $t \geqslant 0$.

PROBLEMS

5.1 Show that in Theorem 5.3.1, fundamental matrix $\Psi(t,\tau)$ can be replaced by matrix $\Phi_u(t,\tau)$ satisfying conditions (3.3.9 a-c).

5.2 Show that the single input *l.t.i.* TD system (5.3.20) where $\mathbf{A}_1 = \mathbf{B}\mathbf{C}'$ for any arbitrary constant vector $\mathbf{C}$ is controllable to the origin if and only if the pair $(\mathbf{A},\mathbf{B})$ is controllable [5.9].

5.3 Prove that Theorem 5.3.1 holds for the case of *l.t.i.* multiple delay systems, *i.e.*, the case where the state equation is given by (5.1.1) with $\mathbf{A}$, $\mathbf{A}_i$, $i=1,2,\ldots,N$, and $\mathbf{B}$ being constant matrices.

5.4 Give examples of third-order *l.t.i.* single-delay systems of the form $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t)$ in which i) $\mathbf{A}_1 = \mathbf{a}\mathbf{b}'$ where $\mathbf{a}$ and $\mathbf{b}$ are arbitrary column vectors, ii) $\rho(\mathbf{B}) = 3$. Check controllability to the origin of these systems using Theorems 5.3.4, 5.3.5 and 5.3.7 and compare these results.

5.5 Consider system (5.3.1) where

$$\mathbf{A} = \begin{bmatrix} 0 & -2 \\ a & t \end{bmatrix}, \quad \mathbf{A}_1 = \begin{bmatrix} 0 & b \\ -1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad h = 2$$

Apply Theorem 5.3.7 to determine constraints for constants a and b such that this system is controllable to the origin.

5.6 Show that Theorem 5.3.7 in the case of time-invariant systems reduces to Theorem 5.3.5.

5.7 Show that matrix $\mathbf{Q}(t)$ in (5.3.26) for the nondelay case ($\mathbf{A}_1$=0) becomes

$$\mathbf{Q}(t) = \left[\ \mathbf{P}_o(t),\ \mathbf{P}_1(t), \ldots,\ \mathbf{P}_{n-1}(t)\ \right]$$

where

$$\mathbf{P}_o(t) = \mathbf{B}(t) \quad \text{and} \quad \mathbf{P}_{k+1}(t) = \dot{\mathbf{P}}_k(t) - \mathbf{A}(t)\mathbf{P}_k(t),\ k = 0,1,2,\ldots,n-2$$

Also show that $\mathbf{Q}(t)$ for the *l.t.i.* case with $\mathbf{A}_1 = \mathbf{0}$ becomes

$$\mathbf{Q} = \left[\ \mathbf{B},\ \mathbf{AB},\ \mathbf{A}^2\mathbf{B}, \ldots,\ \mathbf{A}^{n-1}\mathbf{B}\ \right].$$

5.8 This example due to Olbrot [5.15] indicates that sometimes a TD system cannot distinguish between different initial functions. Consider the system $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-1),\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$ where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},\ \mathbf{A}_1 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix},\ \mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Assume that the initial function $\phi(t)$ is defined as follows: $\mathbf{x}_1(t)$ is arbitrary on [-1,0] with $\mathbf{x}_1(0) = 0$, $\mathbf{x}_2(t) = 0$ *on* [-1,0]. Show that the corresponding solution is $\mathbf{x}_1(t) = \mathbf{x}_2(t) = \mathbf{y}(t) = 0$ for all $t \geqslant 0$. Thus, this system cannot distinguish between the above initial function and the zero initial function.

5.9 Prove Theorems 5.5.1 and 5.5.2.

5.10 Apply Theorem 5.5.3 to the system of Problem 5.8 to show that it is not observable.

5.11 Consider system (5.5.17) where

$$\mathbf{A} = \begin{bmatrix} 0 & -2 \\ a & 1 \end{bmatrix}.\ \mathbf{A}_1 = \begin{bmatrix} 0 & b \\ -1 & 1 \end{bmatrix},\ \mathbf{C} = \begin{bmatrix} 0 & 1 \end{bmatrix},\ h = 2$$

Apply Theorems 5.5.3 and 5.5.4 to determine constraints for constants a and b such that this system is F-observable for the cases where $F = L^p$ and $\mathbf{F}$ is arbitrary.

REFERENCES

[5.1] L. Weiss and R. E. Kalman, "Contributions to linear system theory," *Int. J. Eng. Sci.*, Vol. 3, pp. 141-171, 1965

[5.2] L. Weiss, "On the controllability of delay-differential systems," *SIAM J. Control*, Vol. 5, pp. 575-587, 1967

[5.3] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach*, Pergamon Press, 1986

[5.4] D. H. Chyung and E. B. Lee, "Optimal systems with time delays," *Proc. Third IFAC Congress*, London, 1966

[5.5] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955

[5.6] V. M. Popov, "Pointwise degeneracy of linear, time-invariant, delay-differential equations," *J. Differential Equations*, Vol. 11, pp. 541-561, 1972

[5.7] A. K. Choudhury, "Algebraic and transfer-function criteria of fixed-time controllability of delay-differential systems," *Int. J. Control*, Vol. 16, pp. 1073-1081, 1972

[5.8] F. M. Kirillova and S. V. Churakova, "The controllability problem for linear systems with aftereffect," *Differential Equations*, Vol. 3, pp. 221-225, 1967

[5.9] A. W. Olbrot, "On Controllability of linear systems with time delays in control," *IEEE Trans. Aut. Cont.*, Vol. AC-16, pp. 664-666, 1972

[5.10] A. F. Buckalo, "Explicit conditions for controllability of linear systems with time lag," *IEEE Trans. Aut. Cont.*, Vol. AC-13, pp. 193-195, 1968

[5.11] L. M. Silverman and H. E. Meadows, "Controllability and observability in time-variable linear systems," *SIAM J. Control*, Vol. 5, pp. 64-73, 1967

[5.12] L. Weiss, "An algebraic criterion for controllability of linear systems with time delays," *IEEE Trans. Aut. Cont.*, Vol. AC-15, pp. 443-444, 1970

[5.13] A. Halanay, "On the controllability of linear difference-differential systems," *Mathematical Systems Theory and Economics*, Vol. II, H. W. Kuhn and G. P. Szego, Eds., pp. 329-335, Springer-Verlag, 1970

[5.14] M. C. Delfour and S. K. Mitter, "Controllability, observability and optimal feedback control of affine hereditary differential systems," *SIAM J. Control*, Vol. 10, pp. 298-328, 1972

[5.15] A. W. Olbrot, "A counterexample to observability of linear systems with time-variable delays," *IEEE Trans. Aut. Cont.*, Vol. AC-22, pp. 281-283, 1977

[5.16] M. Sendaula, "Observability of Linear Systems with Time-Variable Delays," *IEEE Trans. Aut. Cont.*, Vol. AC-19, pp. 604-605, 1974

[5.17] M. Malek-Zavarei, "Comments on observability of linear systems with time-variable delays,"" *IEEE Trans. Aut. Cont.*, Vol. AC-27, p. 280, 1982

[5.18] K. P. M. Bhat and H. N. Koivo, "Modal characterizations of controllability and observability in time delay systems," *IEEE Trans. Aut. Cont.*, Vol. AC-21, pp. 292-293, 1976

[5.19] M. Malek-Zavarei, "Observability of linear systems with multiple delays," *Proc. IEEE Conference on Decision and Control*, pp. 1273-1275, Orlando, Florida, 1982

[5.20] A. W. Olbrot, "Observability tests for constant time-lag systems, control and cybernetics," *Pol. Acad. Sci.*, Vol. 4, pp. 71-84, 1975

[5.21] A. F. Morse, "Ring models for delay-differential systems," *IFAC J. Automatica*, Vol. 12, pp. 529-531, 1976

[5.22] E. B. Lee and A. W. Olbrot, "Observability and related structural results for linear hereditary systems," *Int. J. Control*, Vol. 34, pp. 1061-1078, 1981

Part III OPTIMIZATION

CHAPTER 6

OPTIMIZATION OF TIME-DELAY SYSTEMS

6.1 INTRODUCTION

The optimization of non-delay systems has been traditionally performed by applying either the Pontryagin's maximum principle [6.1] or the Bellman's principle of optimality [6.2]. The former application will normally result in a nonlinear two-point boundary-value of (TPBV) problem while the latter requires the solution of a partial differential equation known as the Hamilton-Jacobi-Bellman equation. In either case the solution to the optimal control problem is a very difficult under taking.

The extension of the maximum principle to time-delay systems has been developed by Kharatishvili [6.3, 6.4]. The corresponding TPBV problem for such systems, as will be seen, involves both delay (in state) and advance (in costate) terms whose exact solution is almost impossible. The primary concern of most proposed methods of solving an optimal time-delay problem is to avoid solving the associated TPBV problem. More often than not this TPBV problem is solved by an approximate method leading to a sub-optimal solution. The sub-optimal control of TD systems is discussed in Chapter 7. Our objective in this chapter is to define the optimal control and present some solution methods for TD system optimization. The case of discrete-time TD system is treated in Chapter 7.

6.2 STATEMENT OF THE PROBLEM

Consider a nonlinear TD system described by the following n-dimensional state equation:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t-h_x), \mathbf{u}(t), \mathbf{u}(t-h_u), t) \tag{1}$$

with initial state and control functions

$$\mathbf{x}(t) = \phi(t), \quad -h_x \leqslant t \leqslant t_o \tag{2}$$

$$\mathbf{u}(t) = \eta(t), \quad -h_u \leqslant t \leqslant t_o \tag{3}$$

where t_o is the initial time of the process and $\phi(.)$ and $\eta(.)$ are assumed to be continuous functions. Let the cost function be defined by

$$J(\phi, \mathbf{u}) = F(\mathbf{x}(t_f), t_f) + \int_{t_o}^{t_f} L(\mathbf{x}(t), \mathbf{u}(t))dt \tag{4}$$

where F(.), is the penalty term, L(.) is the cost function's integrand and t_f is the final time. The optimal control of the TD system can now be stated as follows. Find an optimal control function $\mathbf{u}(t)$, $t_o < t < t_f$ which satisfies the state equation (1), associated initial functions (2) and (3) while minimizing the cost functional (4).

Other versions of the optimal TD problem of concern are linear continuous-time or discrete-time plants with quadratic cost function. For the linear quadratic continuous-time case the problem is to find an optimal control $\mathbf{u}(t)$, $t_o < t < t_f$ such that

$$J = ½\, \mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + ½ \int_{t_o}^{t_f} [\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)]dt \tag{5}$$

is minimized, while satisfying

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \sum_{i=1}^{N} A_i \mathbf{x}(t-h_{xi}) + \sum_{j=1}^{R} \mathbf{B}_i \mathbf{u}(t-h_{ui}) \tag{6}$$

$$\mathbf{x}(t) = \phi(t), \quad -h_{xN} < t < t_o \tag{7}$$

$$\mathbf{u}(t) = \eta(t), \quad -h_{uR} < t < t_o \tag{8}$$

where without any loss of generality, it is assumed that $h_{x1} < h_{x2} < \cdots < h_{xN}$ and $h_{u1} < h_{u2} < \cdots < h_{uR}$ and matrix A is $n \times n$, B is $n \times r$, matrix $\mathbf{Q} \geqslant 0$ is $n \times n$ and matrix $R > 0$ is $r \times r$. Clearly when $N = R = 1$, the TD system involves single delays and the solution to the problem would be somewhat easier.

For the case of discrete-time TD system, the problem is as follows. Find an optimal control sequence $\mathbf{u}(k)$, $k_o < k < k_f$ such that a quadratic cost functional

$$J = \tfrac{1}{2}\,\mathbf{x}'(k_f)\mathbf{F}\mathbf{x}(k_f) + \tfrac{1}{2}\sum_{k=k_o}^{k_f-1}\left[\mathbf{x}'(k)\mathbf{Q}\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}\mathbf{u}(k)\right] \tag{9}$$

is minimized, while satisfying

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \sum_{i=1}^{N}\mathbf{A}_i\mathbf{x}(k-h_{xi}) + \sum_{j=1}^{R}\mathbf{B}_j\mathbf{u}(k-h_{uj}) \tag{10}$$

$$\mathbf{x}(k) = \phi(k), \quad -h_{xN} < k < k_o \tag{11a}$$

$$\mathbf{u}(k) = \eta(k), \quad -h_{uR} < k < k_o \tag{11b}$$

Similar definitions and appropriate dimensions hold true for matrices **A**, **B**, $\mathbf{A}_i$, $\mathbf{B}_i$, **Q**, **F** and **R**. The optimal control of this last equation, as discussed earlier, will be considered in Chapter 8 within the context of a large-scale TD system, as well as in Chapter 9 in applications to traffic control and water resources systems.

6.3 THE MAXIMUM PRINCIPLE

As mentioned earlier, the maximum principle of non-delay systems has been extended to the TD case by Kharatishvili [6.3]. He later extended his results to the systems with delay in both the state and the control [6.4]. As in the non-delay case, the maximum principle provides a set of necessary conditions. Let us define a Hamiltonian function

$$\begin{aligned} &H(\mathbf{x}(t), \mathbf{x}(t-h_x), \mathbf{p}(t), \mathbf{u}(t), \mathbf{u}(t-h_u), t) \triangleq \\ &-L(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{p}'(t)f(\mathbf{x}(t), \mathbf{x}(t-h_x), \mathbf{u}(t), \mathbf{u}(t-h_u), t) \end{aligned} \tag{1}$$

and seek the desired necessary conditions for optimality, described as follows [6.3]: if $\mathbf{u}^*$ is the optimal control vector resulting in an optimal state $\mathbf{x}^*(t)$, then there exists costate vector trajectory $\mathbf{p}(t)$ which together with $\mathbf{x}(t)$ and $\mathbf{u}(t)$ satisfies the following state equation:

$$\dot{\mathbf{x}}(t) = \nabla_p H(\cdot) = f(\mathbf{x}(t), \mathbf{x}(t-h_x), \mathbf{u}(t), \mathbf{u}(t-h_u), t) \tag{2a}$$

$$\mathbf{x}(t) = \phi(t), \ -h_x \leqslant t \leqslant t_o \tag{2b}$$

$$\mathbf{u}(t) = \eta(t), \ -h_u \leqslant t \leqslant t_o \tag{2c}$$

costate equation:

$$\dot{\mathbf{p}}(t) = -\nabla_x H(\cdot) - \nabla_{x_d} H(.,s)|_{s=t+h_x}, \quad t_o < t < t_f - h_x \tag{3a}$$

$$= -\nabla_x H(\cdot), \, t_f - h_x < t < t_f \tag{3b}$$

$$\mathbf{p}(t_f) = \nabla_{x(t_f)} F(\mathbf{x}(t_f), t_f), \tag{3c}$$

and maximization equation:

$$0 = \nabla_u H(\cdot) + \nabla_{u_d} H(.,s)|_{s=t+h_u}, \quad t_o < t < t_f - h_u \tag{4a}$$

$$= \nabla_u H(\cdot), \, t_f - h_u < t < t_f \tag{4b}$$

where $\mathbf{x}_d$ and $\mathbf{u}_d$ represent $\mathbf{x}(t-h_x)$ and $\mathbf{u}(t-h_u)$, respectively. A close look at the above conditions indicates that five time arguments are involved, *i.e.*, $t - h_x$, $t - h_u$, t, $t + h_u$ and $t + h_x$. This behavior would indicate that even if one can eliminate $\mathbf{u}(t)$ in favor of $\mathbf{x}(t)$ and $\mathbf{p}(t)$ in (4), the resulting problem would be a set of differential equations with two-point boundary conditions. This TPBV problem involves both delays and advance terms which makes a numerical solution, even for a simple case, formidable.

The application of the above necessary conditions to a simpler singe-delay ($N = 1$, in 6.2.6) linear time-invariant TD system with a quadratic cost functional is presented next [6.5]. Consider

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t) \tag{5a}$$

$$\mathbf{x}(t) = \phi(t), \, t_o - h \leqslant t \leqslant t_o \tag{5b}$$

$$J = \tfrac{1}{2} \int_{t_o}^{t_f} (\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t))\,dt. \tag{6}$$

In a similar fashion to the nonlinear case, the necessary conditions of optimality would become

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{A}_1\mathbf{x}(t-h), \quad t_o \leqslant t \leqslant t_f \tag{7a}$$

$$\mathbf{x}(t) = \phi(t), \quad t_o - h \leqslant t \leqslant t_o \tag{7b}$$

$$\dot{\mathbf{p}}(t) = \mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t) - \mathbf{A}'_1\mathbf{p}(t+h), \, t_o \leqslant t \leqslant t_f - h \tag{8a}$$

$$= \mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t), \, t_f - h < t \leqslant t_f \tag{8b}$$

$$\mathbf{p}(t_f) = 0 \tag{8c}$$

$$0 = -\mathbf{R}\mathbf{u}(t) + \mathbf{B}'\mathbf{p}(t) \tag{9}$$

Now, eliminating $\mathbf{u}(t)$ from (9) and substituting it into the (7) - (8) would lead to the following TPBV problem:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{S}\mathbf{p}(t) + \mathbf{A}_1\mathbf{x}(t-h) \tag{10a}$$

$$\dot{\mathbf{p}}(t) = \mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t) - \mathbf{A}'_1\,\mathbf{p}(t+h), \quad t_o \leqslant t \leqslant t_f - h, \tag{10b}$$

$$= \mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t), \quad t_f - h \leqslant t \leqslant t_f \tag{10c}$$

where $\mathbf{S} = \mathbf{B}\mathbf{R}^{-1}\mathbf{B}'$ and the above problem once again, involves both delay and advance terms. In an attempt to perform a Riccati transformation on this TPBV problem a time-advanced Riccati equation would result, while the state equation would still involve delays. The solution to the problem would still remain difficult. A so-called generalized Riccati approach is, however, discussed in the next section.

1. Example. Consider an *l.t.i.* scalar TD system with an associated quadratic cost functional

$$\dot{x}(t) = -x(t) + x(t-0.1) + u(t) \tag{11a}$$

$$x(t) = 1, \quad -0.1 \leqslant t \leqslant 0 \tag{11b}$$

$$J = \tfrac{1}{2}\int_0^2 (x^2(t) + u^2(t))dt \tag{12}$$

It is desired to develop a TPBV problem for this system.

The Hamiltonian function is given by

$$H(x,p,u) = -\tfrac{1}{2}u^2(t) - \tfrac{1}{2}x^2(t) + p(-x(t) + x(t-0.1) + u(t))$$

The necessary conditions of optimality are given by

$$\dot{x}(t) = -x(t) + x(t-0.1) + u(t), \; 0 \leqslant t \leqslant 2 \tag{13}$$

$$x(t) = 1, \quad -0.1 \leqslant t \leqslant 0$$

$$\dot{p}(t) = x(t) + p(t) - p(t+0.1), \quad 0 \leqslant t \leqslant 1.9 \tag{14a}$$

$$= x(t) + p(t), \; 1.9 \leqslant t \leqslant 2.0 \tag{14b}$$

$$p(2) = 0$$

$$0 = -u + p \tag{15}$$

After the elimination of u, the desired TPBV problem is given by

$$\dot{x}(t) = -x(t) + x(t-0.1) + p(t) \tag{16a}$$

$$x(t) = 1, \quad -0.1 < t < 0 \tag{16b}$$

$$\dot{p}(t) = p(t) - p(t+0.1), \quad 0 \leqslant t \leqslant 1.9 \tag{17a}$$

$$= p(t), \qquad 1.9 \leqslant t \leqslant 2.0 \tag{17b}$$

$$p(2) = 0$$

Two-point boundary-value problems for TD systems will be discussed in Chapter 7 where suboptimal solutions for them will be presented.

6.4 GENERALIZED RICCATI METHOD

Another method for the optimization of linear TD system is based on a linear control law proposed by Krasovskii [6.6,6.7]. This method, which we will refer to as the Generalized Riccati Method, will be discussed in this section.

For the sake of discussion, let us first consider a *l.t.i.* TD system with a single delay in the state:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t), \; t \geqslant 0 \tag{1}$$

with initial state

$$\mathbf{x}(t) = \phi(t), \qquad -h \leqslant t \leqslant 0 \tag{2}$$

where $\phi(t)$ is continuous in the interval $[-h,0]$. For the time being assume the initial time $t_0 = 0$ and the final time $t_f = \infty$. The cost functional which is to be minimized is of quadratic form given below:

$$J = \int_0^\infty [\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)]\, dt \tag{2}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are, respectively, symmetric positive semidefinite and positive definite matrices of proper dimensions. The problem is to determine a control function $\mathbf{u}^*(t)$ which satisfies the state equation (1) and minimizes the cost functional (2).

For the nondelay case, *i.e.*, when $\mathbf{A}_1 = \mathbf{0}$, the above problem reduces to the well-known *state regulator problem*. In this case, under certain conditions such as controllability of the pair (A,B), the optimal control law has been shown [6.8,6.9] to be of linear feedback form given by

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{K}\mathbf{x}(t) \tag{3}$$

where constant matrix **K** is a solution to the following Riccati equation:

$$\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{K} + \mathbf{Q} = \mathbf{0} \tag{4}$$

and the optimal cost functional is

$$J^* = \mathbf{x}'(0)\mathbf{K}\mathbf{x}(0) = \phi'(0)\mathbf{K}\phi(0) \tag{5}$$

For TD system, *i.e.*, when $\mathbf{A}_1$ in (1a) is not identically zero, Krasovskii [6.6] used a sufficient condition for optimality to conclude that, under certain conditions, the optimal control law is linear and that the optimal cost functional is of the form

$$J^* = \phi'(0)\mathbf{W}_1\phi(0) + 2\phi'(0)\int_{-h}^{0}\mathbf{W}_2(s)\phi(s)ds + \int_{-h}^{0}\int_{-h}^{0}\phi'(r)\mathbf{W}_3(r,s)\phi(s)drds \tag{6}$$

However, an explicit characterization of the optimal control was not offered and no information was given on how to obtain the matrices $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_3$ in (6).

Ross and Flügge-Lotz [6.10] used Krasovskii's results to develop explicit conditions for the optimality of *l.t.i.* TD systems similar to those for nondelay systems described above. Simultaneously, Eller, et al [6.11] used a similar approach to develop conditions for optimality of nonstationary linear TD system. In this section we will present these results.

6.4.1 The Linear Time-Invariant Case

Consider a TD system characterized by (1) with cost functional (2). The following theorem provides sufficient conditions for a linear control to be optimal.

1. Theorem [6.10]. The optimal control law for the above system is

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{B}'\left[\mathbf{W}_1\mathbf{x}(t) + \int_{-h}^{0}\mathbf{W}_2(s)\mathbf{x}(t+s)\,ds\right], \quad t \geq 0 \tag{7}$$

provided that $\mathbf{u}^*(.)$ is a stable control law and the following conditions are satisfied:

$$\mathbf{A'W_1 + W_1A - W_1BR^{-1}B'W_1 + W_2(0) + W'_2(0) + Q = 0} \tag{8}$$

$$-\frac{d\mathbf{W}_2(s)}{ds} + \left[\mathbf{A'} - \mathbf{W_1BR^{-1}B'}\right] \mathbf{W}_2(s) + \mathbf{W}_3(0,s) = \mathbf{0} \tag{9}$$

$$\frac{\partial \mathbf{W}_3(r,s)}{\partial r} + \frac{\partial \mathbf{W}_3(r,s)}{\partial s} + \mathbf{W'}_2(r)\mathbf{BR^{-1}B'W}_2(s) = \mathbf{0} \tag{10}$$

$$\mathbf{W}_2(-h) = \mathbf{W_1A_1} \tag{11}$$

$$\mathbf{W}_3(-h,s) = \mathbf{A'}_1\mathbf{W}_2(s) \tag{12}$$

where $-h \leqslant r \leqslant 0$, $-h \leqslant s \leqslant 0$ and $\mathbf{W}_1$ is symmetric positive-definite. Furthermore, the corresponding optimal cost functional will be as given in (6). Δ

The proof of the above theorem will not be given here. It is available in [6.10] where the existence of an exact solution, as well as an approximate solution to equations (8)-(12) have also been discussed.

Note that optimal control (7) has two parts and that the first part of this optimal control is similar to the optimal control (3) for the nondelay case. (Compare algebraic matrix equations (4) and (8).) The integral part of the optimal control (7) accounts for the hereditary effects in the control. Also note that if matrix $\mathbf{A}_1$ in (1a) is identically zero, the optimal control (7) will reduce to that given in (3) for the nondelay case (Problem 6.3).

2. Example. Consider the following *l.t.i.* scalar TD system

$$\dot{\mathrm{x}}(t) = \mathrm{x}(t-1) + \mathrm{u}(t), \quad t \geqslant 0 \tag{13a}$$

$$\mathrm{x}(t) = 1, \quad -1 \leqslant t \leqslant 0 \tag{13b}$$

with the associated cost functional

$$J = \int_0^\infty [\, \mathrm{x}^2(t) + \mathrm{u}^2(t)]\, dt \tag{14}$$

The optimal control (7) in this case becomes

$$\mathrm{u}^*(t) = -W_1\mathrm{x}(t) - \int_{-1}^{0} W_2(s)\mathrm{x}(t+s)\, ds, \quad t \geqslant 0 \tag{15}$$

and the optimal cost functional (6) becomes

$$J^* = W_1 + 2\int_{-1}^{0} W_2(s)\, ds, + \int_{-1}^{0}\int_{-1}^{0} W_3(r,s)\, dr\, ds \tag{16}$$

Note that W_1, $W_2(.)$ and $W_3(.,.)$ are scalar functions. From (8) to (12), they satisfy the following relations:

$$-W_1^2 + 2\, W_2(0) + 1 = 0 \tag{17}$$

$$-\frac{dW_2(s)}{ds} - W_1 W_2(s) + W_3(0,s) = 0 \tag{18}$$

$$\frac{\partial W_3(r,s)}{\partial r} + \frac{\partial W_3(r,s)}{\partial s} + W_2(r) W_2(s) = 0 \tag{19}$$

$$W_2(-1) = W_1 \tag{20}$$

$$W_3(-1,s) = W_2(s) \tag{21}$$

where $-1 \leqslant r \leqslant 0$ and $-1 \leqslant s \leqslant 0$. The above system can be solved (numerically) to obtain W_1, W_2 and W_3. Substitution in (15) and (16) will then result in $\mathbf{u}^*$ and J^*, respectively.

6.4.2 The Linear Time-Varying Case

The state equation of the system in this case is

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h) + \mathbf{B}(t)\mathbf{u}(t), \quad t_0 \leqslant t \leqslant t_f \tag{22a}$$

$$\mathbf{x}(t) = \phi(t), \qquad t_0 - h \leqslant t \leqslant t_0 \tag{22b}$$

where $\phi(t)$ is continuous in $[\ t_0 - h, t_0\ ]$, $\mathbf{x}(t)$ is an n-vector, $\mathbf{u}(t)$ is an r-vector and matrices $\mathbf{A}(t)$, $\mathbf{A}_1(t)$ and $\mathbf{B}(t)$ are continuous in $[t_0, t_f]$. The cost functional to be minimized is

$$J = \mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + \int_{t_0}^{t_f} [\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t)]dt \tag{23}$$

where matrix $\mathbf{F}$ is symmetric and positive semi-definite, matrix $\mathbf{Q}(t)$ is continuous, symmetric and positive semi-definite, and matrix $\mathbf{R}(t)$ is continuous, symmetric and positive-definite. Eller, et-al [6.11] used an optimal cost functional of the form

$$J^* = \phi'(t_0)P_1(t_0)\phi(t_0) + 2\phi(t_0)\int_{-h}^{0} P_2(t_0,s)\phi(t_0+s)ds + \int_{-h}^{0}\int_{-h}^{0} \phi'(t_0+r)P_3(t_0,r,s)\phi(t_0+s)drds \tag{24}$$

as proposed by Krasovskii [6.7] to develop sufficient conditions for a linear control to be optimal. (Compare (24) with (6).) Their results are presented in the following theorem.

3. Theorem [6.11]. The optimal control law for the above system is

$$\begin{aligned} u^*(t) &= -R^{-1}(t)B'(t)[P_1(t)x(t) + \int_{-h}^{0} P_2(t,s)x(t+s)ds], \quad t_0 \leqslant t < t_f - h \\ &= -R^{-1}(t)B'(t)\,[w_1(t)x(t) + \int_{-h}^{t_f-h-t} w_2(t,s)x(t+s)ds\,], \quad t_f - h \leqslant t \leqslant t_f \end{aligned} \tag{25}$$

and the corresponding optimal cost functional is

$$J^* = V(t_0,\phi) \tag{26}$$

where

$$\begin{aligned} V(t,x) &= x'(t)P_1(t)x(t) + 2x'(t)\int_{-h}^{0} P_2(t,s)x(t+s)\,ds \\ &\quad + \int_{-h}^{0}\int_{-h}^{0} x'(t+r)P_3(t,r,s)x(t+s)\,dr\,ds, \quad t_o \leqslant t \leqslant t_f - h \\ &= x'(t)w_1(t)x(t) + 2x'(t)\int_{-h}^{t_f-h-t} w_2(t,s)x(t+s)\,ds \\ &\quad + \int_{-h}^{t_f-h-t}\int_{-h}^{t_f-h-t} x'(t+r)w_3(t,r,s)x(t+s)\,dr\,ds, \quad t_f - h \leqslant t \leqslant t_f \end{aligned} \tag{27}$$

provided that the following conditions are satisfied:

$$\begin{aligned} &\dot{P}_1(t) + A'(t)P_1(t) + P_1(t)A(t) - P_1(t)B(t)R^{-1}(t)B'(t)P_1(t) \\ &\quad + P_2(t,0) + P'_2(t,0) + Q(t) = 0 \end{aligned} \tag{28}$$

$$\frac{\partial P_2(t,s)}{\partial t} - \frac{\partial P_2(t,s)}{\partial s} + \left[A'(t) - P_1(t)B(t)R^{-1}(t)B'(t) \right] P_2(t,s) + P_3(t,0,s) = 0 \tag{29}$$

$$-\frac{\partial P_3(t,r,s)}{\partial t} + \frac{\partial P_3(t,r,s)}{\partial r} + \frac{\partial P_3(t,r,s)}{\partial s} + P'_2(t,r)B(t)R^{-1}(t)B'(t)P_2(t,s) = 0 \tag{30}$$

$$P_2(t,-h) = P_1(t)A_1(t) \tag{31}$$

$$P_3(t,-h,s) = A'_1(t)P_2(t,s) \tag{32}$$

where $t_0 \leqslant t < t_f - h, -h \leqslant r \leqslant 0$ and $-h \leqslant s \leqslant 0$;

$$\dot{w}_1(t) + A'(t)w_1(t) + w_1(t)\,A(t) - w_1(t)B(t)R^{-1}\,(t)B'(t)w_1(t) + Q(t) = 0 \tag{33}$$

$$\frac{\partial w_2(t,s)}{\partial t} - \frac{\partial w_2(t,s)}{\partial s} + \left[A'(t) - w_1(t)B(t)R^{-1}(t)B'(t) \right] w_2(t,s) = 0 \tag{34}$$

$$-\frac{\partial w_3(t,r,s)}{\partial t} + \frac{\partial w_3(t,r,s)}{\partial r} + \frac{\partial w_3(t,r,s)}{\partial s} + w'_2(t,r)B(t)R^{-1}(t)B'(t)w_2(t,s) = 0 \tag{35}$$

where $t_f - h \leqslant t \leqslant t_f, -h \leqslant r \leqslant t_f - h - t$ and $-h \leqslant s \leqslant t_f - h - t$; and

$$w_1(t_f) = F \tag{36}$$

$$P_1(t_f - h) = w_1(t_f - h) \tag{37}$$

$$P_2(t_f - h,s) = w_2(t_f - h,s) \tag{38}$$

$$P_3(t_f - h,r,s) = w_3(t_f - h,r,s) \tag{39}$$

Δ

The proof of the above theorem can be found in Reference [6.11] where the existence, uniqueness and continuity of the solution have been discussed. Also, a numerical method of solution has been presented in that reference which transforms the differential equations (28)-(30) and (33)-(35) to systems of ordinary differential equations and simultaneously integrates them for various values of parameters r and s. Aggarwal [6.12] later developed a more efficient numerical method to solve these equations.

6.5 DYNAMIC PROGRAMMING METHOD

In this section a brief presentation is made on how dynamic programming techniques[6.2] can be used to optimally control a linear TD system. In Chapter 9, an alternative dynamic programming scheme known as discrete - differential dynamic programming will also be introduced. The contents of this section in part follow the treatment of Oguztoreli [6.13].

Consider a *l.t.i.* TD system,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t) \tag{1}$$

with initial function,

$$\mathbf{x}(t) = \phi(t)\ , \quad -h \leqslant t \leqslant t_o \tag{2}$$

where vectors $\mathbf{x}$ and $\mathbf{u}$ are, respectively n and r dimensional and matrices A, B and A_1 have appropriate dimensions. The performance index, to be minimized, is in quadratic form, *i.e.*

$$J(\phi,\mathbf{u}) = \tfrac{1}{2}\mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_o}^{t_f}\Big[\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)\Big]dt \tag{3}$$

It is clear that for every optimal pair $\{\phi^*,\mathbf{u}^*\}\epsilon P$ for the above problem we have

$$J(\phi^*,\mathbf{u}^*) = \min_{\{\phi,\mathbf{u}\}\epsilon P, I[t_o,t_f]} J(\phi,\mathbf{u}) \tag{4}$$

where P is a feasible set for all initial functions and controls and the notation $\{\phi,\mathbf{u}\}\epsilon P, I[t_o,t_f]$ indicates that the minimization *w.r.t.* $\{\phi,\mathbf{u}\}\epsilon P$ is performed on the interval $I[t_o,t_f]$. Accordingly, let us define

$$f(t_f,t_o) = \min_{\{\phi,\mathbf{u}\}\epsilon P, I[t_o,t_f]} \{\tfrac{1}{2}\mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_o}^{t_f}(\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t))dt\} \tag{5}$$

where $\mathbf{x}(t) = \mathbf{x}(t,t_o,\phi,\mathbf{u})$ is a trajectory of (1) which correspond to an admissible pair $\{\phi, \mathbf{u}\}$, *i.e.* $\{\phi, \mathbf{u}\}\epsilon P$.

Let $\mathbf{D}(\sigma,t)$ and $\mathbf{E}(\sigma,t)$ be the kernel matrices corresponding to $\mathbf{x}$ and $\mathbf{u}$ of system (1). Then one has the following representation of the state vector,

$$\mathbf{x}(t,t_o,\phi,\mathbf{u}) = \int_{-h}^{t_o} \mathbf{D}(\sigma,t)\phi(\sigma)d\sigma + \int_{t_o}^{t} \mathbf{E}(\sigma,t)\mathbf{B}\mathbf{u}(\sigma)d\sigma \tag{6}$$

Define the penalty and integrand of (5) as

$$G(t,\tau) = \tfrac{1}{2}\mathbf{x}'(t,\tau,\phi,\mathbf{u})\mathbf{F}\mathbf{x}(t,\tau,\phi,\mathbf{u}) \tag{7a}$$

$$L(t,\tau) = \tfrac{1}{2}\mathbf{x}'(t,\tau,\phi,\mathbf{u})\mathbf{Q}\mathbf{x}(t,\tau,\phi,\mathbf{u}) + \tfrac{1}{2}\mathbf{u}'(t)\mathbf{R}\mathbf{u}(t) \tag{7b}$$

then (5) can be put in the form

$$f(t_f,t_o) = \min_{(\phi\,,\,\mathbf{u})\epsilon P,I[t_o,t_f]} \{G(t_f,t_o) + \int_{t_o}^{t_f} L(t,t_o)dt\} \tag{8}$$

Now one can define

$$f(t_f,\tau) = \min_{(\phi\,,\,\mathbf{u})\epsilon P,I[\tau,t_f]} \{G(t_f,\tau) + \int_{\tau}^{t_f} L(t,\tau)dt\} \tag{9}$$

where $\tau\epsilon[t_o,t_f]$. The properties of this function will be examined by using the Principle of Optimality[6.2] of dynamic programming.

1. Definition. *Principle of Optimality:* An optimal policy has the property that, whatever the initial decision $\mathbf{u}(t)$ and initial state $\mathbf{x}(t)$ are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. Δ

This principle is used in the following analysis. Consider Equation (9) and let us divide the interval $I[\tau,t_f]$ into two parts $I[\tau,\tau + \Delta]$ and $I[\tau + \Delta,t_f]$, where Δ is a small discrete-time interval. Thus one can write

$$f(t_f,\tau) = \min_{(\phi\mathbf{u})\epsilon P,I[\tau,t_f]} \left[G(t_f,\tau) + \int_{\tau}^{\tau+\Delta} L(t,\tau)dt + \int_{\tau+\Delta}^{t_f} L(t,\tau)dt \right] \tag{10}$$

Noting that the kernel matrices $\mathbf{D}(\sigma,t)$ and $\mathbf{E}(\sigma,t)$ are continuous in σ and t and continuously differentiable *w.r.t.* t for $\sigma\epsilon I[-h,t_o]$ and $t \geqslant t_o$, and that the function G(.) and L(.) are continuously differentiable, one has, for small Δ

$$\int_{\tau}^{\tau+\Delta} L(t,\tau)dt = \Delta L(\tau,\tau) + O(\Delta^2) \tag{11}$$

$$\int_{\tau+\Delta}^{t_f} L(t,\tau)dt = \int_{\tau+\Delta}^{t_f} L(t,\tau+\Delta)dt - \Delta\int_{\tau+\Delta}^{t_f} L'_{\tau}(t,\tau)dt + O(\Delta^2) \tag{12}$$

where $L'_{\tau}(t,\tau)$ denotes the partial derivative of $L(t,\tau)$ *w.r.t.* τ, and

$$G(t_f,\tau) = G(t_f,\tau+\Delta) - \Delta G_{\tau}'(t_f,\tau+\Delta) + O(\Delta^2) \tag{13}$$

where G_{τ}' is the partial derivative of G *w.r.t.* τ. Thus

$$f(t_f,\tau) = \min_{\{\phi u\}\epsilon P,I[\tau,t_f]} \{\Delta L(\tau,\tau) - \Delta G_{\tau}'(t_f,\tau+\Delta)$$
$$- \Delta\int_{t+\Delta}^{t_f} L'_{\tau}(t,\tau)dt + O(\Delta^2) + G(t_f,\tau+\Delta) + \int_{\tau+\Delta}^{t_f} L(t,\tau+\Delta)dt\} \tag{14}$$

The Principle of Optimality would yield

$$f(t_f,\tau) = \min_{\{\phi,u\}\epsilon P,I[\tau,\tau+\Delta]} \left\{ \min_{\{\phi,u\}\epsilon P,I[\tau+\Delta,t_f]} \left[\Delta L(\tau,\tau) - \Delta G'_{\tau}(t_f,\tau+\Delta)\right.\right.$$
$$\left.\left.- \Delta\int_{\tau+\Delta}^{t_f} L'_{\tau}(t,\tau)dt + O(\Delta^2) + G(t_f,\tau+\Delta) + \int_{\tau+\Delta}^{t_f} L(t,\tau+\Delta)dt \right] \right\} \tag{15}$$

In accordance to the definition (9) one may write

$$f(t_f,\tau) = \min_{\{\phi,u\}\epsilon P,I[\tau,\tau+\Delta]} \{f(t_f,\tau+\Delta) + \Delta L(\tau,\tau) - \Delta G'_{\tau}(t_f - \Delta,\tau)$$
$$- \Delta\int_{\tau+\Delta}^{t_f} L'_{\tau}(t,\tau)dt + O(\Delta^2)\} \tag{16}$$

Here, due to the principle of optimality, the term $f(t_f,\tau+\Delta)$ on the right-hand side is independent of the choice of the pair $\{\phi,u\}$ for the interval $I[\tau,\tau+\Delta]$. Hence we have

$$f(t_f,\tau) = f(t_f,\tau+\Delta) + \min_{\{\phi,u\}\epsilon P,I[\tau,\tau+\Delta]} \{\Delta L(\tau,\tau) - \Delta G'_{\tau}(t_f - \Delta,\tau)$$
$$- \Delta\int_{\tau+\Delta}^{t_f} L'_{\tau}(t,\tau)dt + O(\Delta^2)\} \tag{17}$$

This relation is the *discrete form of the dynamic programming condition* for an optimal pair.

If the term $f(t,\tau+\Delta)$ on the right-hand side of (17) is transferred to the left, dividing both sides by Δ and passing to the limit as $\Delta \to 0$ one gets

$$f'_\tau(t_f,\tau) = \min_{\{\phi,u\}\epsilon P} \{L(\tau,\tau) - G'_\tau(t_f,\tau) - \int_\tau^{t_f} L'_\tau(t,\tau)dt\} \tag{18}$$

where $\tau\epsilon I[t_o,t_f]$. Now, by the relation (8) one obtains

$$f(t_o,t_o) = \min_{\phi(t)\epsilon\Phi} G(\phi(t_o),t_o) \tag{19}$$

Now assuming that G(x,t) is a continuous function and since the set Φ of all feasible initial functions $\phi(t)$ is compact, the minimum of the right-hand set of (19) exists. This would mean that the differentiability of G(x,t) permits one to compute this minimum through the methods of calculus. Let us denote this minimum by γ, *i.e.*

$$\gamma = \min_{\phi(t)\epsilon\Phi} G(\phi(t_o),t_o) \tag{20}$$

Let ϕ^* correspond to the point at which the function $G(\phi(t_o),t_o)$ attains its minimum value γ, *i.e.*

$$G(\phi^*,t_o) = \gamma \tag{21}$$

Now, let us define the sets

$$\Phi^* = \{\phi(t)\epsilon\Phi \mid \phi(t_o) = \phi^*\} \tag{22}$$

$$P^* = \{\{\phi,u\}\epsilon P | \phi\epsilon\Phi^*,\ u\epsilon U\} \tag{23}$$

Since the trajectories of (1) are continuous, one can replace the space P by the space $P*$ in (18). Thus

$$f'_\tau(t_f,\tau) = \min_{\{\phi\ ,\ u\}\epsilon P*} \{L(\tau,\tau) - G'_\tau(t_f,\tau) - \int_\tau^{t_f} L'_\tau\ (t,\tau)\ dt\} \tag{24}$$

Therefore, the solution of this optimization problem is equivalent to the solution of the functional Equation (24) subject to the initial condition[6.13]

$$f(t_o,t_o) = \gamma \tag{25}$$

where γ is defined by (21).

It is clear that (24) implies two steps. One is the minimization of the difference term in the right-hand braces and the other is to equate the minimum value to the left-hand side of (24). Oguztoreli[6.13] has defined explicit expressions for the right-hand side of (24). For further insight in the application of dynamic programming to TD systems the interested reader can consult Reference [6.13] and Section 9.4.

6.6 TIME-OPTIMAL CONTROL

The time-optimal control problem is that of determining a control function which drives the system to a desired state in minimum time. This is a problem of interest in many control systems. Many authors have investigated the time-optimal control of nondelay systems (*e.g.* see references [6.17-6.23]). In this section we are concerned with time-optimal control of TD systems. This problem has also been investigated by several authors [6.1, 6.3, 6.14, 6.15, 6.24, 6.25]. The development in this section relies basically on the results of Reference [6.14].

Consider the *l.t.i.* system described by state Equation (6.5.1) repeated here for convenience:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t),\ t \geqslant 0 \tag{1a}$$

$$\mathbf{x}(t) = \phi(t),\ -h \leqslant t \leqslant t_o \tag{1b}$$

where $\mathbf{x}(t)$ is the n-dimensional continuous state vector, $\mathbf{u}(t)$ is the r-dimensional control vector and h is a positive constant delay.

The control $\mathbf{u}(t)$ is assumed to be limited in magnitude. Thus, from (1a), we can assume with no loss of generality that the components $u_k(t)$ of the control vector $\mathbf{u}(t)$ are all in the r-dimensional unit cube, *i.e.*

$$|u_k| \leqslant 1,\ k = 1,2\ ,\ldots,\ r \tag{2}$$

We denote the set of control vectors with the above property by U:

$$U = \{\mathbf{u}(t) |\ |u_k(t)| \leqslant 1,\ k = 1,2\ ,\ldots,\ r\} \tag{3}$$

Vectors $\mathbf{u}(t) \epsilon U$ are said to be *admissible control vectors.* Also, we denote the set of all real-valued n-dimensional continuous vector functions in the interval $[-h,t_o]$ by Φ. Vector functions $\phi(t) \epsilon \Phi$ are said to be *admissible initial functions.*

Now consider an arbitrary n-dimensional continuous vector function $\mathbf{z}(t)$, $t \geqslant t_o$, and let the trajectory $\mathbf{z}(t)$ represent a moving target. The time-optimal control problem is to hit target $\mathbf{z}(t)$, *i.e.* to achieve

$$\mathbf{x}(t,t_o,\phi,\mathbf{u}) = \mathbf{z}(t) \tag{4}$$

in minimum time, where $\mathbf{x}(t,t_o,\phi,\mathbf{u})$ indicates the solution of system (1) at time t. More precisely, let an admissible pair $\{\phi,\mathbf{u}\}$ exist such that the trajectories $\mathbf{x}(t,t_o,\phi,\mathbf{u})$ and $\mathbf{z}(t)$ satisfy

$$\mathbf{x}(t,t_o,\phi,\mathbf{u}) \neq \mathbf{z}(t) \text{ for } t<T \tag{5a}$$

$$\mathbf{x}(T,t_o,\phi,\mathbf{u}) = \mathbf{z}(T) \text{ for some } T>t_o \tag{5b}$$

It is clear that the value of time T depends on the pair $\{\phi,\mathbf{u}\}$, *i.e.* we have

$$T = T(\phi,\mathbf{u}) \tag{6}$$

Let

$$T^* = \min_{\phi\epsilon\Phi,\mathbf{u}\epsilon U} T(\phi,\mathbf{u}) \tag{7}$$

The time-optimal control problem under consideration is then to determine an admissible control $\mathbf{u}(t)\epsilon U$ and an admissible initial function $\phi\epsilon\Phi$ for which $T(\phi,\mathbf{u}) = T^*$.

In the case of nondelay systems, *i.e.* where $\mathbf{A}_1 = \mathbf{0}$ in (1a), the time-optimal control problem has been extensively investigated [6.17-6.23]. It is well known that the time-optimal control of a normal[1] *nth-order* system with real poles and with control of limited magnitude is of the bang-bang type. This control is unique and can be achieved by at most *n-1* switching reversals of the control [6.21]. Further, if the system is stable, any initial state in the state space can be driven to the origin in minimum time by the application of the time-optimal control. For unstable systems, however, the initial state must fall in a controllable region which is a subspace of the state space [6.23].

Time-optimal control of TD systems has also been investigated [6.1,6.3,6.14,6.24,6.25]. However, in all these contributions the initial functions $\phi(t)$ has been assumed to be fixed. Here, we consider $\phi(t)$ to range over an admissible set Φ.

1. Nondelay system $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$ where $\mathbf{B} = [b_1,b_2,\cdots,b_r]$ is said to be *normal* if each one of the systems $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + b_i\mathbf{u}(t)$, $i=1,2,\ldots,r$ is completely controllable [6.21].

Such a set is a subset of the set of all real-valued n-dimensional continuous vector functions in the interval $[t_o-h, t_o]$.

It has been shown[6.14,6.18] that if there is a pair of functions $\phi\epsilon\Phi$ and $\mathbf{u}\epsilon U$ such that (4) is satisfied, then there is a pair of functions $\phi^*\epsilon\Phi$ and $\mathbf{u}^*\epsilon U$ which is optimal, *i.e.*

$$\mathbf{x}(T^*, t_o, \phi^*, \mathbf{u}^*) = \mathbf{z}(T^*) \tag{8}$$

where T^* is defined by (7). Note that the optimal pair $\{\phi^*, \mathbf{u}^*\}$ is not necessarily unique, but the optimal time T^* is unique. Further, it has been shown[6.18] that if system (1) is *normal* (to be defined shortly), then the optimal control $\mathbf{u}^*(t)$ is of the bang-bang type. That is, if the set U^* is defined as

$$U^* = \{\mathbf{u}(t) \mid |u_k(t)| = 1, \quad k = 1,2,\ldots,r; \quad t \geqslant t_o\} \tag{9}$$

then $\mathbf{u}^*(t) \epsilon U^*$. We will follow, in part, References [6.13] and [6.14] to determine optimal controls $\mathbf{u}^*(t)$ and optimal initial functions $\phi^*(t)$.

In Chapter 3 we showed that the solution to system (1) has the form

$$\begin{aligned} \mathbf{x}(t, t_o, \phi, \mathbf{u}) = & \int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau \\ & + \int_{t_o}^{t} \Phi_u(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau, \; t \geqslant t_o \end{aligned} \tag{10}$$

where fundamental matrix $\Phi(t,\tau)$ satisfies (3.2.18a,b) and matrix $\Phi_u(t,\tau)$ satisfies (3.3.9a,b,c). We wish to match the trajectories $\mathbf{x}(t, t_o, \phi, \mathbf{u})$ and $\mathbf{z}(t)$ in minimum time. Thus we want to achieve (4) at some (minimum) time t, *i.e.* from (10),

$$\int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau + \int_{t_o}^{t} \Phi_u(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau = \mathbf{z}(t), \quad t \geqslant t_o \tag{11}$$

Define the vector functional or the left-hand side of (11) as $\Omega(t,\phi,\mathbf{u})$, *i.e.*

$$\Omega(t,\phi,\mathbf{u}) = \int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(t)d\tau + \int_{t_o}^{t} \Phi_u(t,\tau)\mathbf{B}(\tau)u(\tau)d\tau \tag{12}$$

Also define the sets

$$\Gamma(t) = \{\Omega(t,\phi,\mathbf{u}) \mid \phi\epsilon\Phi, \mathbf{u}\epsilon U\} \tag{13}$$

$$\Gamma^*(t) = \{\Omega(t,\phi,\mathbf{u}) \mid \phi\epsilon\Phi, \mathbf{u}\epsilon U^*\} \tag{14}$$

Thus (11) is equivalent to

$$\Omega(T,\phi,\mathbf{u}) = \mathbf{z}(T) \tag{15}$$

where $T = T(\phi,\mathbf{u})$ is the smallest solution of (4) which is greater than t_o. It can be shown[6.13,6.14] that $\Gamma(t)$ is convex and compact and that $\Gamma(t) = \Gamma^*(t)$. Thus any control achievable by a control vector in U is also achievable by a control vector in U^*. This then establishes that if there is an optimal control, there is always a *bang-bang* control that is optimal.

6.6.1 Optimal Controls

The following theorem establishes the form of an optimal control.

1. **Theorem** [6.13]. All optimal control functions $\mathbf{u}^*(t)$ are of the form

$$\mathbf{u}^*(t) = sgn\,[\Psi\Phi_u(T^*,t)\mathbf{B}(t)],\ \ t_o \leqslant t \leqslant T^* \tag{16}$$

where Ψ is some nonzero n-dimensional row vector depending on $\phi^*(t)$.
Proof. We have, from (12),

$$\Omega(T^*,\phi^*,\mathbf{u}) = \mathbf{z}(T^*) = \int_{t_o-h}^{t_o} \Phi(T^*,\tau)\phi^*(\tau)d\tau$$
$$+ \int_{t_o}^{T^*} \Phi_u(T^*,\tau)B(\tau)\mathbf{u}(\tau)d\tau \tag{17}$$

It has been shown [6.18] that an n-dimensional row vector Ψ always exist such that

$$\Psi\Omega(T^*,\phi,\mathbf{u}) \leqslant \Psi\Omega(T^*,\phi^*,\mathbf{u}^*) \tag{18}$$

for all $\phi\epsilon\Phi$ and $\mathbf{u}\epsilon U$. Thus from (15) with $\phi = \phi^*$ and (18) we have

$$\int_{t_o}^{T^*} \Psi\Phi_u\ (T^*,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau \leqslant \int_{t_o}^{T^*} \Psi\Phi_u(T^*,\tau)\mathbf{B}(\tau)\mathbf{u}^*(\tau)d\tau \tag{19}$$

for all $\mathbf{u}\epsilon U$. Now let

$$\lambda(t) = \Psi\Phi_u(T^*,t)\mathbf{B}(t) \tag{20}$$

Thus for $\mathbf{u}^*\epsilon U$ we have

$$\int_{t_o}^{T^*} \Psi\Phi_u(T^*,\tau)\mathbf{B}(\tau)\mathbf{u}^*(\tau)d\tau \leqslant \sum_{j=1}^{r} \int_{t_o}^{T^*} |\lambda_j(\tau)|d\tau$$

$$= \sum_{j=1}^{r} \int_{t_o}^{T^*} \lambda_j(\tau)\ sgn(\lambda_j(\tau))d\tau \tag{21}$$

where $\lambda_j(t)$ denotes the jth component of vector $\lambda(t)$. Hence it is clear that on any interval of positive length where $\lambda_j(t) \neq 0$, we have

$$\mathbf{u}_j^*(t) = sgn(\lambda_j(t)) \tag{22}$$

and the theorem is proved.

2. Definition [6.13]. System (1) is said to be *normal* if for $\Psi \neq 0$, no component of $\Psi\Phi_u(T^*,t)\mathbf{B}(t)$ is identically zero on an interval of positive length. Δ
Thus, according to the above definition, system (1) is normal if the components of $\Phi_u(T^*,t)\mathbf{B}(t)$ are linearly independent on each interval of positive length. Therefore, if system (1) is normal, the control vector $\mathbf{u}^*(t)$ in (22) will be unique. This establishes the following theorem.

3. Theorem [6.13, 6.18]. If system (1) is normal, the optimal control $\mathbf{u}^*(t)$ is of the bang-bang type and is unique no matter what the vector $\Psi \neq \mathbf{0}$ is. Δ
Note that if system (1) is not normal, the optimal control is still of the form (16). However, it is not necessarily unique.

6.6.2 Optimal Initial Functions

Now let us determine the form of the optimal initial functions. Using (10), (8) can be written as

$$\int_{t_o-h}^{t_o} \Phi(T^*,\tau)\phi^*(\tau)d\tau = \mathbf{g}(T^*) \tag{23a}$$

where

$$\mathbf{g}(T^*) = \mathbf{z}(T^*) - \int_{t_o}^{T^*} \Phi_u(T^*,\tau)\mathbf{B}(\tau)\mathbf{u}^*(\tau)d\tau \tag{23b}$$

Assuming that the optimal control $\mathbf{u}^*(t)$ has been determined, $\mathbf{g}(T^*)$ is completely known. Thus to obtain the optimal initial function $\phi^*(t)$, (23a) must be solved. This is a Fredholm integral equation of the first kind[6.24]. To solve it, consider the following homogeneous Fredholm equations of the second kind:

$$\int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau = \lambda\phi(t) \tag{24a}$$

$$\int_{t_o-h}^{t_o} \Phi^*(t,\tau)\zeta(\tau)d\tau = \lambda\zeta(t), \ \ t_o-h<\tau<t \tag{24b}$$

Let $\{\lambda_k\}$ be the set of eigenvalues of the above equations and $\{\phi^k(t)\}$ and $\{\zeta^k(t)\}$ be the complete sets of biorthonormalized principal eigenfunctions of $\Phi(t,\tau)$ and $\Phi^*(t,\tau)$, respectively. That is, we have

$$\Phi(t,\tau)\phi^k(t) = \lambda\phi^k(t) \tag{25a}$$

$$\Phi^*(t,\tau)\zeta^k(t) = \lambda^k\zeta^k(t) \tag{25b}$$

$$(\phi^i,\zeta^j) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{25c}$$

Also consider the integral equations

$$\int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau = 0 \tag{26a}$$

$$\int_{t_o-h}^{t_o} \Phi^*(t,\tau)\phi(\tau)d\tau = 0, \ \ t_o-h<\tau<t \tag{26b}$$

Let $\{\phi^{oi}(t)\}$ and $\{\xi^{oi}(t)\}$ indicate the sets of nontrivial solutions of (26a,b). Assume that the solvability condition

$$(\xi^{oi}(t),g(t)) = 0 \tag{27}$$

is satisfied for each $\phi^{oi}(t)$, where

$$g(t) = \int_{t_o-h}^{t_o} \Phi(t,\tau)\phi(\tau)d\tau \tag{28}$$

Let

$$g(t) = \sum_{i=1}^{\infty} c_i\ \phi^i(t), \ \ c_i = (\xi^k,g) \tag{29}$$

be the expression of $g(t)$ *w.r.t.* the biorthonormal system $\{\phi^i,\xi^i\}$. Then the following theorem establishes the form of the optimal initial functions.

4. Theorem. [6.14] Let $\{\lambda_k\}$ be the set of eigenvalues of the integral equations (24) and $\{\phi^k(t)\}$, $\{\xi^k(t)\}$ be the complete sets of eigenfunctions of $\Phi(t,\tau)$ and its transpose

$\Phi^*(t,\tau)$. Let $\{\phi^{oi}(t)\}$ and $\{\xi^{oi}(t)\}$ be the sets of nontrivial solutions of the integral equations (26). If the solvability conditions (27) is satisfied for each $\xi^{oi}(t)$, then optimal initial functions $\phi^*(t)$ are given by

$$\phi^*(t) = \sum_{i=1}^{\infty} \frac{c_i}{\lambda_i} \phi^i(t) + \tilde{\phi}(t) \tag{30}$$

where the coefficients c_i are defined in (29) and $\tilde{\phi}(t)$ is a linear combination of the functions $\phi^{oi}(t)$. Δ

The discussion in this section has been for *l.t.i* single-delay system (1). However, the development for linear time-varying multidelay systems directly follows the steps in this section.

PROBLEMS

6.1 For the nonlinear TD system

$$\dot{x} = -x^2 + x(t-0.5), \; t>0$$
$$x(t) = 1, \quad -0.5 \leqslant t \leqslant 0$$

and cost functional,

$$J = \tfrac{1}{2} \int_0^2 (x^2 + u^2)$$

find the necessary TPBV problem.

6.2 Repeat Problem 6.1 for the linear TD system

$$\dot{x} = -x + x(t-0.2) + u, \; 0 \leqslant t \leqslant 1.$$

and initial function $\phi(t) = 1, \; -0.2 \leqslant t \leqslant 0.$

6.3 Show that if matrix A_1 in (6.4.1a) is identically zero, the system of equations (6.4.8) - (6.4.12) will collapse to (6.4.4) and, as a result, the optimal control in (6.4.7) and the optimal cost functional (6.4.6) will reduce to those in (6.4.3) and (6.4.5), respectively.

6.4 Show that in (6.4.19) - (6.4.23) it can be assumed that $\mathbf{P}_1(t)$ is symmetric and $\mathbf{P}_3(t,r,s) = \mathbf{P}'(t,s,r)$ with no loss of generality.

6.5 Show that in the time-invariant case, the optimal control (6.4.25) reduces to (6.4.7) and the optimal cost functional (6.4.24) reduces to (6.4.6).

6.6 Solve the system of equations (17)-(21) and determine the optimal control and the optimal cost functional for Example 6.4.2.

6.7 Determine the optimal control and the optimal cost functional for the system of Problem 6.2 using the generalized Riccati method.

6.8 Determine the discrete-time principle of optimality relation for Problem 6.2. Use a quadratic cost function $J = \int_0^1 (x^2 + u^2)dt.$

6.9 Repeat Problem 6.8 for the *TD* system of Problem 6.1.

6.10 Assume that matrix $A_1 = 0$ in (6.6.1) (nondelay system). Develop steps parallel to those in Section 6.6 for time-optimal control of nondelay systems. Prove all the corresponding lemmas whose proofs were omitted in Section 6.6. Compare the results with existing results (*e.g.* with [6.22]).

6.11 Show the following property of the functional $\Omega(t,\phi,\mathbf{u})$ defined in (6.6.12): if $\phi\epsilon\Phi$ and $\mathbf{u}\epsilon U$ are in sufficiently small neighborhoods of $\bar{\phi}\epsilon\Phi$ and $\bar{\mathbf{u}}\epsilon U$, respectively, then for any $\epsilon > 0$ there exists a $\delta>0$ such that

$$\|\Omega(\bar{t},\bar{\phi},\bar{\mathbf{u}}) - \Omega(t,\phi,\mathbf{u})\|<\epsilon$$

for each $\Omega(\bar{t},\bar{\phi},\bar{\mathbf{u}})$ and for all $\bar{t}-\delta<t<\bar{t}$ [6.13].

6.12 Consider a linear time-varying multidelay system. Develop steps parallel to those of Section 6.6 for time-optimal control of this system.

6.13 Show that the functional $T(\phi,\mathbf{u})$ in (6.6.6) is continuous in both its arguments [6.13].

6.14 Prove that $\mathbf{z}(T^*)$ is a boundary point of the set $\Gamma(T^*)$ where $\mathbf{z}(t)$, $\Gamma(.)$ and T^* are defined in Section 6.6 [6.13]. Using this result and the convexity of the set $\Gamma(T^*)$ prove (6.6.18).

6.15 Show that Definition 6.2 of TD system normality is consistent with the definition of nondelay system normality given in the footnote in Section 6.6.

REFERENCES

[6.1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes,* translated by K. N. Trirogoff, ed. by L. W. Neustadt, Wiley Interscience, 1962

[6.2] R. Bellman, *Dynamic Programming,* Princeton University Press, Princeton, N.J. 1957

[6.3] G. L. Kharatishvili, "The Maximum Principle in the Theory of Optimal Process with Time-lags," *Dolk. Akad. Nauk. USSR,* Vol. 136, pp. 39-43, 1961

[6.4] G. L. Kharatishvili, "A Maximum Principle in External Problems With Delays," *Mathematical Theory of Control,* A. V. Balakrishnan and L. W. Neustadt (eds.) Academic Press, 1967

[6.5] E. B. Lee, "Generalized Quadratic Optimal Controllers for Linear Hereditary Systems," *IEEE Trans. Aut.* Contr., vol. AC-25, pp. 528-531, 1980

[6.6] N. N. Krasovskii, "Analytic Constructions of an Optimal Regulator in a System with Time Lags," *Prichl. Mat. Mech.* (USSR), Vol. 26, No. 1, 1962

[6.7] N. N. Krasovskii, "Optimal Processes in Systems with Time-Lags." *Proc. IFAC Congress,* Basel, Switzerland pp. 327-332, 1963

[6.8] R. E. Kalman, "Contributions to the Theory of Optimal Control," *Bull. Soc. Math. Mexico,* Vol. 5, pp. 101-119, 1960

[6.9] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach,* Pergamon Press, 1986

[6.10] D. W. Ross and I. Flügge-Lotz, "An Optimal Control Problem for Systems with Differential-Difference Equation Dynamics," *SIAM J. Control,* Vol. 7, pp. 608-623, 1969

[6.11] D. H. Eller, J. K. Aggarwal and H. T. Banks, "Optimal Control of Linear Time-Delay Systems," *IEEE Trans. Aut. Control,* Vol. AC-14, No. 6, pp. 678-687, 1969

[6.12] J. K. Aggarwal, "Computation of Optimal Control for Time-Delay Systems," *IEEE Trans. Aut. Control,* Vol. AC-15, pp. 683-685, 1970

[6.13] M. N. Oguztoreli, *Time-Lag Control Systems,* Academic Press, 1966

[6.14] M. N. Oguztoreli, A Time Optimal Control Problem for Systems Described by Differential Difference Equations," *SIAM J. Control,* Vol. 1, pp. 290-310, 1963

[6.15] E. Shimemura, "Existence and Uniqueness of Time-Optimal Control in a Linear System with a Delay," *Bull. Sci. and Eng. Res. Lab.* Waseda Univ. Japan, Vol. 26, pp. 35-41, 1964

[6.16] R. Bellman, I. Glicksberg and O. Gross, "On the Bang-Bang Control Problem," *Quart. Appl. Math.,* Vol. 14, pp. 11-18, 1956

[6.17] V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko and L. S. Pontryagin, "The Maximum Principle in the Theory of Optimal Processes of Control" *Automation and Remote Control,* pp. 1004-1008, 1960

[6.18] J. P. LaSalle, "The Time Optimal Control Problem," *Contributions to the Theory of Nonlinear Oscillations,* Vol. 5, pp.1-24, Princeton University Press, 1961

[6.19] L. W. Neustadt, "Synthesizing Time Optimal Control Systems," *J. Math. Anal. Appl.,* Vol. 1, pp. 484-493, 1960

[6.20] M. Athanassiades, "Bang-Bang Control for Tracking Systems," *IRE Trans. Aut. Control,* Vol. AC-7, pp. 77-78, 1962

[6.21] M. Athans and P. L. Falb, *Optimal Control,* McGraw-Hill, 1966

[6.22] M. Malek-Zavarei, "Time-Optimal Control of a Third-Order Plant," *Kybernetika,* Vol. 16, pp. 71-85, 1980

[6.23] M. Malek-Zavarei, "Time-Optimal Control of a Class of Unstable Third-Order Plants," *J. Franklin Institute,* Vol. 309, pp. 125-134, 1980

[6.24] R. Bellman and R. Kalaba, "Reduction of Dimensionality, Dynamic Programming and Control Processes," *J. Basic Eng. (Trans. A.S.M.E.),* Vol. 83, pp. 82-84, 1961

[6.25] A. T. Fuller, "Simplified Time-Optimal Switching Functions for Plants Containing Logs," *Int. J. Control,* Vol. 18, pp. 1141- , 1973

[6.26] J. T. Oden, *Applied Functional Analysis,* Prentice-Hall, 1979

CHAPTER 7

SUBOPTIMAL CONTROL OF TIME-DELAY SYSTEMS

7.1 INTRODUCTION

The optimization of TD systems was discussed in the previous chapter where we pointed out the computational difficulties in determining the exact optimal control. This chapter is concerned with methods of determining suboptimal control for TD systems. Such methods involve solving a simpler, often nondelay, subproblem repeatedly. The suboptimal solution moves closer to the optimal solution as the number of repetitions increases.

One approach for determining suboptimal control for TD systems is based on the concept of optimal control sensitivity [7.1-7.3]. In this approach the control is expanded into a MacLaurin series in some parameter. The coefficients of the truncated series are then computed from the optimization of some related nondelay systems. Another approach is to treat the delay terms in the state equations as extra perturbing inputs, thus converting the problem to a nondelay problem [7.4-7.5]. This method can be applied to determine suboptimal control for a certain class of nonlinear TD systems as well. Yet another approach for suboptimal control of TD systems is to transform TD systems into equivalent nondelay systems. This approach which was applied to linear TD systems independently by Bate [7.6] and Slater and Wells [7.7], allows only the integral of a quadratic form of the control, but not the state, in the cost functional.

In this chapter we will describe the former two approaches for determining suboptimal control for TD systems. The sensitivity of system performance to parameter variations will also be discussed.

7.2 SENSITIVITY APPROACH

The concept of *continuation* or *sensitivity* has been used for the solution of boundary-value problems [7.8-7.11]. This concept was first applied to determine optimal control sensitivity of nondelay systems by Werner and Cruz [7.12] and by Sannuti and Kokotovic' [7.13]. Inoue *et al.* [7.14] also used this approach to obtain suboptimal control for stationary linear systems with small delay in the state. They expanded the control into a MacLaurin series in the delay and obtained the series coefficients from the solution of simple two-point boundary-value problems. Jamshidi and Malek-Zavarei [7.1] used a MacLaurin series expansion of the control in a sensitivity parameter and obtained the series coefficients from nondelay system computations. They obtained the suboptimal control in the form of an exact feedback term and a truncated-series forward term for stationary linear systems with time delay. Chan and Perkins [7.2] also used the same technique, but obtained an open-loop suboptimal control. Malek-Zavarei [7.3] later applied a sensitivity approach to obtain the suboptimal control for nonstationary linear systems with delays in both the state and the control and with quadratic cost. He obtained the suboptimal control in the form of an exact closed-loop and truncated-series open-loop terms, where all the series coefficients were calculated in a recursive manner from nondelay system computations.

We will review the methods of References [7.1] and [7.3] in this section. The former reference is concerned with *l.t.i.* systems with a single delay in the state. The latter deals with linear time-varying systems with delays in the state and the control.

7.2.1 The Linear Time-Invariant Case

Consider the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t) \tag{1a}$$

$$\mathbf{x}(t) = \phi(t), \quad t_0 - h \leqslant t \leqslant t_0 \tag{1b}$$

where $\mathbf{x}$ and $\mathbf{u}$ are, respectively, the state and the control vectors; $\mathbf{A}$, $\mathbf{A}_1$ and $\mathbf{B}$ are matrices of appropriate dimensions, $\phi(t)$ is the initial function, t_0 is the initial process time and h is the time delay. A control $\mathbf{u}$ must be found that minimizes the quadratic cost functional

$$J = \tfrac{1}{2}\int_{t_0}^{t_f}[\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)]dt \tag{2}$$

where t_f is the final time and matrices $\mathbf{Q}$ and $\mathbf{R}$ satisfy the usual state-regulator conditions [7.15].

The Hamiltonian function for the problem is

$$H(\mathbf{x},\mathbf{p},\mathbf{u}) = \tfrac{1}{2}[\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)] + \mathbf{p}'(t)[\mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t)] \tag{3}$$

where $\mathbf{p}(t)$ is the costate vector. The necessary and sufficient conditions for optimality are (See Section 6.3.):

$$\dot{\mathbf{x}}(t) = \nabla_p H(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t),\ t_0 \leqslant t \leqslant t_f \tag{4a}$$

$$\dot{\mathbf{p}}(t) = -\nabla_x H(t) - \nabla_{x_d} H(t+h),\ \ t_0 \leqslant t \leqslant t_f - h \tag{4b}$$

$$= -\nabla_x H(t),\ \ t_f - h \leqslant t \leqslant t_f \tag{4c}$$

$$0 = \nabla_u H(t) \tag{4d}$$

where $\mathbf{x}_d(t)$ represents the delayed state, *i.e.* $\mathbf{x}(t-h)$. Calculating the gradients, (4) becomes

$$\dot{\mathbf{p}}(t) = -\mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t) - \mathbf{A}_1'\mathbf{p}(t+h),\ t_0 \leqslant t \leqslant t_f - h \tag{5a}$$

$$= -\mathbf{Q}\mathbf{x}(t) - \mathbf{A}'\mathbf{p}(t),\ t_f - h \leqslant t \leqslant t_f \tag{5b}$$

$$0 = \mathbf{R}\mathbf{u}(t) + \mathbf{B}'\mathbf{p}(t) \tag{5c}$$

Equations (4a) and (5), with initial function (1b) form a system of coupled first-order vector differential equations involving both delay and advance terms. The exact solution of this system is, in general, extremely difficult.

Now consider the family of TPBV problems

$$\dot{\mathbf{x}}(t,\epsilon) = \mathbf{A}\mathbf{x}(t,\epsilon) + \epsilon\mathbf{A}_1\mathbf{x}(t-h,\epsilon) + \mathbf{B}\mathbf{u}(t,\epsilon),\ t_0 \leqslant t \leqslant t_f \tag{6a}$$

$$\dot{\mathbf{p}}(t,\epsilon) = -\mathbf{Q}\mathbf{x}(t,\epsilon) - \mathbf{A}'\mathbf{p}(t,\epsilon) - \epsilon\mathbf{A}_1'\mathbf{p}(t+h,\epsilon),\ t_0 \leqslant t \leqslant t_f - h \tag{6b}$$

$$= -\mathbf{Q}\mathbf{x}(t,\epsilon) - \mathbf{A}'\mathbf{p}(t,\epsilon),\ t_f - h \leqslant t \leqslant t_f \tag{6c}$$

$$\mathbf{u}(t,\epsilon) = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{p}(t,\epsilon),\ t_0 \leqslant t \leqslant t_f \tag{6d}$$

with boundary conditions

$$\mathbf{x}(t,\epsilon) = \phi(t)\ ,\ t_0 - h \leqslant t \leqslant t_0 \tag{6e}$$

$$\mathbf{p}(t_f,\epsilon) = 0 \tag{6f}$$

where $0 \leqslant \epsilon \leqslant 1$ is a real scalar parameter. It is assumed that the solution to the above system exists for each ϵ and is unique. Note that for $\epsilon = 1$, the above system becomes identical to the original problem. For $\epsilon = 0$, this system reduces to

$$\dot{\mathbf{x}}^{(0)}(t) = \mathbf{A}\mathbf{x}^{(0)}(t) + \mathbf{B}\mathbf{u}^{(0)}(t) \tag{7a}$$

$$\dot{\mathbf{p}}^{(0)}(t) = -\mathbf{Q}\mathbf{x}^{(0)}(t) - \mathbf{A}'\mathbf{p}^{(0)}(t) \tag{7b}$$

$$\mathbf{u}^{(0)}(t) = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{p}^{(0)}(t) \tag{7c}$$

for $t_0 \leqslant t \leqslant t_f$. Following (6e) - (6f) the necessary boundary conditions are,

$$\mathbf{x}^{(0)}(t_0,0) = \phi(t_0) \tag{7d}$$

$$\mathbf{p}^{(0)}(t_f) = 0 \tag{7e}$$

where

$$\mathbf{x}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{x}(t,\epsilon) \tag{8a}$$

$$\mathbf{p}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{p}(t,\epsilon) \tag{8b}$$

$$\mathbf{u}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{u}(t,\epsilon) \tag{8c}$$

Equations (7a-e) form a linear TPBV problem with no delay or advance terms. Such a TPBV problem is encountered in the state regulator problem and its solution is well known [7.15].

The plan for obtaining a suboptimal control $\mathbf{u}^*(t)$ is as follows. Assuming that $\mathbf{u}(t,\epsilon)$, $\mathbf{x}(t,\epsilon)$ and $\mathbf{p}(t,\epsilon)$ are infinitely differentiable *w.r.t.* ϵ at $\epsilon = 0$, their MacLaurin series expansions in ϵ are

$$\mathbf{u}(t,\epsilon) = \sum_{i=0}^{\infty} \frac{\epsilon^i}{i!} \mathbf{u}^{(i)}(t) \tag{9}$$

$$\mathbf{x}(t,\epsilon) = \sum_{i=0}^{\infty} \frac{\epsilon^i}{i!} \mathbf{x}^{(i)}(t), \tag{10}$$

$$\mathbf{p}(t,\epsilon) = \sum_{i=0}^{\infty} \frac{\epsilon^i}{i!} \mathbf{p}^{(i)}(t), \tag{11}$$

where the superscript i denotes ith-order differentiation *w.r.t.* ϵ evaluated at $\epsilon = 0$. The optimal control $\mathbf{u}(t)$, $t_0 \leqslant t \leqslant t_f$, can be obtained by determining the coefficients $\mathbf{u}^{(i)}(t)$ and, assuming convergence, calculating the infinite sum in (9) for $\epsilon = 1$. An *Nth-order suboptimal control* $\mathbf{u}^*(t)$ can be obtained by truncating the infinite series in (9) with $\epsilon = 1$ to the Nth term. We will establish a recursive method to determine $\mathbf{u}^{(i)}(t)$, $i = 0,1,2,\ldots$, $t_0 \leqslant t \leqslant t_f$.

Substitution of (6d) in (6a) results in

$$\dot{\mathbf{x}}(t,\epsilon) = \mathbf{A}\mathbf{x}(t,\epsilon) + \epsilon\mathbf{A}_1\mathbf{x}(t-h,\epsilon) - \mathbf{S}\mathbf{p}(t,\epsilon), \quad t_0 \leqslant t \leqslant t_f \tag{12}$$

where

$$\mathbf{S} = \mathbf{B}\mathbf{R}^{-1}\mathbf{B}' \tag{13}$$

Differentiating (12), (6b-d) *w.r.t.* ϵ and letting $\epsilon \rightarrow 0$ yields

$$\dot{\mathbf{x}}^{(1)} = \mathbf{A}\mathbf{x}^{(1)} + \mathbf{A}_1\mathbf{x}^{(0)}(t-h) - \mathbf{S}\mathbf{p}^{(1)}, \quad t_0 \leqslant t \leqslant t_f \tag{14a}$$

$$\dot{\mathbf{p}}^{(1)} = -\mathbf{Q}\mathbf{x}^{(1)} - \mathbf{A}'\mathbf{p}^{(1)} - \mathbf{A}_1'\mathbf{p}^{(0)}(t+h), \; t_0 \leqslant t \leqslant t_f - h \tag{14b}$$

$$= -\mathbf{Q}\mathbf{x}^{(1)} - \mathbf{A}'\mathbf{p}^{(1)}, \; t_f - h \leqslant t \leqslant t_f \tag{14c}$$

$$\mathbf{u}^{(1)} = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{p}^{(1)} \tag{14d}$$

where dependence on t is dropped for convenience, with boundary conditions

$$\mathbf{x}^{(1)}(t_0) = 0 \tag{14e}$$

$$\mathbf{p}^{(1)}(t_f) = 0 \tag{14f}$$

The systems of equations (14) also represents a linear TPBV problem, with the difference that the delay and the advance terms appear not as dependent variables, but rather as forcing functions resulting from zeroth-order terms $\mathbf{x}^{(0)}$ and $\mathbf{p}^{(0)}$. The solution to the zeroth-order TPBV problem (7) is as follows [7.15]:

$$\dot{\mathbf{x}}^{(0)} = (\mathbf{A}-\mathbf{S}\mathbf{K})\mathbf{x}^{(0)}, \; \mathbf{x}^{(0)}(t_0) = \phi(t_0) \tag{15a}$$

$$\mathbf{p}^{(0)} = \mathbf{K}\mathbf{x}^{(0)} \tag{15b}$$

$$\mathbf{u}^{(0)} = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{K}\mathbf{x}^{(0)} \tag{15c}$$

where matrix $\mathbf{K}$ is the symmetric positive-definite solution of the Riccati equation

$$\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{S}\mathbf{K} + \mathbf{Q} = 0 \tag{16}$$

Therefore, the solution of (16) and substitution in (15) will result in the zeroth-order terms that are the forcing functions of the 1st-order TPBV problem (14).

The system of equations (14) can be decoupled by letting

$$\mathbf{p}^{(1)} = \mathbf{K}\mathbf{x}^{(1)} + \mathbf{g}_1 \tag{17}$$

where $\mathbf{g}_1$ is an *adjoint vector* obtained by substituting (17) and its time derivative in

(14) as follows:

$$\dot{g}_1(t) = -(A - SK)'g_1(t) - K\delta_1(t) + \sigma_1(t),\ t_0 \leqslant t \leqslant t_f - h \tag{18a}$$
$$= -(A - SK)'g_1(t) - K\delta_1(t),\ t_f - h \leqslant t \leqslant t_f \tag{18b}$$

where, by (17) and (14f),

$$g_1(t_f) = 0 \tag{18c}$$

and

$$\delta_1(t) = A_1x^{(0)}(t-h),\ \sigma_1(t) = -A'_1p^{(0)}(t+h) \tag{18d}$$

Thus (18) can be solved by integrating it backward in time and $g_1(t)$, $t_0 \leqslant t \leqslant t_f$, can be determined. Substitution of (17) in (14a) yields

$$\dot{x}^{(1)}(t) = (A - SK)x^{(1)}(t) - Sg_1 + A_1x^{(0)}(t-h) \tag{19}$$

Thus $x^{(1)}(t)$ can also be obtained from the above equation and boundary condition (14e). Therefore, using (14d) and (17), the coefficient $u^{(1)}(t)$ can be determined as follows:

$$u^{(1)} = -R^{-1}B'(Kx^{(1)} + g_1) \tag{20}$$

In a similar fashion, the *ith*-order terms can be obtained from the solution of the $(i-1)st$ set of equations. That is,

$$u^{(i)} = -R^{-1}B'(Kx^{(i)} + g_i) \tag{21}$$

where

$$\dot{x}^{(i)} = (A - SK)x^{(i)} - Sg_i + A_1x^{(i-1)}(t-h) \tag{22}$$
$$\dot{g}_i = -(A - SK)'g_i - K\delta_i + \sigma_i,\ t_0 \leqslant t \leqslant t_f - h \tag{23a}$$
$$= -(A - SK)'g_i - K\delta_i,\ t_f - h \leqslant t \leqslant t_f \tag{23b}$$

where

$$\delta_i = A_1x^{(i-1)}(t-h) \tag{23c}$$

and

$$\sigma_i = -\,A'_1 p^{(i-1)}(t+h) \tag{23d}$$

Thus from (9) and (10) we have

$$\begin{aligned} u(t,\epsilon) &= -\sum_{i=0}^{\infty} \frac{\epsilon^i}{i!} R^{-1}B'[Kx^{(i)}(t) + g_i(t)] \\ &= -\,R^{-1}B'[Kx(t,\epsilon) + \sum_{i=1}^{\infty} \frac{\epsilon^i}{i!} g_i(t)] \end{aligned} \tag{24}$$

Note that the state feedback portion of the control in (24) is exact, *i.e.* the first term contains, in effect, an infinite number of control sensitivity vectors. An Nth-order suboptimal control is obtained by truncating the infinite series in (24) to N terms and setting $\epsilon=1$:

$$u^*(t) = -\,R^{-1}B'[Kx(t) + \sum_{i=1}^{N} \frac{g_i(t)}{i!}] \tag{25}$$

Figure 1 shows the block diagram of the proposed suboptimal controller.

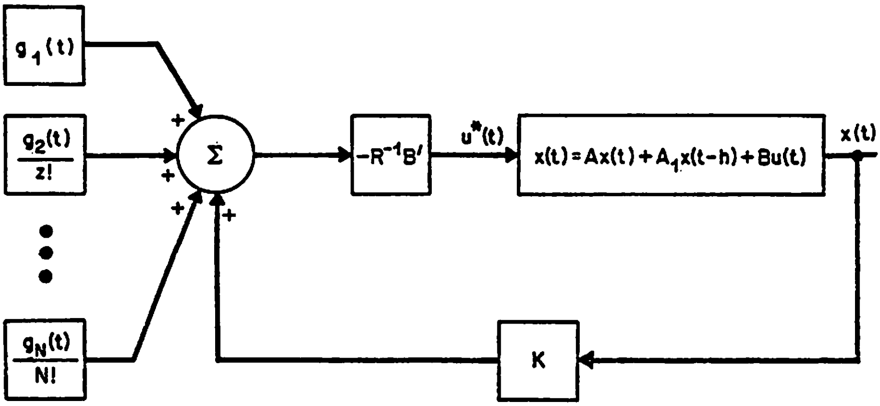

Figure 1. Block Diagram of the Suboptimally Controlled System

A summary of the computational procedure follows:

Step 1: Solve the algebraic matrix Riccati equation (16).

Step 2: Obtain the zeroth-order terms $x^{(0)}$, $p^{(0)}$ and $u^{(0)}$ from (15). Set $i=1$

Step 3: Using the $(i-1)st$ order terms, solve the ith adjoint vector equation (23).

Step 4: Obtain $x^{(i)}$ and $u^{(i)}$ from (22) and (21).

Step 5: If $i < N$, let $i = i + 1$ and go to Step 3; otherwise stop.

One has to prove that the infinite sum in (24) always converges. Such convergence can be easily proved for small values of ϵ. It has not been possible to show convergence for the case of $\epsilon = 1$, in general. However, we have not encountered any case where this series diverged.

1. Example[7.1]. We will apply the proposed method to the first-order system

$$\dot{x} = x + x(t-1) + u, \quad 0 \leqslant t \leqslant 2 \tag{26a}$$

$$x(t) = 1, \ -1 \leqslant t \leqslant 0 \tag{26b}$$

$$J = \int_0^2 (x^2 + u^2)dt \tag{27}$$

The Riccati equation (16) reduces to the following algebraic equation:

$$2K - K^2/2 + 2 = 0 \tag{28}$$

whose positive solution is $K = 2(1 + \sqrt{2})$. The zeroth-order equations are

$$\dot{x}^{(0)} = (A - SK)x^{(0)} = -(\sqrt{2})x^{(0)}, \ x^{(0)}(t) = e^{-\sqrt{2}t} \tag{29a}$$

$$p^{(0)} = Kx^{(0)} = 2(1 + \sqrt{2})e^{-\sqrt{2}t} \tag{29b}$$

$$u^{(0)} = -R^{-1}B'Kx^{(0)} = -(1 + \sqrt{2})e^{-\sqrt{2}t} \tag{29c}$$

This completes the first two steps of the procedure. The differential equation for the first order adjoint vector g_1 is

$$\dot{g}_1 = \sqrt{2}g_1 - 2(1 + \sqrt{2})(e^{\sqrt{2}} + e^{-\sqrt{2}})e^{-\sqrt{2}t}, \ 0 \leqslant t \leqslant 1 \tag{30a}$$

$$= \sqrt{2}g_1 - 2(1 + \sqrt{2})e^{\sqrt{2}}e^{-\sqrt{2}t}, \ 1 \leqslant t \leqslant 2 \tag{30b}$$

whose solution is

$$g_1(t) = c_1e^{\sqrt{2}t} - (1 + 1/\sqrt{2})(e^{\sqrt{2}} + e^{-\sqrt{2}})e^{-\sqrt{2}t}, \ 0 \leqslant t \leqslant 1 \tag{31a}$$

$$= c_2e^{\sqrt{2}t} - (1 + 1/\sqrt{2})e^{\sqrt{2}}e^{-\sqrt{2}t}, \ 1 \leqslant t \leqslant 2 \tag{31b}$$

where constants c_1 and c_2 can be determined by using boundary conditions at $t = 2$ and $t = 1$:

$$c_1 = (2 + \sqrt{2})e^{-3\sqrt{2}}, \ c_2 = (1 + (\sqrt{2})/2)e^{-3\sqrt{2}} \tag{31c}$$

For $N = 1$, the first-order suboptimal control, using (25), becomes

$$u^*(t) = -[(1+\sqrt{2})x(t) + g_1(t)] \tag{32}$$

The state equation (26a) can then be rewritten as

$$\dot{x}(t) = -\sqrt{2}x(t) + x(t-1) - g_1(t) \tag{33}$$

The resulting state, control and cost functional are depicted in Figure 2. These results compare very closely with those obtained by Eller, *el al.*[7.16] for the same example.

7.2.2 The Linear Time-Varying Case

Consider the nonstationary linear TD system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h_x) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{B}_1(t)\mathbf{u}(t-h_u),\ t \geqslant t_0, \tag{34a}$$

$$\mathbf{x}(t) = \phi(t),\ t_0 - h_x \leqslant t \leqslant t_0, \tag{34b}$$

$$\mathbf{u}(t) = \eta(t),\ t_0 - h_u \leqslant t \leqslant t_0, \tag{34c}$$

where $\mathbf{A}(t)$, $\mathbf{A}_1(t)$, $\mathbf{B}(t)$ and $\mathbf{B}_1(t)$ are real, piecewise continuous matrices of appropriate dimensions defined on the appropriate intervals, t_0 is the initial process time, $\phi(t)$ and $\eta(t)$ are specified initial functions, and h_x and h_u are constant positive scalars where, for convenience, it is assumed that

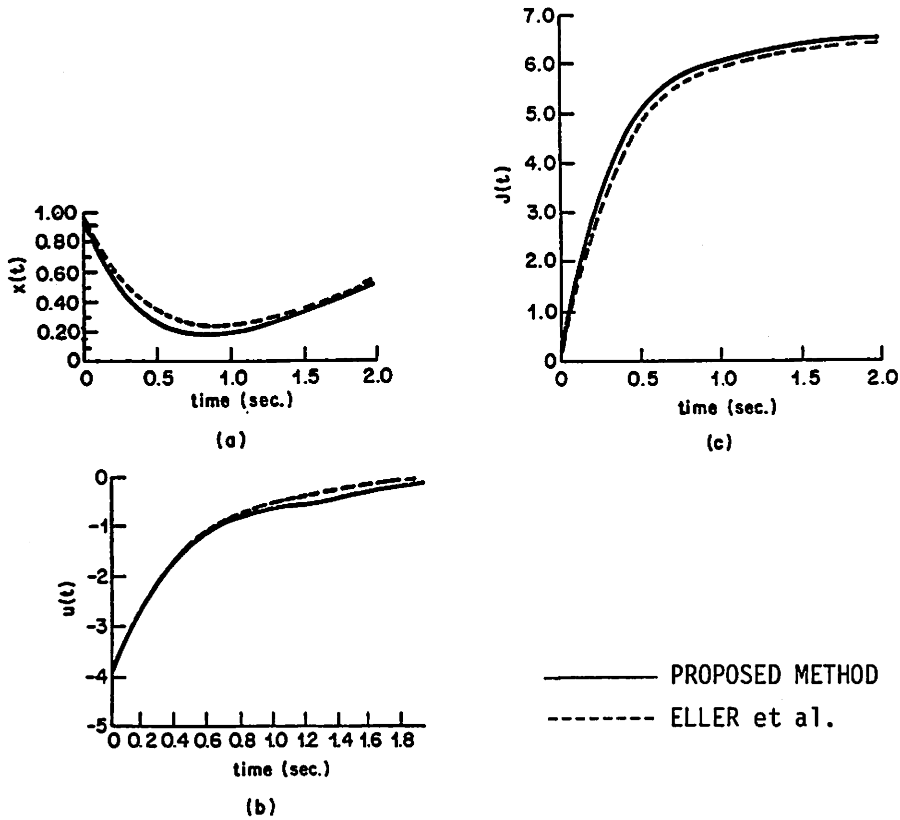

Figure 2. Suboptimal State, Control and Cost Functional in Example 1
— proposed method, - - - -Eller *et al.* [7.16]

$$h_u \leqslant h_x \tag{34d}$$

The cost functional, to be minimized, is

$$J = \tfrac{1}{2}\mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_0}^{t_f}[\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t)]dt, \tag{35}$$

where matrix $\mathbf{F}$ is symmetric positive semidefinite, matrix $\mathbf{Q}(t)$ is symmetric positive semidefinite and piecewise continuous for $t_0 \leqslant t \leqslant t_f$, and matrix $\mathbf{R}(t)$ is symmetric positive definite and piecewise continuous for $t_0 \leqslant t \leqslant t_f$. The problem is to find a control $\mathbf{u}(t)$, $t_0 \leqslant t \leqslant t_f$, which for fixed final time t_f and free final state $\mathbf{x}(t_f)$ will minimize the cost functional J in (35).

The sensitivity approach to solve the above problem is an extension of the approach used for the *l.t.i.* case. The Hamiltonian function for the problem is

$$H(\mathbf{x},\mathbf{p},\mathbf{u}) = \tfrac{1}{2}[\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t)] + \mathbf{p}'(t)[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{A}_1(t)\mathbf{x}(t-h_x) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{B}_1(t)\mathbf{u}(t-h_u)], \tag{36}$$

where $\mathbf{p}(t)$ is the costate vector. The necessary and sufficient conditions for optimality are (see Section 6.3):

$$\dot{\mathbf{p}}(t) = -H_x - H_y(t+h_x),\ t_0 \leqslant t \leqslant t_f - h_x, \tag{37a}$$
$$\dot{\mathbf{p}}(t) = -H_x,\ t_f - h_x \leqslant t \leqslant t_f, \tag{37b}$$
$$\mathbf{p}(t_f) = \mathbf{F}\mathbf{x}(t_f),\ \mathbf{p}(t_f - h_x^+) = \mathbf{p}(t_f - h_x^-), \tag{37c}$$
$$0 = H_u + H_v(t+h_u),\ t_0 \leqslant t \leqslant t_f, \tag{37d}$$
$$0 = H_u,\ t_f - h_u \leqslant t \leqslant t_f, \tag{37e}$$

and (34), where

$$\mathbf{y} \triangleq \mathbf{x}(t-h_x),\ \mathbf{v} \triangleq \mathbf{u}(t-h_u), \tag{37f}$$

and the subscripts denote gradients. Calculating the gradients, we obtain

$$\dot{\mathbf{p}}(t) = -\mathbf{Q}(t)\mathbf{x}(t) - \mathbf{A}'(t)\mathbf{p}(t) - \mathbf{A}_1'(t+h_x)\mathbf{p}(t+h_x),\ t_0 \leqslant t \leqslant t_f - h_x, \tag{38a}$$
$$\dot{\mathbf{p}}(t) = -\mathbf{Q}(t)\mathbf{x}(t) - \mathbf{A}'(t)\mathbf{p}(t),\ t_f - h_x \leqslant t \leqslant t_f, \tag{38b}$$
$$\mathbf{p}(t_f) = \mathbf{F}\mathbf{x}(t_f),\ \mathbf{p}(t_f - h_x^+) = \mathbf{p}(t_f - h_x^-), \tag{38c}$$
$$0 = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}'(t)\mathbf{p}(t) + \mathbf{B}_1'(t+h_u)\mathbf{p}(t+h_u),\ t_0 \leqslant t \leqslant t_f - h_u, \tag{38d}$$
$$0 = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}'(t)\mathbf{p}(t),\ t_f - h_u \leqslant t \leqslant t_f \tag{38e}$$

Equations (34) and (38) form a linear TPBV problem with time-varying coefficients involving both delay and advance terms. The exact solution of this problem, and therefore, determining the exact optimal control $\mathbf{u}(t)$ for $t_0 \leqslant t \leqslant t_f$ is, in general, extremely difficult, if not impossible.

Now consider the family of TPBV problems

$$\dot{\mathbf{x}}(t,\epsilon) = \mathbf{A}(t)\mathbf{x}(t,\epsilon) + \epsilon\mathbf{A}_1(t)\mathbf{x}(t - h_x,\epsilon) + \mathbf{B}(t)\mathbf{u}(t,\epsilon)_1(t \tag{39a}$$

$$\dot{\mathbf{p}}(t,\epsilon) = -\mathbf{Q}(t)\mathbf{x}(t,\epsilon) - \mathbf{A}'(t)\mathbf{p}(t,\epsilon)$$
$$- \epsilon\mathbf{A}_1'(t + h_x)\mathbf{p}(t + h_x,\epsilon), \quad t_0 \leqslant t \leqslant t_f - h_x, \tag{39b}$$

$$\dot{\mathbf{p}}(t,\epsilon) = -\mathbf{Q}(t)\mathbf{x}(t,\epsilon) - \mathbf{A}'(t)\mathbf{p}(t,\epsilon), \quad t_f - h_x \leqslant t \leqslant t_f, \tag{39c}$$

$$\mathbf{u}(t,\epsilon) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}(t,\epsilon)$$
$$- \epsilon\mathbf{R}^{-1}(t)\mathbf{B}_1'(t + h_u)\mathbf{p}(t + h_u,\epsilon), \quad t_0 \leqslant t \leqslant t_f - h_u, \tag{39d}$$

$$\mathbf{u}(t,\epsilon) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}(t,\epsilon), \quad t_f - h_u \leqslant t \leqslant t_f, \tag{39e}$$

with boundary conditions

$$\mathbf{x}(t,\epsilon) = \phi(t), \quad t_0 - h_x \leqslant t \leqslant t_0, \tag{39f}$$

$$\mathbf{p}(t_f,\epsilon) = \mathbf{F}\mathbf{x}(t_f,\epsilon), \quad \mathbf{p}(t_f - h_x^+,\epsilon) = \mathbf{p}(t_f - h_x^-,\epsilon), \tag{39g}$$

$$\mathbf{u}(t,\epsilon) = \eta(t), \quad t_0 - h_u \leqslant t \leqslant t_0, \tag{39h}$$

where $0 \leqslant \epsilon \leqslant 1$ is a real scalar parameter. It is assumed that the solution to the above TPBV problem exists for every value of ϵ and is unique. Note that, for $\epsilon = 1$, this TPBV problem is identical to the original one. For $\epsilon = 0$, it becomes

$$\dot{\mathbf{x}}^{(0)}(t) = \mathbf{A}(t)\mathbf{x}^{(0)}(t) + \mathbf{B}(t)\mathbf{u}^{(0)}(t), \tag{40a}$$

$$\dot{\mathbf{p}}^{(0)}(t) = -\mathbf{Q}(t)\mathbf{x}^{(0)}(t) - \mathbf{A}'(t)\mathbf{p}^{(0)}(t), \tag{40b}$$

$$\mathbf{u}^{(0)}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}^{(0)}(t), \tag{40c}$$

for $t_0 \leqslant t \leqslant t_f$, with boundary conditions

$$\mathbf{x}^{(0)}(t_0) = \phi(t_0), \tag{40d}$$

$$\mathbf{p}^{(0)}(t_f) = \mathbf{F}\mathbf{x}^{(0)}(t_f), \tag{40e}$$

where

$$\mathbf{x}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{x}(t,\epsilon), \tag{41a}$$

$$\mathbf{p}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{p}(t,\epsilon), \tag{41b}$$

$$\mathbf{u}^{(0)}(t) \triangleq \lim_{\epsilon \to 0} \mathbf{u}(t,\epsilon) \tag{41c}$$

Equation (40) forms a standard linear TPBV problem without delay and advance terms which is encountered in the state regulator problem. Again, the plan for obtaining the optimal control is to expand $\mathbf{u}(t,\epsilon)$ in MacLaurin series and to determine the coefficients from nondelay system calculations. An Nth-order suboptimal control $\mathbf{u}^*(t)$ is obtained by truncating the series to the Nth term. More specifically, assume that $\mathbf{u}(t,\epsilon)$, $\mathbf{x}(t,\epsilon)$ and $\mathbf{p}(t,\epsilon)$ are infinitely differentiable *w.r.t.* ϵ at $\epsilon = 0$. Then these functions can be expanded as given in (9), (10) and (11). In sequel, a recursive method, based on nondelay system optimization, will be established to determine $\mathbf{u}^{(i)}(t)$, $i = 0,1,2,\ldots,N$ for $t_0 \leqslant t \leqslant t_f$.

Substitution of (39d-e) in (39a) results in

$$\dot{\mathbf{x}}(t,\epsilon) = \mathbf{A}(t)\mathbf{x}(t,\epsilon) + \epsilon\mathbf{A}_1(t)\mathbf{x}(t-h_x,\epsilon) - [\mathbf{S}_1(t) + \epsilon^2\mathbf{S}_2(t)]\mathbf{p}(t,\epsilon) - \epsilon\mathbf{S}_3(t)\mathbf{p}(t+h_u,\epsilon) - \epsilon\mathbf{S}_4(t)\mathbf{p}(t-h_u,\epsilon),\quad t_0 \leqslant t \leqslant t_f - h_u, \tag{42a}$$

$$\dot{\mathbf{x}}(t,\epsilon) = \mathbf{A}(t)\mathbf{x}(t,\epsilon) + \epsilon\mathbf{A}_1(t)\mathbf{x}(t-h_x,\epsilon) - \mathbf{S}_1(t)\mathbf{p}(t,\epsilon) - \epsilon\mathbf{S}_4(t)\mathbf{p}(t-h_u,\epsilon),\quad t_f - h_u \leqslant t \leqslant t_f, \tag{42b}$$

where

$$\mathbf{S}_1(t) = \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}'(t), \tag{43}$$

$$\mathbf{S}_2(t) = \mathbf{B}_1(t)\mathbf{R}^{-1}(t-h_u)\mathbf{B}_1'(t), \tag{44}$$

$$\mathbf{S}_3(t) = \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}_1'(t+h_u), \tag{45}$$

$$\mathbf{S}_4(t) = \mathbf{B}_1(t)\mathbf{R}^{-1}(t-h_u)\mathbf{B}'(t-h_u). \tag{46}$$

Differentiating (42) and (39b-e) i times, $i = 1,2,\ldots,$ *w.r.t.* ϵ and letting $\epsilon \to 0$ yields

$$\dot{\mathbf{x}}^{(i)}(t) = \mathbf{A}(t)\mathbf{x}^{(i)}(t) - \mathbf{S}_1(t)\mathbf{p}^{(i)}(t) + \mathbf{A}_1(t)\mathbf{x}^{(i-1)}(t-h_x) - \mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) - \mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u),\quad t_0 \leqslant t \leqslant t_f - h_u, \tag{47a}$$

$$\dot{\mathbf{x}}^{(i)}(t) = \mathbf{A}(t)\mathbf{x}^{(i)}(t) - \mathbf{S}_1(t)\mathbf{p}^{(i)}(t) + \mathbf{A}_1(t)\mathbf{x}^{(i-1)}(t-h_x) - \mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u),\quad t_f - h_u \leqslant t \leqslant t_f, \tag{47b}$$

$$\dot{\mathbf{p}}^{(i)}(t) = -\mathbf{Q}(t)\mathbf{x}^{(i)}(t) - \mathbf{A}'(t)\mathbf{p}^{(i)}(t)$$
$$- \mathbf{A}_1'(t + h_x)\mathbf{p}^{(i-1)}(t + h_x), \ t_0 \leqslant t \leqslant t_f - h_x, \tag{47c}$$
$$\dot{\mathbf{p}}^{(i)}(t) = -\mathbf{Q}(t)\mathbf{x}^{(i)}(t) - \mathbf{A}'(t)\mathbf{p}^{(i)}(t), \ t_f - h_x \leqslant t \leqslant t_f, \tag{47d}$$
$$\mathbf{u}^{(i)}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}^{(i)}(t)$$
$$- \mathbf{R}^{-1}(t)\mathbf{B}_1'(t + h_u)\mathbf{p}^{(i-1)}(t + h_u), \ t_0 \leqslant t \leqslant t_f - h_u, \tag{47e}$$
$$\mathbf{u}^{(i)}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}^{(i)}(t), \ t_f - h_u \leqslant t \leqslant t_f, \tag{47f}$$

with boundary conditions

$$\mathbf{x}^{(i)}(t) = 0, \ t_0 - h_x \leqslant t \leqslant t_0, \tag{47g}$$
$$\mathbf{p}^{(i)}(t_f) = \mathbf{F}\mathbf{x}^{(i)}(t_f), \tag{47h}$$
$$\mathbf{u}^{(i)}(t) = 0, \ t_0 - h_u \leqslant t \leqslant t_0. \tag{47i}$$

The appearance of $\mathbf{p}^{(i-1)}(t - h_u)$ in (47a) requires the knowledge of $\mathbf{p}^{(i-1)}(t)$ for $t_0 - h_u \leqslant t \leqslant t_0$. This can be determined by the analytic extension of (40c) and (47e) to $t_0 - h_u$. Thus, (40c) and (34c) yield

$$\mathbf{B}'(t)\mathbf{p}^{(0)}(t) = -\mathbf{R}(t)\eta(t), \ t_0 - h_u \leqslant t \leqslant t_0, \tag{48}$$

and (47e) and (47i) yield

$$\mathbf{B}'(t)\mathbf{p}^{(i)}(t) = -\mathbf{B}_1'(t + h_u)\mathbf{p}^{(i-1)}(t + h_u), \ t_0 - h_u \leqslant t \leqslant t_0, \ i = 1,2,.... \tag{49}$$

From (46), (48), (49) and (44) one obtains, for $t_0 \leqslant t \leqslant t_0 + h_u$,

$$-\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t - h_u) = \mathbf{B}_1(t)\eta(t - h_u), \ i = 1, \tag{50a}$$
$$-\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t - h_u) = \mathbf{S}_2(t)\mathbf{p}^{(i-2)}(t), \ i = 2,3,.... \tag{50b}$$

Therefore (47a,b) become

$$\dot{\mathbf{x}}^{(i)}(t) = \mathbf{A}(t)\mathbf{x}^{(i)}(t) - \mathbf{S}_1(t)\mathbf{p}^{(i)}(t) + \mathbf{A}_1(t)\mathbf{x}^{(i-1)}(t - h_x) + \mathbf{F}_i(t), \tag{51a}$$

where

$$
\begin{aligned}
\mathbf{F}_i(t) &= -\mathbf{S}_3(t)\mathbf{p}^{(0)}(t+h_u) + \mathbf{B}_1(t)\eta(t-h_u),\ t_0 \leqslant t \leqslant t_0+h_u,\ i=1, && \text{(51b)}\\
&= -\mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) + \mathbf{S}_2(t)\mathbf{p}^{(i-2)}(t),\ t_0 \leqslant t \leqslant t_0+h_u,\\
&\qquad i=2,3,\ldots, && \text{(51c)}\\
&= -\mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) - \mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u),\ t_0+h_u \leqslant t \leqslant t_f-h_u,\\
&\qquad i=1,2,\ldots, && \text{(51d)}\\
&= -\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u),\ t_f-h_u \leqslant t \leqslant t_f,\\
&\qquad i=1,2,\ldots, && \text{(51e)}
\end{aligned}
$$

and $\mathbf{S}_1(t)$, $\mathbf{S}_2(t)$, $\mathbf{S}_3(t)$ and $\mathbf{S}_4(t)$ are given by (43) to (46). Note that $\mathbf{F}_i(t)$ involves only values from $(i-1)st$ calculations.

The unique solution to the zeroth-order TPBV problem (40) is [7.15]

$$\dot{\mathbf{x}}^{(0)}(t) = [\mathbf{A}(t) - \mathbf{S}_1(t)\mathbf{K}(t)]\mathbf{x}^{(0)}(t),\ \ \mathbf{x}^{(0)}(t_0) = \phi(t_0), \tag{52a}$$

$$\mathbf{p}^{(0)}(t) = \mathbf{K}(t)\mathbf{x}^{(0)}(t), \tag{52b}$$

$$\mathbf{u}^{(0)}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{p}^{(0)}(t), \tag{52c}$$

for $t_0 \leqslant t \leqslant t_f$, where $\mathbf{S}_1(t)$ is given by (43) and $\mathbf{K}(t)$ is the symmetric positive-definite solution of the differential matrix Riccati equation

$$\dot{\mathbf{K}}(t) + \mathbf{A}'(t)\mathbf{K}(t) + \mathbf{K}(t)\mathbf{A}(t) - \mathbf{K}(t)\mathbf{S}_1(t)\mathbf{K}(t) + \mathbf{Q}(t) = 0, \tag{53a}$$

with boundary condition

$$\mathbf{K}(t_f) = \mathbf{F} \tag{53b}$$

Differential equations (47c,d) and (51a) describing the state vector $\mathbf{x}(.)$ and the costate vector $\mathbf{p}(.)$ can be uncoupled by defining adjoint vectors $\mathbf{g}_i(t)$, $i = 0,1,2,\ldots$, as follows:

$$\mathbf{p}^{(i)}(t) = \mathbf{K}(t)\mathbf{x}^{(i)}(t) + \mathbf{g}_i(t),\ t_0 \leqslant t \leqslant t_f,\ i = 0,1,2,\ldots. \tag{54}$$

Equation (52b) implies that

$$\mathbf{g}_0(t) = \mathbf{0},\ t_0 \leqslant t \leqslant t_f \tag{55}$$

Also (47h) and (53b) imply that

$$\mathbf{g}_i(t_f) = \mathbf{0},\ i = 0,1,2,\ldots. \tag{56}$$

Substituting (54) in (51a) yields

$$\dot{\mathbf{x}}^{(i)}(t) = [\mathbf{A}(t) - \mathbf{S}_1(t)\mathbf{K}(t)]\mathbf{x}^{(i)}(t) - \mathbf{S}_1(t)\mathbf{g}_i(t) + \mathbf{A}_1(t)\mathbf{x}^{(i-1)}(t-h_x) + \mathbf{F}_i(t), \tag{57a}$$

$$\mathbf{x}^{(i)}(t_0) = 0 \tag{57b}$$

where $\mathbf{F}_i(t)$ is given by (51b-e).

In order to obtain $\mathbf{g}_i(t), i = 1,2,\ldots,$ for $t_0 \leqslant t \leqslant t_f$, differentiate (54) *w.r.t.* time and use (47c,d), (57), (53a) and (54) to obtain

$$\dot{\mathbf{g}}_i(t) = -[\mathbf{A}(t) - \mathbf{S}_1(t)\mathbf{K}(t)]'\mathbf{g}_i(t) - \mathbf{K}(t)\mathbf{A}_1(t)\mathbf{x}^{(i-1)}(t-h_x) + \mathbf{G}_i(t), \tag{58a}$$

with boundary condition (56), where

$$\begin{aligned}
\mathbf{G}_i(t) &= -\mathbf{A}_1'(t+h_x)\mathbf{p}^{(0)}(t+h_x) + \mathbf{K}(t)\mathbf{S}_3(t)\mathbf{p}^{(0)}(t+h_u) \\
&\quad - \mathbf{K}(t)\mathbf{B}_1(t)\eta(t-h_u), \quad t_0 \leqslant t \leqslant t_0 + h_u, \; i = 1, &(58\text{b})\\
&= -\mathbf{A}_1'(t+h_x)\mathbf{p}^{(i-1)}(t+h_x) + \mathbf{K}(t)\mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) \\
&\quad - \mathbf{K}(t)\mathbf{S}_2(t)\mathbf{p}^{(i-2)}(t), \quad t_0 \leqslant t \leqslant t_0 + h_u, \; i = 2,3,\ldots, &(58\text{c})\\
&= -\mathbf{A}_1'(t+h_x)\mathbf{p}^{(i-1)}(t+h_x) + \mathbf{K}(t)\mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) \\
&\quad + \mathbf{K}(t)\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u), \quad t_0 + h_u \leqslant t \leqslant t_f - h_x, \; i = 1,2,\ldots, &(58\text{d})\\
&= \mathbf{K}(t)\mathbf{S}_3(t)\mathbf{p}^{(i-1)}(t+h_u) \\
&\quad + \mathbf{K}(t)\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u), \quad t_f - h_x \leqslant t \leqslant t_f - h_u, \; i = 1,2,\ldots, &(58\text{e})\\
&= \mathbf{K}(t)\mathbf{S}_4(t)\mathbf{p}^{(i-1)}(t-h_u), \quad t_f - h_u \leqslant t \leqslant t_f, i = 1,2,\ldots.. &(58\text{f})
\end{aligned}$$

Note that the homogeneous parts of (57a) and (58a) are adjoint. Also note that the forcing function of (58a) for each i is obtained exclusively from $(i-1)st$ calculations and known functions.

Now we are ready to formulate the suboptimal control. Substitution of (47e,f) in (9) and noting (52c) yields

$$\begin{aligned}
\mathbf{u}(t,\epsilon) &= -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\sum_{i=0}^{\infty}(\epsilon^i/i!)\mathbf{p}^{(i)}(t) \\
&\quad - \mathbf{R}^{-1}(t)\mathbf{B}_1'(t+h_u)\sum_{i=0}^{\infty}(\epsilon^i/i!)\mathbf{p}^{(i-1)}(t+h_u), \quad t_0 \leqslant t \leqslant t_f - h_u, &(59\text{a})\\
\mathbf{u}(t,\epsilon) &= -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\sum_{i=0}^{\infty}(\epsilon^i/i!)\mathbf{p}^{(i)}(t), \quad t_f - h_u \leqslant t \leqslant t_f &(59\text{b})
\end{aligned}$$

Using (54) and (10), (59) becomes

$$\mathbf{u}(t,\epsilon) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)[\mathbf{K}(t)\mathbf{x}(t,\epsilon) + \sum_{i=0}^{\infty}(\epsilon^i/i!)\mathbf{g}_i(t)]$$
$$- \mathbf{R}^{-1}(t)\mathbf{B}_1'(t + h_u)[\mathbf{K}(t + h_u)\sum_{i=1}^{\infty}(\epsilon^i/i!)\mathbf{x}^{(i-1)}(t + h_u)$$
$$+ \sum_{i=1}^{\infty}(\epsilon^i/i!)\mathbf{g}_{i-1}(t + h_u)], \quad t_0 \leqslant t \leqslant t_f - h_u, \tag{60a}$$

$$\mathbf{u}(t,\epsilon) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)[\mathbf{K}(t)\mathbf{x}(t,\epsilon) + \sum_{i=0}^{\infty}(\epsilon^i/i!)g_i(t)], \quad t_f - h_u \leqslant t \leqslant t_f \tag{60b}$$

For $\epsilon = 1$, $\mathbf{u}(t,\epsilon)$ given in (60) is the exact optimal control provided that the infinite sums converge. An Nth-order suboptimal control $\mathbf{u}^*(t)$ can be obtained by truncating the infinite series in (60) with $\epsilon = 1$ to the Nth term:

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)[\mathbf{K}(t)\mathbf{x}(t) + \mathbf{U}_n(t)]$$
$$- \mathbf{R}^{-1}(t)\mathbf{B}_1'(t + h_u)[\mathbf{K}(t + h_u)\mathbf{X}_n(t + h_u) + \mathbf{V}_n(t + h_u)],$$
$$t_0 \leqslant t \leqslant t_f - h_u, \tag{61a}$$

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)[\mathbf{K}(t)\mathbf{x}(t) + \mathbf{U}_n(t)], \quad t_f - h_u \leqslant t \leqslant t_f \tag{61b}$$

where

$$\mathbf{X}_n(t) = \sum_{i=1}^{N}[\mathbf{x}^{(i-1)}(t)/i!], \tag{61c}$$

and, by (55),

$$\mathbf{U}_n(t) = \sum_{i=1}^{N}\mathbf{g}_i(t)/i!, \quad \mathbf{V}_n(t) = \sum_{i=2}^{N}\mathbf{g}_i(t)/i! \,. \tag{61d}$$

Note that the above suboptimal control includes an exact feedback term.

The computational procedure to determine the suboptimal control (61) can be summarized as follows:

Step 1: Solve the Riccati differential equation (53).

Step 2: Solve (52) to obtain $\mathbf{x}^{(0)}(t)$, $\mathbf{p}^{(0)}(t)$ and $\mathbf{u}^{(0)}(t)$. Set $i = 1$.

Step 3: Calculate the forcing function of (58a) for $i = 1$.

Step 4: Solve (58) to obtain adjoint vector $\mathbf{g}_i(t)$.

Step 5: Calculate the forcing function of (57).

Step 6: Solve (57) to obtain $\mathbf{x}^{(i)}(t)$.

Step 7: Calculate $\mathbf{p}^{(i)}(t)$ from (54).

Step 8: If $i < N$, let $i = i + 1$ and go to Step 3; otherwise go to Step 9.

Step 9: Calculate $\mathbf{U}_n(t)$, $\mathbf{V}_n(t)$ and $\mathbf{X}_n(t)$ from (61c,d). Calculate suboptimal control $\mathbf{u}^*(t)$ from (61a,b). Stop.

Various methods of solving the differential and algebraic matrix Riccati equations have been surveyed in Reference [7.18].

Several extensions to the above procedure are also possible. One such extension is to systems with multiple delays in the state and the control. Consider a system with state delays h_x, $i = 1,2,\ldots,N$, and control delays h_u, $i = 1,2,\ldots,R$. Then, (37a,b) should be replaced by

$$\dot{\mathbf{p}}(t) = -H_{x_i} - H_{y_i}(t + h_{x_i}), \; t_0 \leqslant t \leqslant t_f - h_{x_i}, \tag{62a}$$

$$\dot{\mathbf{p}}(t) = -H_{x_i}, \; t_f - h_{x_i} \leqslant t \leqslant t_f \tag{62b}$$

for $i = 1,2,\ldots,j$, and (37d,e) should be replaced by

$$0 = H_{u_i} + H_{u_i}(t + h_{u_i}), \; t_0 \leqslant t \leqslant t_f - h_{u_i} \tag{63a}$$

$$0 = H_{u_i}, \; t_f - h_{u_i} \leqslant t \leqslant t_f \tag{63b}$$

for $i = 1,2,\ldots,R$ where

$$\mathbf{y}_i = \mathbf{x}(t - h_{x_i}), \text{ and } \mathbf{v}_i = \mathbf{u}(t - h_{u_i}). \tag{63c}$$

Then, (57), (58) and (61) must be modified accordingly.

Another extension to the procedure is not requiring the final state, $\mathbf{x}(t_f)$, to be free. The final state may be specified by introducing a penalty term in the form of an appropriate large terminal cost in the cost functional (35).

2. Example[7.3]. Consider the second-order linear time-varying multi-delay system expressed by a state equation similar to (34) where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix}, \mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{B}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \phi(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$h_{x_1} = 1, \; h_{x_2} = 0.8, \; h_u = 0.5, \; \eta(t) = 5(t+1), \; t_0 = 0, \tag{64}$$

with the cost functional (35) where

$$\mathbf{F} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{R} = 1/(t+2), \; t_f = 3 \tag{65}$$

Assume that the final state is free. Thus, the state equations are

$$\dot{x}_1 = x_2(t) + x_1(t-1), \; t \geqslant 0, \tag{66a}$$

$$\dot{x}_2 = t x_1(t) + 2x_1(t-1) + x_2(t-1) + u(t) - u(t-0.5), \; t \geqslant 0, \tag{66b}$$

$$x_1(t) = 1, \quad -1 \leqslant t \leqslant 0, \; x_2(t) = 1, \quad -0.8 \leqslant t \leqslant 0, \tag{66c}$$

$$u(t) = 5(t+1), \quad -0.5 \leqslant t \leqslant 0, \tag{66d}$$

and the cost functional is

$$J = \tfrac{1}{2}x_1^2(3) + x_2^2(3) + \tfrac{1}{2}\int_0^3 [2x_1^2(t) + 2x_1(t)x_2(t) + x_2^2(t) + u^2(t)/(t+2)]dt \tag{67}$$

All differential equations were solved by the finite- difference method using a step size $\Delta t = 0.05$. Variations of $x_1(t), x_2(t)$ and $u(t)$ for $N = 2$ and $N = 10$ are plotted in Figure 3. The value of the cost functional corresponding to each N is shown in Table 1.

Table 1. Results for Example 2.

N	J
2	118.3228000
3	43.5914154
4	28.3378601
5	25.0539551
6	24.2544098
7	24.0693359
8	24.0292816
9	24.0214233
10	24.0200500

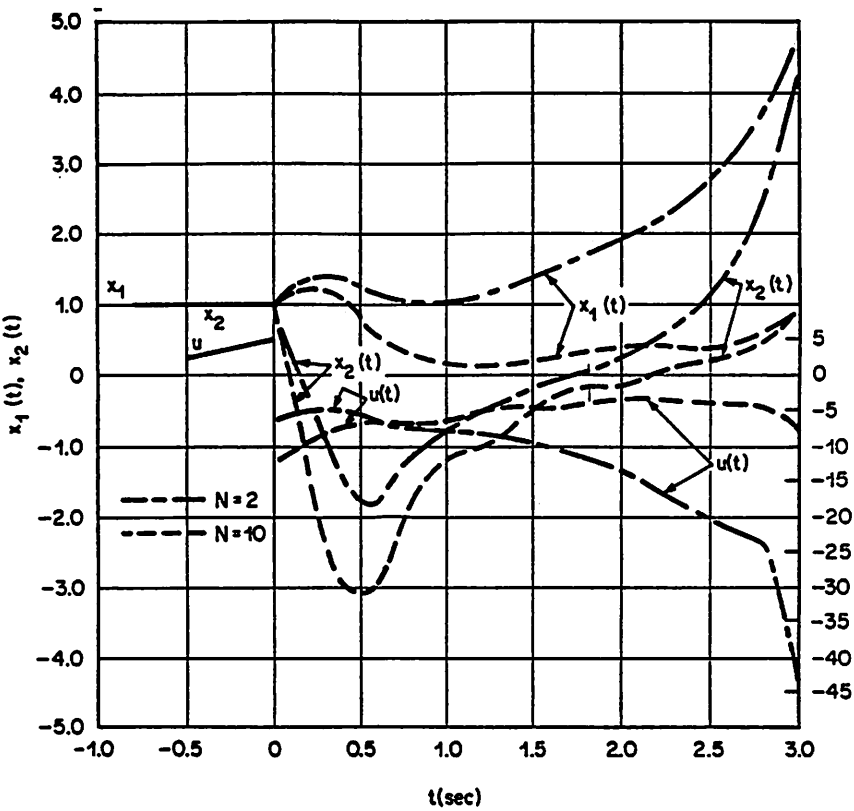

Figure 3. Variations of the State and the Control in Example 2

7.3 NONDELAY CONVERSION APPROACH

The approach to be presented in this section is based on treating the delay terms in the state equation of the system as extra perturbing inputs. Thus, the TD problem will be converted to a nondelay problem. This approach was first used by Gracovetsky and Vidyasagar [7.4] for systems with a single state delay. The method presented in this section is based on reference [7.5]. It is a fast-converging iterative method to obtain suboptimal control for linear systems with multiple state and control delays and with quadratic cost. In this method, the delay terms are treated as extra inputs to linear nondelay systems which must be optimized in each iteration. The suboptimal control obtained by the method is partly closed-loop and partly open-loop. The procedure can be extended to the case with time-varying delays. It can also be extended to the case where the delay terms are nonlinear, provided that some of the nonlinearities satisfy the Lipschitz condition.

7.3.1 Linear Systems with Multiple Constant Delays

Consider the linear multiple-delay system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{A}_i(t)\mathbf{x}(t-h_{xi}) + \mathbf{B}(t)\mathbf{u}(t)$$

$$+ \sum_{i=1}^{R}\mathbf{B}_i(t)\mathbf{u}(t-h_{ui}),\ t \geqslant t_0, \tag{1a}$$

$$\mathbf{x}(t) = \phi(t), \qquad t_o - \Delta_x \leqslant t \leqslant t_o, \tag{1b}$$

$$\mathbf{u}(t) = \eta(t), \qquad t_o - \Delta_u \leqslant t \leqslant t_o, \tag{1c}$$

where $\mathbf{x}(t)$ and $\mathbf{u}(t)$ are, respectively, the state and the control vectors; $\mathbf{A}(t)$, $\mathbf{A}_i(t)$, $i=1,2,\ldots,N$, $\mathbf{B}(t)$ and $\mathbf{B}_i(t)$, $i=1,2,\ldots,R$, are real, piecewise continuous matrices of appropriate dimensions defined on the appropriate intervals; t_0 is the initial process time; $\phi(t)$ and $\eta(t)$ are specified initial functions; h_{xi}, $i=1,2,\ldots,N$, and h_{ui}, $i=1,2,\ldots,R$, are constant positive scalars, and

$$\Delta_x = \max_i h_{x_i} \tag{1d}$$

$$\Delta_u = \max_i \ h_{u_i} \tag{1e}$$

The cost functional to be minimized is the same as that in (7.2.35), repeated here for convenience:

$$J = ½\mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + ½\int_{t_o}^{t_f} [\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t)+\mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t)]dt \tag{2}$$

The problem is to find a control $\mathbf{u}(t)$, $t_o \leqslant t \leqslant t_f$, which for fixed final time t_f and free final state $\mathbf{x}(t_f)$ minimizes the cost functional J in (2).

The following theorem is essential in obtaining a suboptimal control for the problem under consideration.

1. Theorem. Consider the sequence of linear state equations

$$\dot{\mathbf{x}}_k(t) = \mathbf{A}(t)\mathbf{x}_k(t) + \sum_{i=1}^{N}\mathbf{A}_i(t)\mathbf{x}_{k-1}(t-h_{x_i}) + \mathbf{B}(t)\mathbf{u}_k(t)$$
$$+ \sum_{i=1}^{R}\mathbf{B}_i(t)\mathbf{u}_{k-1}(t-h_{u_i}), \quad t \geqslant t_o, \quad k=1,2,\ldots, \tag{3a}$$

with

$$\mathbf{x}_o(t) = \Psi(t,t_o)\mathbf{y}(t_o), \quad t \geqslant t_o, \tag{3b}$$
$$\mathbf{u}_o(t) = \mathbf{f}(t), \quad t \geqslant t_o, \tag{3c}$$

$$\mathbf{x}_k(t) = \phi(t), \quad t_o - \Delta_x \leqslant t \leqslant t_o, \quad k=0,1,2,\ldots, \tag{3d}$$
$$\mathbf{u}_k(t) = \eta(t), \quad t_o - \Delta_u \leqslant t \leqslant t_o, \quad k=0,1,2,\ldots, \tag{3e}$$

and the sequence of associated cost functionals

$$J_k = \tfrac{1}{2}\mathbf{x}'_k(t_f)\mathbf{F}\mathbf{x}_k(t_f) + \tfrac{1}{2}\int_{t_o}^{t_f}[\mathbf{x}'_k(t)\mathbf{Q}(t)\mathbf{x}_k(t) + \mathbf{u}'_k(t)\mathbf{R}(t)\mathbf{u}_k(t)]dt,$$
$$k=0,1,2,\ldots, \tag{4}$$

where $\mathbf{f}(t)$ is any arbitrary continuous function and $\Psi(t,\tau)$ is the transition matrix corresponding to the matrix $\mathbf{A}(t) - \mathbf{S}(t)\mathbf{K}(t)$ [7.19], where

$$\mathbf{S}(t) = \mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}'(t) \tag{5}$$

and $\mathbf{K}(t)$ is the symmetric positive-definite solution of the matrix Riccati differential equation

$$\dot{\mathbf{K}}(t) + \mathbf{K}(t)\mathbf{A}(t) + \mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{S}(t)\mathbf{K}(t) + \mathbf{Q}(t) = 0 \tag{6a}$$

with boundary condition

$$\mathbf{K}(t_f) = \mathbf{F}. \tag{6b}$$

Suppose that for the kth optimization problem, the optimal state trajectory is $\mathbf{x}_k^*(t)$, and the optimal control is $\mathbf{u}_k^*(t)$. Then, the sequences $\{\mathbf{x}_k^*(t)\}$ and $\{\mathbf{u}_k^*(t)\}$ uniformly converge, respectively, to $x^*(t)$ and $u^*(t)$ the optimal state trajectory and the optimal control for the optimization problem given by (1) and (2).

In order to prove the above theorem, the following lemma will first be established.

2. Lemma. Consider the delay-differential equation

$$\dot{x}(t) = A(t)x(t) + \sum_{i=1}^{N} A_i(t)x(t-h_{x_i}), \quad t \geqslant t_o, \tag{7a}$$

$$x(t) = \phi(t), \quad t_o - \Delta_x \leqslant t \leqslant t_o, \tag{7b}$$

where $A(t)$ and $A_i(t)$, $i=1,2,\ldots,N$, are piecewise continuous matrices and Δ_x is defined in (1d). Then the sequence of vector functions $\{x_k(t)\}$ defined by

$$x_o(t) = \Phi(t,t_o)\phi(t_o), \quad t \geqslant t_o, \tag{8a}$$

$$x_k(t) = \Phi(t,t_o)\phi(t_o) + \int_{t_o}^{t} \Phi(t,\tau)\sum_{i=1}^{N} A_i(\tau)x_{k-1}(\tau-h_{x_i})\, d\tau, \quad t \geqslant t_o, \quad k=1,2,\ldots, \tag{8b}$$

$$x_k(t) = \phi(t), \quad t_o - \Delta_x \leqslant t \leqslant t_o, \quad k=0,1,2,\ldots, \tag{8c}$$

where $\Phi(t,\tau)$ is the transition matrix corresponding to $A(t)$ [7.19], converges uniformly to the solution of (7).

Proof of Lemma 2. Consider $\{x_k(t)\}$ as a sequence in $C^N[t_0-\Delta_x,t_f]$, the space of continuous n-vector-valued functions. (n is the dimension of x.) Equation (8b), for $k=1$ and 2, yields

$$x_2(t) - x_1(t) = \int_{t_o}^{t} \Phi(t,\tau)\sum_{i=1}^{N} A_i(\tau)[x_1(\tau-h_{x_i}) - x_o(\tau-h_{x_i})]\, d\tau, \quad t \geqslant t_0 \tag{9}$$

Therefore, we have

$$\|x_2(t) - x_1(t)\| \leqslant M\sum_{i=1}^{N} N_i \int_{t_o+h_{x_i}}^{t} \|x_1(\tau-h_{x_i}) - x_o(\tau-h_{x_i})\| dt, \quad t_o \leqslant t \leqslant t_f \tag{10a}$$

where

$$M = \sup_{t,\tau \in [t_o,t_f]} \|\Phi(t,\tau)\|, \tag{10b}$$

$$N_i = \sup_{\tau \in [t_o,t_f]} A_i(\tau)\|. \tag{10c}$$

It is convenient to choose a norm $\|\,.\,\|$ such that $\|\Phi(t_o,t_o)\| = \|I\| = 1$. This then guarantees that $M \geqslant 1$. From (8) we have

$$\begin{aligned}\|x_1(t) - x_o(t)\| &\leqslant M \sum_{i=1}^{N} N_i \int_{t_o}^{t} \|x_o(\tau - h_{x_i})\| d\tau \\ &= M \sum_{i=1}^{N} N_i \left[\int_{t_o}^{t_o+h_{x_i}} \|\phi(\tau - h_{x_i})\| d\tau + \int_{t_o+h_{x_i}}^{t} \|\Phi(\tau - h_{x_i}, t_o)\phi(t_o)\| dt \right] \\ &\leqslant M \sum_{i=1}^{N} N_i [L h_{x_i} + ML(t - t_o - h_{x_i})] \leqslant LM^2N(t-t_o), \quad t_o \leqslant t \leqslant t_f \end{aligned} \tag{11a}$$

where

$$L = \sup_{t \in [t_o - \Delta_x, t_o]} \|\phi(t)\|, \tag{11b}$$

$$N = \sum_{i=1}^{R} N_i. \tag{11c}$$

Hence, (10a) yields

$$\|x_2(t) - x_1(t)\| \leqslant LM^3N^2[(t-t_o)^2/2], \quad t_o \leqslant t \leqslant t_f \tag{12}$$

and, by induction,

$$\|x_k(t) - x_{k-1}(t)\| \leqslant LM^{k+1}N^k[(t-t_o)^k/k!], \quad t_o \leqslant t \leqslant t_f \tag{13}$$

Applying the triangle inequality we have, for any r,

$$\begin{aligned}\|x_{k+r}(t) - x_k(t)\| &\leqslant \sum_{i=k+1}^{k+r} LM^{i+1}N^i[t-t_o)^i/i!] \\ &\leqslant LM[MN(t-t_o)]^{k+1} \exp[MN(t-t_o)]/k!, \quad t_o \leqslant t \leqslant t_f \end{aligned} \tag{14}$$

Therefore, the sequence $\{x_k(t)\}$ is a Cauchy sequence in $C^N[t_0-\Delta_x, t_f]$. From the completeness of this space it follows that this sequence is uniformly convergent [7.20,7.21]. Since r is arbitrary, the limit of this sequence is clearly the solution to (7), and the lemma is proved.

Proof of Theorem 1. Let us define

$$\mathbf{h}_{k-1}(t) = \sum_{i=1}^{N}\mathbf{A}_i(t)\mathbf{x}_{k-1}(t-h_{x_i}) + \sum_{i=1}^{R}\mathbf{B}_i(t)u_{k-1}(t-h_{u_i}), \qquad t \geqslant t_o. \tag{15}$$

Note that $\mathbf{h}_{k-1}(t)$ is known in the kth iteration and acts like an extra perturbing input in the kth state equation. Then, the Hamiltonian function for the kth optimization problem is

$$\begin{aligned} H_k = {} & \tfrac{1}{2}[\mathbf{x}'_k(t)\mathbf{Q}(t)\mathbf{x}_k(t) + \mathbf{u}'_k(t)\mathbf{R}(t)\mathbf{u}_k(t)] \\ & + \mathbf{p}'_k[\mathbf{A}(t)\mathbf{x}_k(t) + \mathbf{B}(t)\mathbf{u}_k(t) + \mathbf{h}_{k-1}(t)], \end{aligned} \tag{16}$$

where $\mathbf{p}_k(t)$ is the *costate vector* for the *kth* optimization problem. The necessary and sufficient conditions for optimality are (see Section 6.3):

$$\partial H_k/\partial \mathbf{x}_k(t) = -\dot{\mathbf{p}}_k(t) = \mathbf{Q}(t)\mathbf{x}_k(t) + \mathbf{A}'(t)\mathbf{p}_k(t), \tag{17a}$$

$$\partial H_k/\partial \mathbf{u}_k(t) = 0 = \mathbf{R}(t)\mathbf{u}_k(t) + \mathbf{B}'(t)\mathbf{p}_k(t), \tag{17b}$$

$$\partial H_k/\partial \mathbf{p}_k(t) = \dot{\mathbf{x}}_k(t) = \mathbf{A}(t)\mathbf{x}_k(t) + \mathbf{B}(t)\mathbf{u}_k(t) + \mathbf{h}_{k-1}(t), \tag{17c}$$

for $t_o \leqslant t \leqslant t_f$, and

$$\mathbf{p}_k(t_f) = \mathbf{F}\mathbf{x}_k(t_f) \tag{17d}$$

where $k=1,2,\ldots$. Equations (17a) and (17c) can be uncoupled by defining the *adjoint vectors* $\mathbf{g}_k(t)$, $k=1,2,\ldots$, as follows:

$$\mathbf{p}_k(t) = \mathbf{K}(t)\mathbf{x}_k(t) + \mathbf{g}_k(t), \qquad t_o \leqslant t \leqslant t_f \tag{18}$$

Using (17d), (18) implies

$$\mathbf{g}_k(t_f) = 0. \tag{19}$$

Equations (17) and (18) imply that

$$\begin{aligned} \dot{\mathbf{p}}_k(t) &= [\dot{\mathbf{K}}(t) + \mathbf{K}(t)\mathbf{A}(t) - \mathbf{K}(t)\mathbf{S}(t)\mathbf{K}(t)]\mathbf{x}_k(t) \\ &\quad -\mathbf{K}(t)\mathbf{S}(t)\mathbf{g}_k(t) + \mathbf{K}(t)\mathbf{h}_{k-1}(t) + \dot{\mathbf{g}}_k(t) \\ &= -[\mathbf{A}'(t)\mathbf{K}(t) + \mathbf{Q}(t)]\mathbf{x}_k(t) - \mathbf{A}'(t)\mathbf{g}_k(t), \end{aligned} \tag{20}$$

where $\mathbf{S}(t)$ is given by (5). Since (20) is true for all vectors $\mathbf{x}_k(t)$ and $\mathbf{g}_k(t)$, $\mathbf{K}(t)$ satisfies (6a) and we have

$$\dot{\mathbf{g}}_k(t) = -[\mathbf{A}(t) - \mathbf{S}(t)\mathbf{K}(t)]'\mathbf{g}_k(t) - \mathbf{K}(t)\mathbf{h}_{k-1}(t). \tag{21}$$

From (17b) and (18), the optimal control for the kth optimization problem can be written as

$$\mathbf{u}_k^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{K}(t)\mathbf{x}_k^*(t) - \mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{g}_k(t) \tag{22}$$

hence, from (17c) and (22), the optimal state trajectory $\mathbf{x}_k^*(t)$ is the solution to

$$\dot{\mathbf{x}}_k^*(t) = [\mathbf{A}(t) - \mathbf{S}(t)\mathbf{K}(t)]\mathbf{x}_k^*(t) - \mathbf{S}(t)\mathbf{g}_k(t) + \mathbf{h}_{k-1}(t), \quad t_o \leqslant t \leqslant t_f \tag{23}$$

From (21), note that $\mathbf{g}_k(t)$ depends on known functions and $\mathbf{h}_{k-1}(t)$. Also note that the homogeneous parts of (21) and (23) are adjoint. The solution to (23) with boundary condition (3d) is

$$\mathbf{x}_k^*(t) = \Psi(t,t_o)\phi(t_o) + \int_{t_o}^{t}\Psi(t,\tau)[-\mathbf{S}(\tau)\mathbf{g}_k(\tau) + \mathbf{h}_{k-1}(\tau)]\, d\tau,$$
$$t_o \leqslant t \leqslant t_f, \quad k=1,2,\ldots \tag{24}$$

where $\Psi(t,\tau)$ is the state transition matrix corresponding to the matrix $\mathbf{A}(t) - \mathbf{S}(t)\mathbf{K}(t)$. Comparison of (3b) and (24) with (8a) and (8b) shows that the sequence $\{\mathbf{x}_k^*(t)\}$ converges uniformly. Also, the sequences $\{\mathbf{u}_k^*(t)\}$ and $\{\mathbf{g}_k(t)\}$ converge because, from (22), (21) and (15) these sequences are related to $\{\mathbf{x}_k^*(t)\}$ by continuous transformations. From Lemma 2, the limit of the sequence $\{\mathbf{x}_k^*(t)\}$ is the solution to

$$\dot{\mathbf{x}}^*(t) = [\mathbf{A}(t) - \mathbf{S}(t)\mathbf{K}(t)]\mathbf{x}^*(t) - \mathbf{S}(t)\mathbf{g}(t) + \sum_{l=1}^{N}\mathbf{A}_l(t)x^*(t-h_{xl})$$
$$+ \sum_{l=1}^{R}\mathbf{B}_l(t)\mathbf{u}^*(t-h_{ul}) \tag{25a}$$

$$\mathbf{x}^*(t) = \phi(t), \quad t_o - \Delta_x \leqslant t \leqslant t_o, \tag{25b}$$

$$\mathbf{u}^*(t) = \eta(t), \quad t_o - \Delta_u \leqslant t \leqslant t_o, \tag{25c}$$

where $\mathbf{x}^*(t)$, $\mathbf{u}^*(t)$, and $\mathbf{g}(t)$ are, respectively, the limits of the sequences $\{\mathbf{x}_k^*(t)\}$, $\{\mathbf{u}_k^*(t)\}$, $\{\mathbf{g}_k(t)\}$. From (22),

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{K}(t)\mathbf{x}^*(t) - \mathbf{R}^{-1}(t)\mathbf{B}'(t)\mathbf{g}(t), \quad t \geqslant t_o. \tag{26}$$

Substituting $\mathbf{u}^*(t)$ in (26) for $\mathbf{u}(t)$ in (1a) and comparing the result with (25a) shows that $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$ are, respectively, the optimal state trajectory and the optimal control for the optimization problem given by (1) and (2). Thus, Theorem 1 is proved. Δ

The above procedure yields the control and the state trajectory which are very close to optimal if a very large number of iterations are performed. A *Pth-order suboptimal control* $\mathbf{u}_p^*(t)$ is obtained if P iterations are performed. The computational procedure to determine the *Pth* order suboptimal control can be summarized as follows:

Step 1: Solve the Riccati differential equation (6).

Step 2: Choose an arbitrary vector function $\mathbf{f}(t)$. Set $k=1$.

Step 3: Calculate $\mathbf{h}_{k-1}(t)$ from (3) and (15).

Step 4: Solve (21) with boundary condition (19) to obtain $\mathbf{g}_k(t)$.

Step 5: Solve (23) with boundary condition (3d) to obtain $\mathbf{x}^*(t)$.

Step 6: Calculate $\mathbf{u}_k^*(t)$ from (22).

Step 7: If $k < P$, let $k=k+1$ and go to Step 3; otherwise store $\mathbf{x}_p^*(t)$ and $\mathbf{u}_p^*(t)$.

3. **Example** [7.5]. We will consider Example 7.2.2 here. That is, the state equation of the system under consideration and the cost functional are given by equations (7.2.66a-d) and (7.2.67).

To determine a suboptimal control for this example, the function $\mathbf{f}(t)$ in (3c) was taken to be $-5(t+1)$. All the differential equations were solved by the finite-difference method using a step size $\Delta t = 0.05$. The results are summarized in Table 1. From iteration No. 5 on, the results compare favorably with those of Section 7.2. However, it is noticed that the convergence is faster with the method of this section.

7.3.2 Extensions to Other Systems

The results can be extended directly to the case where the delays h_{xi}, $i=1,2,\ldots,N$, and h_{ui}, $i=1,2,\ldots,R$, are time-varying. They can also be extended to nonlinear multiple delay systems, provided that the nonlinearities involving the state variables satisfy the Lipschitz condition. To clarify this point, consider the case where the state equation (1a) is of the form

Table 1. Results for Example 3

Iteration number k	$\sup_{0 \leq t \leq 3} \lVert x_{k+1}(t) - x_k(t) \rVert$	$\sup_{0 \leq t \leq 3} \lVert u_{k+1}(t) - u_k(t) \rVert$	Value of cost functional J
1	1.655	14.190	69.9648437
2	0.927	6.639	46.4596710
3	0.429	2.963	26.6414032
4	0.056	0.538	25.9881439
5	0.011	0.115	24.8826416
6	2.98×10^{-3}	2.08×10^{-2}	24.4650513
7	1.21×10^{-3}	4.84×10^{-3}	24.1571863
8	5.06×10^{-4}	2.68×10^{-3}	22.0519323
9	2.15×10^{-4}	1.40×10^{-3}	22.0472483

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \sum_{i=1}^{N} \mathbf{f}_i(\mathbf{x}(t-h_{xi}(t))) + \mathbf{B}(t)\mathbf{u}(t) + \sum_{i=1}^{R} \mathbf{g}_i(\mathbf{u}(t-h_{ui}(t))), \quad t \geq t_o, \tag{27}$$

where positive constants F_i exist such that, for all vectors α_1 and α_2,

$$\lVert \mathbf{f}_i(\alpha_1) - \mathbf{f}_i(\alpha_2) \rVert \leq F_i \lVert \alpha_1 - \alpha_2 \rVert, \quad i=1,2,\ldots,N. \tag{28}$$

In this case, the proofs of the lemma and the theorem proceed as before, except that N_i in (10) must be replaced by F_1, $i=1,2,\ldots,N$. Jamshidi and Malek-Zavarei [7.21] used this approach to determine suboptimal control for large-scale TD systems.

7.4 SUBOPTIMAL CONTROL OF NONLINEAR TIME-DELAY SYSTEMS

Thus far, in this chapter all the suboptimal control methods discussed have been limited to linear TD systems. There are very few computational algorithms which are devoted to optimize a nonlinear TD process. One such schemes is due to Jamshidi [7.22] which is briefly discussed here. Suboptimal control of nonlinear TD systems with multiple delays in both control and state variables within the context of large-scale TD systems are, however, discussed in some detail in Section 8.5.

The main object of this section is to apply methods of series expansion (see Section 2.5) and transformation to obtain a suboptimal control for nonlinear TD systems. The proposed method consists of three stages. The first stage is a Taylor series expansion of the originally nonlinear system state equation about the nominal state and control functions. The resulting linear TD system with time-varying coefficients is then transformed into a linear nondelay, nonhomogeneous system using the basic approach of Bate [7.6]. The third stage is the application of the maximum principle to the transformed system leading to the desired suboptimal control. This control can then be applied to the original TD system to obtain the state.

Consider a nonlinear TD system described by

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t),\ \mathbf{x}(t-h_x),\ \mathbf{u}(t),\ \mathbf{u}(t-h_u)) \tag{1}$$

with initial functions,

$$\mathbf{x}(t) = \mathbf{x}_0(t)\ ,\ -h_x \leqslant t \leqslant 0 \tag{2a}$$

$$\mathbf{u}(t) = \mathbf{u}_0(t)\ ,\ -h_u \leqslant t \leqslant 0 \tag{2b}$$

where $\mathbf{x}(t)$ and $\mathbf{u}(t)$ are $n\times 1$ and $r\times 1$ state and control vectors, respectively, and h_x and h_u are delays in state and control. The function $\mathbf{f}(\cdot)$ is assumed to be a continuously differentiable function of its arguments. The optimal control problem is formulated as follows. it is required to find an optimal control vector $\mathbf{u}^*$ which satisfies (1) and (2) while minimizing an energy cost functional

$$J = Q(\mathbf{x}(t_f)) + \int_0^{t_f} G(\mathbf{u}(t), t)dt \tag{3}$$

As it has been mentioned earlier, the application of the maximum principle to the above problem requires the solution of a *TPBV* problem with both delay and advance terms which is computationally difficult.

The first stage of the proposed scheme is to follow the discussion of Section 2.5 by linearizing (1) about nominal trajectory $(\mathbf{x}_n,\ \mathbf{u}_n)$. Let $(\delta\mathbf{x},\ \delta\mathbf{u})$ be the trajectories of deviations between $(\mathbf{x},\mathbf{u})$ and $(\mathbf{x}_n,\ \mathbf{u}_n)$, *i.e.*,

$$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_n \tag{4a}$$

$$\delta\mathbf{u} = \mathbf{u} - \mathbf{u}_n \tag{4b}$$

Assume that $(\mathbf{x}_n,\ \mathbf{u}_n)$ satisfy (1), *i.e.*,

$$\dot{\mathbf{x}}_n = \mathbf{f}(\mathbf{x}_n, \mathbf{x}_n(t-h_x), \mathbf{u}_n, \mathbf{u}_n(t-h_u)) \tag{5}$$

A Taylor series expansion of (1) about $(\mathbf{x}_n, \mathbf{u}_n)$ will result in a linear time - varying TD system

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) + \mathbf{B}(t)\delta\mathbf{x}(t-h_x) + \mathbf{C}(t)\delta\mathbf{u}(t) + \mathbf{D}(t)\delta\mathbf{u}(t-h_u) \tag{6}$$

with initial functions

$$\delta_{\mathbf{x}_0}(t) = \mathbf{x}_0(t) - \mathbf{x}_n(t), \ -h_x \leqslant t \leqslant 0 \tag{7a}$$

$$\delta\mathbf{u}_0(t) = \mathbf{u}_0(t) - u_n(t), \ -h_u \leqslant t \leqslant 0 \tag{7b}$$

and

$$\mathbf{A}(t) \triangleq \partial\mathbf{f}(.)/\partial\mathbf{x}_n \ , B(t) \triangleq \partial\mathbf{f}(.)/\partial\mathbf{x}_n(t-h_x) \tag{8a}$$

$$\mathbf{C}(t) \triangleq \partial\mathbf{f}(.)/\partial\mathbf{u}_n \ , \mathbf{D}(t) \triangleq \partial\mathbf{f}(.)/\partial\mathbf{u}_n(t-h_u) \tag{8b}$$

The following theorem provides reduction of linear TD system (7) to a nondelay one.

1. Theorem. Dynamical systems I and II, given below, have the same optimal control *w.r.t.* the following cost functional:

$$\hat{J} = q(\delta\mathbf{x}(t_f)) + \int_0^{t_f} g(\mathbf{u}(t), t)dt \tag{9}$$

System I:

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) + \mathbf{B}(t)\delta\mathbf{x}(t-h_x) + \mathbf{C}(t)\delta\mathbf{u}(t) + \mathbf{D}(t)\delta\mathbf{u}(t-h_u) \tag{10}$$

$$\delta\mathbf{x}(t) = \delta x_0(t), \ -h_x \leqslant t \leqslant 0 \tag{11a}$$

$$\delta\mathbf{u}(t) = \delta u_0(t), \ -h_u \leqslant t \leqslant 0 \tag{11b}$$

System II:

$$\delta\dot{\mathbf{z}} = \mathbf{K}(t)\delta\mathbf{u}(t), \quad 0 < t < t_f \tag{12a}$$

$$\delta\mathbf{z}(0) = \mathbf{L}(0)\delta\mathbf{x}(0) + \int_{-h_x}^{0} \mathbf{L}(t+h_x)\mathbf{B}(t+h_x)\delta\mathbf{x}(t)dt$$
$$+ \int_{-h_u}^{0} \mathbf{L}\ (t+h_u)\ \mathbf{D}\ (t+h_u)\delta\mathbf{u}(t)dt \tag{12b}$$

where the $n \times n$ adjoint matrix $\mathbf{L}$ is obtained from

$$\dot{\mathbf{L}}(t) = \mathbf{L}(t)\mathbf{A}(t) + \mathbf{L}(t+h_x)\mathbf{B}(t+h_x), \quad t \leqslant t_f \tag{13a}$$

$$\mathbf{L}(t_f) = I \tag{13b}$$

$$\mathbf{L}(t) = 0, \quad t \geqslant t_f \tag{13c}$$

and

$$\mathbf{K}(t) = \mathbf{L}(t)\mathbf{C}(t) + \mathbf{L}(t+h_u)\mathbf{D}(t+h_u) \tag{14}$$

The proof of this theorem, with some minor modifications, has been given by Bate [7.6]. This is essentially the second state of the design. The n-dimensional vector $\delta\mathbf{z}$, called *pseudo-state*, can be expressed as

$$\delta\mathbf{z}(t) = \mathbf{L}(t)\delta\mathbf{x}(t) + \int_{t-h_x}^{t} \mathbf{L}(\sigma+h_x)\mathbf{B}(\sigma+h_x)\delta\mathbf{x}(\sigma)d\sigma$$
$$+ \int_{t-h_u}^{t} \mathbf{L}(\sigma+h_u)\mathbf{D}(\sigma+h_u)\delta\mathbf{u}(\sigma)d\sigma \tag{15}$$

From (13) and (15) it follows that

$$\delta\mathbf{z}(t_f) = \delta\mathbf{x}(t_f) \tag{16}$$

The ordinary optimal control problem given by (9)-(11), (15) and (16) can now be formulated using the necessary conditions of the maximum principle. Define the Hamiltonian function

$$H(\delta\mathbf{z}, \delta\mathbf{p}, \delta\mathbf{u}, t) = -g(\delta\mathbf{u}, t) + \delta\mathbf{p}'\ \mathbf{K}(t)\delta\mathbf{u} \tag{17}$$

where $\delta\mathbf{p}$ is the co-state vector of $\delta\mathbf{z}$, whose equation is obtained from

$$\delta\dot{\mathbf{p}} = -\partial H/\partial(\delta\mathbf{z}(t)) = 0, \ \delta\mathbf{p}(t_f) = q_{\delta_z}(\delta\mathbf{z}(t_f)) \tag{18}$$

Hence, the costate $\delta\mathbf{p}$ is

$$\delta \mathbf{p}(t) = q_{\delta z}(\delta z(t_f)) = constant \tag{19}$$

The optimality condition of the Hamiltonian H is given by

$$0 = \partial H/\partial(\delta \mathbf{u}) = -g_{\delta u}(\delta \mathbf{u}, t) + \mathbf{K}'\delta \mathbf{p}(t) = -g_{\delta_u}(\delta \mathbf{u}, t) + \mathbf{K}' q_{\delta z}(\delta \mathbf{z}(t_f)) \tag{20}$$

From the above equation an open-loop expression for the optimal control can be obtained. For a quadratic cost functional $g(\delta \mathbf{u}, t) = 1/2\delta \mathbf{u}'\mathbf{R}\delta \mathbf{u}$, $q(\delta \mathbf{z}(t_f)) = \delta \mathbf{z}'(t_f)$ $\mathbf{M}\delta \mathbf{z}(t_f)$, (20) will lead to the optimal closed - loop control [7.15]

$$\delta \mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{K}'(t)\mathbf{P}(t)\delta \mathbf{z}(t) \tag{21}$$

where $\mathbf{P}(t)$ satisfies the following Riccati-type equation:

$$\dot{\mathbf{P}}(t) = \mathbf{P}(t)\mathbf{K}(t)\mathbf{R}^{-1}\mathbf{K}'(t)\mathbf{P}(t) \tag{22a}$$

$$\mathbf{P}(t_f) = \mathbf{M} \tag{22b}$$

Substituting (15) into (21) yields the proposed suboptimal designed control vector

$$\delta \mathbf{u}^*(t) = -\mathbf{K}_1(t)\delta \mathbf{x}(t) - \int_{t-h_x}^{t} \mathbf{K}_2(t, \sigma)\delta \mathbf{x}(\sigma)d\sigma - \int_{t-h_u}^{t} \mathbf{K}_3(t, \sigma)\delta \mathbf{u}(\sigma)d\sigma \tag{23}$$

where

$$\mathbf{K}_1(t) = \mathbf{R}^{-1}\,\mathbf{K}'(t)\mathbf{P}(t)\mathbf{L}(t) \tag{24a}$$

$$\mathbf{K}_2(t,\sigma) = \mathbf{R}^{-1}\mathbf{K}'(t)\mathbf{P}(t)\mathbf{L}(\sigma+h_x)\mathbf{B}(\sigma+h_x) \tag{24b}$$

$$\mathbf{K}_3(t,\sigma) = \mathbf{R}^{-1}\mathbf{K}'(t)\mathbf{P}(t)\mathbf{L}(\sigma+h_u)\mathbf{D}(\sigma+h_u) \tag{24c}$$

It is noted that the control law (23) has a direct feedback portion and two on-line portions in $\delta \mathbf{x}(t)$ and $\delta \mathbf{u}(t)$, delayed during τ seconds preceding the instant t. Figures 1 and 2 present the proposed sub-optimal control design using the linearized and original models, respectively.

2. Computational Algorithm - Optimal Control. The proposed design method for a quadratic cost functional is summarized by the following steps:

1. Integrate (13) backward in time and store $\mathbf{L}(t)$, $-h_x < t < t_f$.

2. Compute matrix $\mathbf{K}(t)$ using (14).

3. Integrate (22) and store the Riccati matrix $\mathbf{P}(t)$.

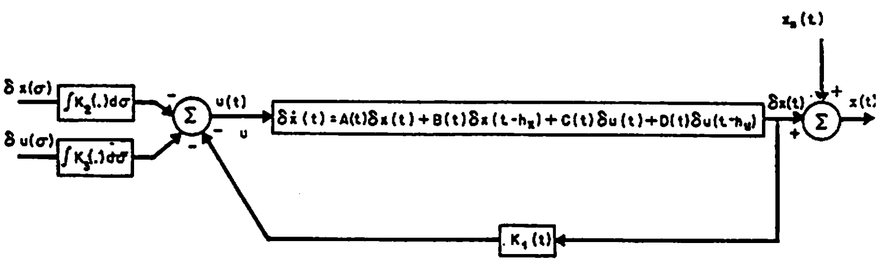

Figure 1. Proposed Suboptimal Control Design Using Linearized Model

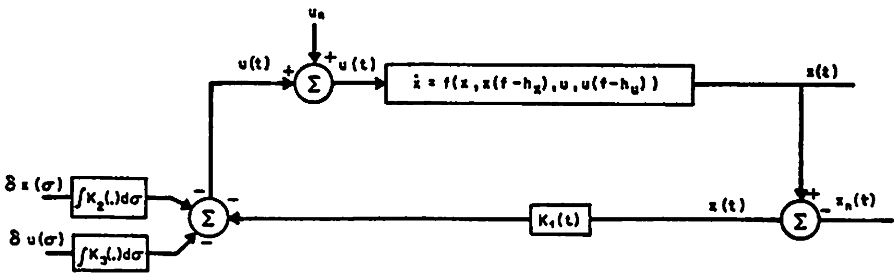

Figure 2. Proposed Suboptimal Control Design Using Original Model

4. For instant t and dummy argument σ obtain the feedback matrices $\mathbf{K}_1(t)$, $\mathbf{K}_2(t, \sigma)$ and $\mathbf{K}_3(t,\sigma)$ using (24) and store.

5. Integrate the terms $\mathbf{K}_2(t,\sigma)\delta\mathbf{x}(\sigma)$ and $\mathbf{K}_3(t, \sigma)\delta\mathbf{u}(\sigma)$ in the interval $t-h_u$ and t using (23) to obtain $\delta\mathbf{u}^*(t)$.

6. Set $t = t + \Delta t = t + mt$.

7. If t $< t_f$ repeat step 4.

8. $\mathbf{u}^*(t) = \delta\mathbf{u}^*(t) + \mathbf{u}_n(t)$

It is noted that the above procedure assumes that the step sizt t is an integral multiple of delay h_x, *i.e.*, $\Delta t = mh_x, m > 1$.

3. Computational Algorithm - Adjoint Matrix. The evaluation of adjoint matrix $\mathbf{L}(t)$ in (13) requires some thought, especially for stable matrices **A** and **B**. This equation can be integrated in two basic ways, depending on the nature of matrices **A, B, C** *and* **D**.

(i) For constant matrices **A**, **B**, **C** and **D**, (13) can be changed to [7.6]

$$\frac{d}{d\alpha}\mathbf{L}_n(\alpha) = \mathbf{L}_n(\alpha)\mathbf{A} + \mathbf{L}_{n-1}(\alpha)\mathbf{B} \tag{25}$$

$$\mathbf{L}_{-1}(\alpha) = 0 \tag{26a}$$

$$\mathbf{L}_0(0) = \mathbf{I} \tag{26b}$$

$$\mathbf{L}_n(0) = \mathbf{L}_{n-1}(\tau) \tag{26c}$$

where

$$\mathbf{L}_n(\alpha) = \mathbf{L}(t_f - n\tau - \alpha) \tag{26d}$$

This sequence of nonhomogeneous differential equations can be solved as follows [7.6]:

$$\mathbf{L}_n(\sigma) = \left[\mathbf{L}_{n-1}(\tau) + \int_0^{\alpha} \mathbf{L}_{n-1}(\sigma')\mathbf{B}exp(-\mathbf{A}\sigma')d\sigma'\right]exp(\mathbf{A}\alpha) \tag{27}$$

(ii) For time-varying matrices $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$ and $\mathbf{D}(t)$, the most reasonable method is the numerical integration of the differential-difference equation (13). This method could have been applied for the case where all the matrices were constant, but the method introduced above is more convenient.

4. Example. The design method outlined in the previous sections is applied to a numerical example [7.22]. Consider a first-order nonlinear TD system

$$\dot{x}(t) = -2x^2(t) + x(t)x(t-0.1) + u(t)u(t-0.1) \tag{28}$$

with initial functions

$$x_0(t) = 10(t + 0.1),\ -0.1 \leqslant t \leqslant 0 \tag{29a}$$

$$u_0(t) = 1,\ -0.1 \leqslant t \leqslant 0 \tag{29b}$$

and the cost functional

$$J = 1/2(x(1) - 1)^2 + 1/2\int_0^1 (u(t) - 1)^2 dt \tag{30}$$

Let the nominal trajectories be $(x_n, u_n) = (1, 1)$. The linearized model for system (28) using the above nominal trajectories become

$$\delta\dot{x} = 3\delta x + \delta x(t - 0.1) + \delta u + \delta u(t - 0.1) \tag{31a}$$

$$\delta x(t) = 10t, \quad -0.1 \leqslant t \leqslant 0.0 \tag{31b}$$

$$\delta u(t) = 0, \ -0.1 \leqslant t \leqslant 0.0 \tag{31c}$$

This completes the first step of the design. Following the computational steps outlined above, the adjoint variable $\ell(t)$ is obtained by solving

$$\dot{\ell}(t) = 3\ell(t) - \ell(t + 0.1) \tag{32a}$$

$$\ell(1) = 1 \tag{32b}$$

which, according to the procedure outlined, is reduced to the following sequence of differential equations:

$$\dot{\ell}_n(t) = -3\ell_n(t) + \ell_{n-1}(t) \tag{33a}$$

$$\ell_{-1}(t) = 0, \ \ell_0(0) = 1, \ \ell_n(0) = \ell_{n-1}(0.1) \tag{33b}$$

Solving these series of differential equations leads to a value $\ell(t) = exp(3(t-1))$ for $0<t<1$. The value of $k(t)$ defined in (14), can be obtained as $k(t) = exp(3(t-1))+exp(3(t-0.9))$. The second stage provides the transformation of (31) into an ordinary differential equation

$$\dot{z} = k(t)u(t) \tag{34a}$$

with

$$z(0) = \ell(0)x(0) + \int_{-0.1}^{0} \ell(t+0.1)x(t)dt \tag{34b}$$

and cost functional

$$J = \frac{1}{2}x^2(1) + \frac{1}{2}\int_0^1 u^2(t)dt \tag{35}$$

For the nondelay optimal control problem given by (34) and (35), the application of the linear regulator theory leads to the following equations:

$$\dot{p}(t) = 0, p(1) = z(1) \tag{36a}$$

$$p(t) = -P(t)z(t) \tag{36b}$$

$$u(t) = -k(t)P(t)z(t) \tag{37}$$

$$\dot{z}(t) = -k^2(t)P(t)z(t) \tag{38}$$

where p is the constant co-state and P is the solution of the Riccati equation

$$\dot{P}(t) = -k^2(t)P^2(t), P(1) = 1 \tag{39}$$

This problem was simulated on a digital computer. The Riccati variable $P(t)$ and pseudo-state $z(t)$ are shown in Figures 3 and 4. Using the optimal control obtained in (37) the state $x(t)$ and control $u(t)$ are shown in Figures 5 and 6. The feedback law outlined in (23) and (24), and the optimal closed-loop control can be similarly obtained.

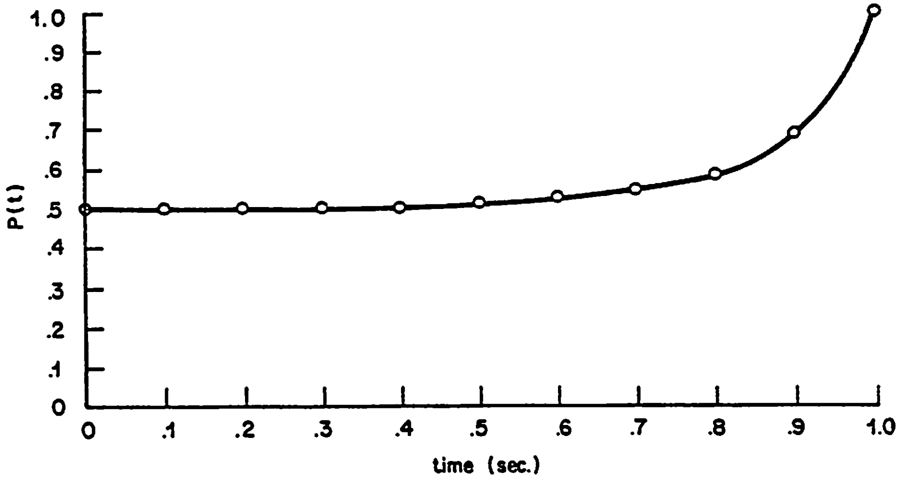

Figure 3. The Riccati Variable $P(t)$ Versus Time

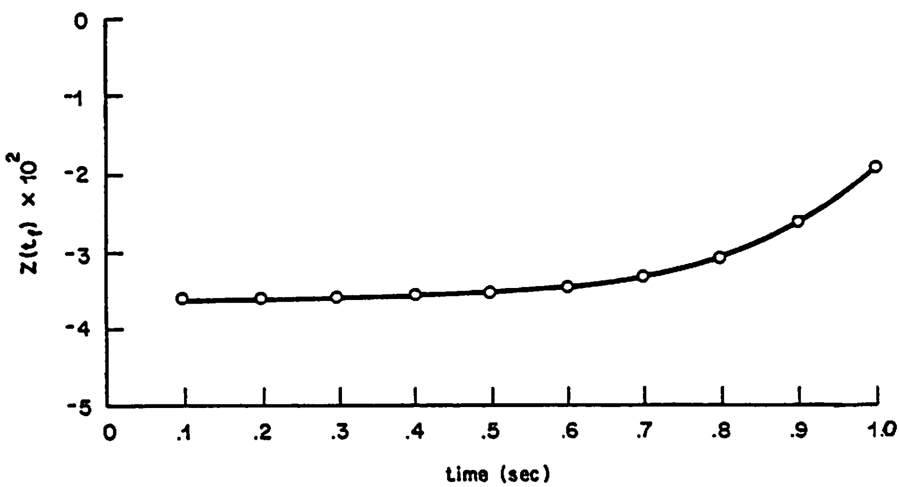

Figure 4. The Pseudo-State $z(t)$ Versus Time

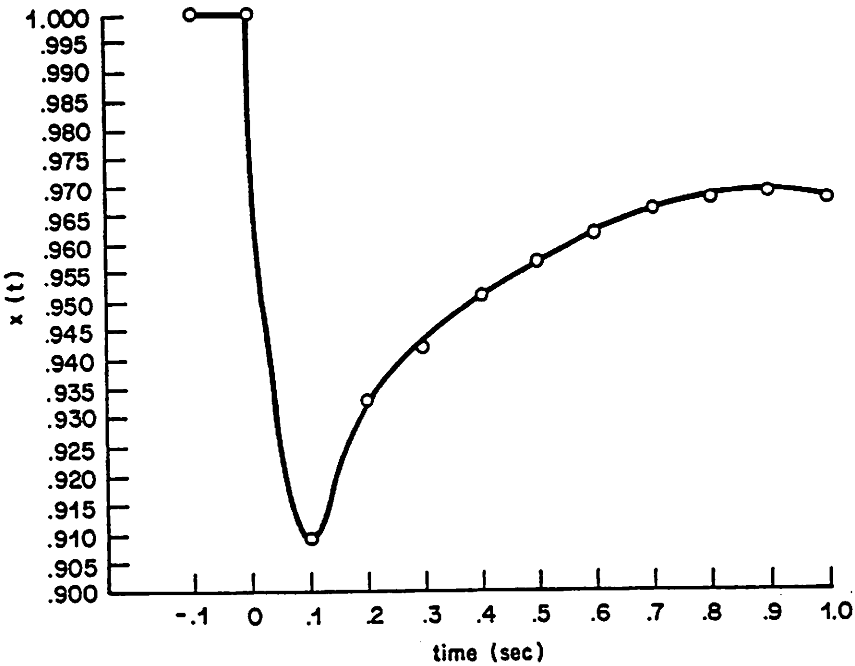

Figure 5. State $x(t) = \delta x(t) + x_n(t)$ Versus Time

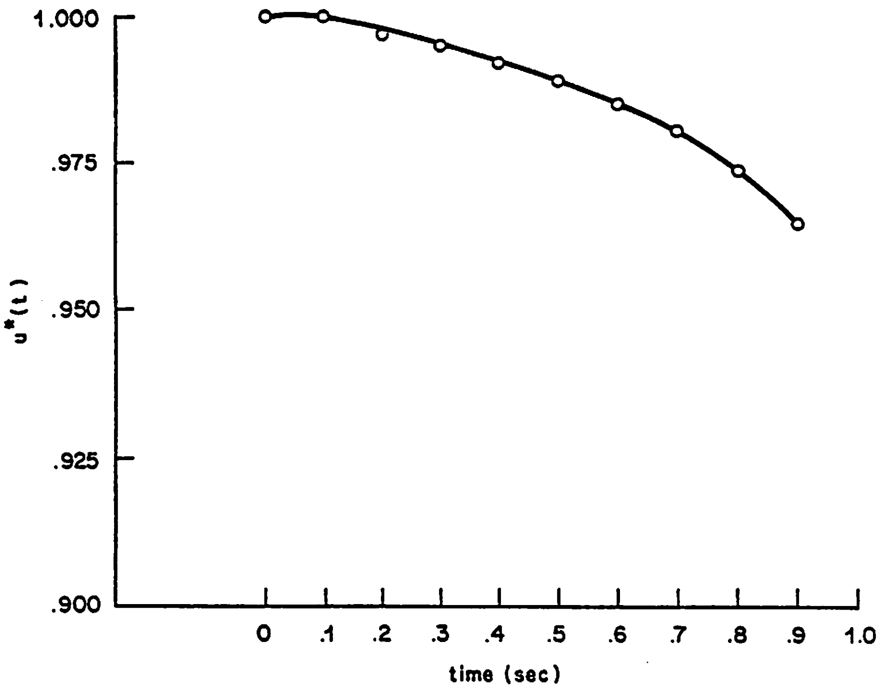

Figure 6. Control $u(t) = \delta u(t) + u_n(t)$ Versus Time

7.5 SINGULAR PERTURBATION METHOD

In this section the singular perturbation method of mathematical analysis of stiff differential equations is used to solve an optimal TD control problem. The bulk of the material of this section is based on the works of Sannuti [7.23], Inoue et. al. [7.14], and Jamshidi [7.24].

Consider a linear stationary TD system described by the following differential-difference equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{A}_1\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t) \tag{1a}$$

with initial function

$$\mathbf{x}(t) = \mathbf{x}_0(t), \quad -h \leqslant t \leqslant 0 \tag{1b}$$

where $\mathbf{x}$, $\mathbf{u}$, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{A}_1$ are vectors and matrices of appropriate dimensions. The optimal control problem is to find a control vector $\mathbf{u}^*(t)$ such that a quadratic cost functional

$$J(\mathbf{u}) = 1/2 \int_0^{t_f} (\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t))dt \tag{2}$$

is minimized, where $\mathbf{Q}$ and $\mathbf{R}$ are positive semi-definite and positive definite, respectively, $\mathbf{x}(t_f)$ is free and (1) is satisfied. The above problem was solved by the "sensitivity" approach in Section 7.2. In this section, the problem (1) - (2) is converted into a nondelay singularly - perturbed system. This is achieved by dividing the delay h into m equal subinterval of time and denoting

$$\mathbf{z}^i(t) = \mathbf{x}(t - i\epsilon), \ \epsilon = h/m \tag{3}$$

for $i = 1,...,m$. Under this change of variables, system (1a) can be approximated by high-order nondelay differential equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0\mathbf{x}(t) + \mathbf{A}_1\mathbf{z}^m(t) + \mathbf{B}\mathbf{u}(t), \ \mathbf{x}(0) = \mathbf{x}_0(0) \tag{4a}$$

$$\epsilon\dot{\mathbf{z}}^1(t) = \mathbf{x}(t) - \mathbf{z}^1(t), \ \mathbf{z}^1(0) = \mathbf{x}_0(-\epsilon) \tag{4b}$$

$$\vdots$$

$$\epsilon\dot{\mathbf{z}}^i(t) = \mathbf{z}^{i-1}(t) - \mathbf{z}^i(t), \ z^i(0) = \mathbf{x}_0(-i\epsilon) \tag{4c}$$

$$\vdots$$

$$\epsilon\dot{\mathbf{z}}^m(t) = \mathbf{z}^{m-1}(t) - \mathbf{z}^m(t), \ \mathbf{z}^m(0) = \mathbf{x}_0(-m\epsilon) \tag{4d}$$

The above model can be rewritten in a more compact form as

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0\mathbf{x}(t) + \mathbf{A}_1\mathbf{z}^m(t) + \mathbf{B}\mathbf{u}(t),\ \mathbf{x}(0) = \mathbf{x}_0(0) \tag{5a}$$

$$\epsilon\dot{\mathbf{z}}(t) = \mathbf{C}_1\mathbf{x}(t) + \mathbf{C}_2\mathbf{z}(t),\ \mathbf{z}(0) = \mathbf{z}_0(0) \tag{5b}$$

where $\mathbf{C}_1$ *and* $\mathbf{C}_2$ are the following block matrices:

$$\mathbf{C}_1 = \begin{bmatrix} \mathbf{I} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \tag{6}$$

$$\mathbf{C}_2 = \begin{bmatrix} -I & 0 & . & . & 0 \\ I & -I & 0 & & . \\ 0 & I & -I & & . \\ & & \cdots & & . \\ & \cdot & \cdots & & . \\ & \cdot & & \cdots & 0 \\ 0 & . & . & 0 \ I & -I \end{bmatrix}$$

$$\mathbf{z} = \left[\mathbf{z}^1 \vdots\ \mathbf{z}^2 \vdots \cdots \vdots\ \mathbf{z}^m\right]',\ \mathbf{z}(0) = \left[\mathbf{x}_0(-\epsilon) \vdots \cdots \vdots\ \mathbf{x}_0(-m\epsilon)\right]' \tag{7}$$

$\mathbf{I} = \mathbf{I}_n$ is an $n \times n$ identity matrix and $\mathbf{A}$, $\mathbf{A}_1$ and $\mathbf{B}$ are matrices defined earlier. From (6) it is clear that $\mathbf{C}_2$ is a nonsingular matrix and hence the singularly-perturbed system can be reduced to a lower-order (slow subsystem) system by letting $\epsilon \to 0$, *i.e.*,

$$\dot{\mathbf{x}} = \mathbf{A}_0\mathbf{x} + \mathbf{A}_1\mathbf{z}^m + \mathbf{B}\ \mathbf{u} \tag{8a}$$

$$0 = \mathbf{C}_1\mathbf{x} + \mathbf{C}_2\ \mathbf{z} \tag{8b}$$

Eliminating $\mathbf{z}$ from algebraic equation (8b) in terms of $\mathbf{x}$, *i.e.*,

$$\mathbf{z} = -\ \mathbf{C}_2^{-1}\ \mathbf{C}_1\ \mathbf{x} = \left[\mathbf{x}\ \mathbf{x} \cdots \mathbf{x}\right]' \tag{9}$$

and noting (7) yields, $\mathbf{z}^m = \mathbf{x}$, hence the reduced-order model (8a) is given by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{10a}$$

$$\mathbf{x}(0) = \mathbf{x}_0(0) \tag{10b}$$

where $\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1$. Thus, the TD system of (1) can be approximated by an unretarded system (10). Following the theory of singular perturbation for near-optimum control design [7.13] and the discussion of Section 7.2, the near-optimum control

$$\mathbf{u}(t,\epsilon) = \mathbf{u}^r(t,0) + \left.\frac{\partial \mathbf{u}^r(t,\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} \epsilon \tag{11}$$

can be obtained for the approximate system (10) as well as for the original TD system (1). Using Tupciev Theorem [7.25] the continuity and differentiability of $\mathbf{u}^r(t,\epsilon)$ *w.r.t.* ϵ at $\epsilon=0$ can be checked. It is easily found that the reduced-order (zeroth-order) optimal control $\mathbf{u}^r(t,0)$ coincides with $\mathbf{u}^*(t,0)$ - the zeroth-order term of the MacLaurin's expansion of the optimal control.

It should, however, be noted that in some applications where the initial transients are of great importance, the loss of initial conditions of $\mathbf{z}^i(0)$, $i=1, \ldots, m$ must be compensated by computing the left "boundary layer" of the singularly-perturbed system (5). For a discussion on this subject the interested reader can consult Reference [7.26].

The second term of expansion in (11) can be obtained from [7.14]:

$$\left.\frac{\partial \mathbf{u}^r(t,\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{p}^{(1)}(t) \tag{12}$$

where $\mathbf{p}^{(1)}(t)$ is the solution of the following *TPBV* problem (see Equation 7.2.14):

$$\dot{\mathbf{x}}^{(1)}(t) = \mathbf{A}\mathbf{x}^{(1)}(t) - \mathbf{S}\,\mathbf{p}^{(1)}(t) - m\mathbf{A}_1\dot{\bar{\mathbf{x}}} \tag{13a}$$

$$\dot{\mathbf{p}}^{(1)}(t) = -\mathbf{Q}\,\mathbf{x}^{(1)}(t) - \mathbf{A}'\mathbf{p}^{(1)}(t) - m\mathbf{A}'_1\dot{\bar{\mathbf{p}}} \tag{13b}$$

$$\mathbf{x}^{(1)}(0) = \mathbf{A}_1\left[\sum_{i=0}^{m} \mathbf{x}_0(-i\epsilon) - m\mathbf{x}_0(0)\right] \tag{14a}$$

$$\mathbf{p}^{(1)}(t_f) = \mathbf{0} \tag{14b}$$

where

$$\bar{\mathbf{x}}(t) = \lim_{h\to 0} \mathbf{x}(t,h),\ 0 \leqslant t \leqslant t_f \tag{15a}$$

$$\bar{\mathbf{p}}(t) = \lim_{h\to 0} \mathbf{p}(t,h),\ 0 \leqslant t \leqslant t_f \tag{15b}$$

1. Example. Consider a first-order linear TD system which is a special case of an example considered by Inoue *et al.* [7.14]:

$$\dot{x}(t) = x(t) + 0.5x(t-0.1) + u(t) \tag{16}$$

with a quadratic cost functional:

$$J(u) = 1/2 \int_0^2 (5x^2 + u^2)dt \tag{17}$$

For a two-subinterval division of the delay $h=0.1$, (16) is rewritten in a singularly-perturbed fashion as

$$\dot{x} = x + 0.5z^2 + u, \quad x(0) = 1.0 \tag{18}$$

$$\epsilon\dot{z}^1 = x - z^1, \quad z^1(0) = 0.15 \tag{19a}$$

$$\epsilon\dot{z}^2 = z^1 - z^2, \quad z^2(0) = 0.20 \tag{19b}$$

The reduced (zeroth-) order version of (18)-(19) is given by

$$\dot{x} = x + 0.5z^2 + u = 1.5x + u \tag{20a}$$

$$z^1 = z^2 = x \tag{20b}$$

The near-optimal control provided a cost functional which is given by Figure 1. Δ

A detailed design example dealing with a 3-stand, 2-delay cold rolling mill system using the above procedure is given in Section 9.2.

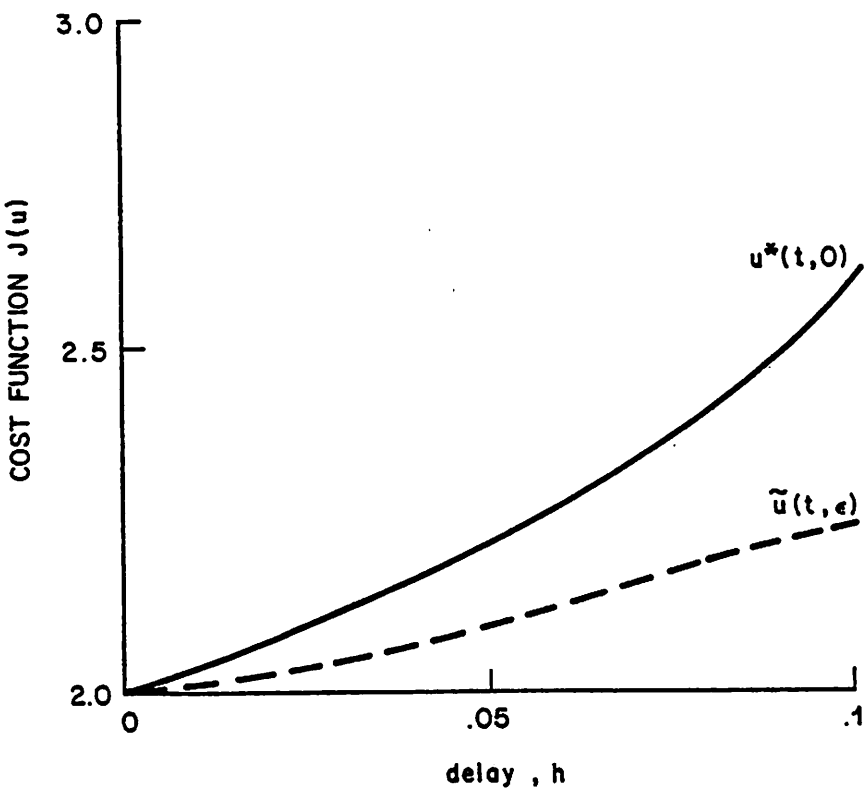

Figure 1. A Comparison of Performance Indices for Reduced-Order and Near-Optimum Control

7.6 SENSITIVITY TO PARAMETER VARIATIONS

The parameters of physical systems are often subject to variations. Such variations may occur because of uncertainties in the state measurement or in the system parameters, or because of the choice of mathematical model [7.27]. The sensitivity of an optimally designed system to variations in the system parameters is of interest. In the case of nondelay systems, this problem has been investigated by several authors [7.27-7.31]. Dorato [7.28] studied performance index sensitivity to small plant parameter variations by introducing the *performance index sensitivity vector*. The effects of parameter variations on the output of linear and nonlinear feedback systems were studied by Cruz and Perkins who introduced the concept of *comparison sensitivity function* [7.29]. Werner and Cruz [7.12] showed a relation between approximations in optimal control and the optimal performance index of the system under the influence of uncertain parameters. They showed that a truncation up to the nth term of the series expansion of control in plant parameters corresponds to a truncation up to the $(2n+1)$st term of the series expansion of the optimal cost functional. Zinober and Fuller [7.30] considered the effect of plant parameter variations on a system with nominally time-optimal feedback controller and showed that small deviations of plant parameters from their nominal values cause instability, and as the order of the system increases, the deviation that takes the system to the verge of instability becomes smaller.

McClamroch *et al.* [7.27] studied the sensitivity of optimal cost functional to large plant parameter variations in linear systems by introducing the concept of ρ-sensitivity. They called a system ρ-sensitive *w.r.t.* a certain "class" of variations if the value of the cost functional did not increase by more than a factor of ρ for variations in that class. This concept was later utilized by Malek-Zavarei and Jamshidi [7.31] to investigate the sensitivity of linear TD systems to parameter variations.

In this section, which is based on the results of Reference [7.31], we will study the sensitivity of the cost functional *w.r.t.* parameter variations in the state equation of optimally designed linear time-delay systems. We will consider two cases depending on the nature of these variations. In one case the delay portion of the state equations is ignored completely, while in the other, the delay itself is set equal to zero. For both types of variations we will obtain conditions for ρ-sensitivity. We will also consider the effects of parameter variations on the stability of the perturbed system.

7.6.1 Sensitivity With Respect To Plant Parameter Variations

Consider a *l.t.i.* nondelay system described by the state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{1}$$

where $\mathbf{x}$ and $\mathbf{u}$ are the state and the control vectors, respectively. Assume that due to plant parameter variations, system (1) changes to a TD system described by the following state equation:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{E}\mathbf{x}(t-h) + \mathbf{B}\mathbf{u}(t), \tag{2a}$$

$$\mathbf{x}(t) = \phi(t), \ -h \leqslant t \leqslant 0, \tag{2b}$$

In the above state equation, h is the time delay which is assumed to be known and matrix $\mathbf{E}$ represents the error (which is not known). Assume that it belongs to some class of errors ξ where $0 \epsilon \xi$ so that the possibility of an errorless system is also included. The cost functional, which is to be minimized, is

$$J = \int_0^\infty [\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)]dt, \tag{3}$$

where the constant matrices $\mathbf{Q}$ and $\mathbf{R}$ satisfy the usual state regulator conditions [7.15]. Of course, when system (1) changes to system (2), the optimal control and the value of the optimal cost functional also change accordingly. The objective is to find the relation between changes in system (1) and the corresponding changes in the value of the optimal cost functional. For physical reasons, there is no need to assume any change in the matrices $\mathbf{Q}$ and $\mathbf{R}$ [7.27].

Since the introduction of the error matrix $\mathbf{E}$ gives rise to a time-delay system, the knowledge of its initial function is essential for sensitivity analysis. Since the error matrix $\mathbf{E}$ is not known, the initial function $\phi(t)$ must be approximated. The optimal control for system (1) with cost functional (3) is

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{B}'\mathbf{W}_o\mathbf{x}(t) \tag{4}$$

where $\mathbf{W}_o$ is the solution to the Riccati equation [7.15]

$$\mathbf{A}'\mathbf{W}_o + \mathbf{W}_o\mathbf{A} - \mathbf{W}_o\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{W}_o + \mathbf{Q} = 0. \tag{5}$$

Hence, assuming the system has been running up to $t = 0$, a justifiable approximation would be

$$\phi(t) = exp[(\mathbf{A}-\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{W}_o)t]\mathbf{x}_o \tag{6}$$

In order to characterize the relation between changes in the optimal cost functional and the error matrix $\mathbf{E}$, the concept of ρ-sensitivity as defined in [7.27] is extended to linear TD systems by the following definition.

1. Definition. For some real number ρ and some class of errors ξ, system (2) and cost functional (3) are said to be *ρ-sensitive* if for each initial state $\mathbf{x}_0$, we have $J_e \leqslant \rho J_0$ for all matrices $\mathbf{E}\epsilon\xi$ where J_0 and J_e denote the optimal values of cost functional J in (3) evaluated, respectively, along the trajectories of (1) and (2a) with initial function (6). Δ

The following theorem establishes the relation between changes in the optimal cost functional and the error matrix **E**.

2. Theorem. System (2a) with initial function (6) and cost functional (3) is ρ-sensitive in the sense of Definition 1 if and only if for each error matrix $\mathbf{E}\epsilon\xi$, the matrix $\rho\mathbf{W}_0 - \mathbf{W}_1(\mathbf{E})$ is positive semi-definite, where $\mathbf{W}_0$ satisfies (5) and

$$\mathbf{W}_1(\mathbf{E}) = \mathbf{w}_1 + 2\int_{-h}^{0}\mathbf{w}_2(s)\ exp[(\mathbf{A}-\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{W}_0)s]ds + \int_{-h}^{0}\int_{-h}^{0} exp[(\mathbf{A}'-\mathbf{W}_0\mathbf{B}\mathbf{R}^{-1}\mathbf{B}')s]\ exp[(\mathbf{A}-\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{W}_0)r]\mathbf{w}_3(r,s)\ drds, \tag{7}$$

$$\mathbf{A}'\mathbf{w}_1 + \mathbf{w}_1\mathbf{A} - \mathbf{w}_1\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{w}_1 + \mathbf{w}_2(0) + \mathbf{w}'_2(0) + \mathbf{Q} = 0, \tag{8}$$

$$-\frac{d\mathbf{w}_2(s)}{ds} + \mathbf{A}'\mathbf{w}_2(s) - \mathbf{w}_1\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{w}_2(s) + \mathbf{w}'_3(0,s) = 0, \quad -h \leqslant s \leqslant 0, \tag{9}$$

$$\frac{\partial\mathbf{w}_3(r,s)}{\partial r} + \frac{\partial\mathbf{w}_3(r,s)}{\partial s} + \mathbf{w}'_2(r)\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{w}_2(s) = 0, \quad -h \leqslant r \leqslant 0,\ -h \leqslant s \leqslant 0, \tag{10}$$

with boundary conditions

$$\mathbf{w}_1\mathbf{E} = \mathbf{w}_2(-h) \tag{11}$$

and

$$\mathbf{E}'\mathbf{w}_2(s) = \mathbf{w}_3(-h,s), \quad -h \leqslant s \leqslant 0. \tag{12}$$

Proof. For the errorless system ($\mathbf{E} = 0$), the optimal value of the cost functional (3) is

$$J_0 = \mathbf{x}'_0\mathbf{W}_0\mathbf{x}_0 \tag{13}$$

where $\mathbf{W}_0$ satisfies (5). When $\mathbf{E} \neq 0$, the optimal cost functional has the form

$$J_e = x'_0 w_1 x_0 + 2x'_0 \int_{-h}^{0} w_2(s)x(s)ds + \int_{-h}^{0}\int_{-h}^{0} x'(s)w_3(r,s)x(r)drds, \tag{14}$$

where matrices w_1, w_2 and w_3 satisfy relations (8)-(12). (See Theorem 6.4.1.) Using (6), one can write

$$J_e = x'_0 W_1(E) x_0 \tag{15}$$

where $W_1(E)$ is given by (7). Hence $J_e \leqslant \rho J_0$ if and only if the matrix $\rho W_0 - W_1(E)$ is positive semi definite.

3. **Example.** Consider the *l.t.i.* system

$$\dot{x} = -2x(t) + u(t), \; x(0) = 1, \tag{16}$$

with cost functional

$$J = \int_0^\infty (5x^2+u^2)dt. \tag{17}$$

Assume that because of plant parameter variations it changes to the TD system

$$\dot{x} = -2x(t) + \alpha x(t-1) + u(t), \tag{18}$$

where α represents the error. The assumed initial function for (18) will be

$$\phi(t) = exp[(-2-W_0)t] \tag{19}$$

where W_0 is the solution to

$$-4W_0 - W_0^2 + 5 = 0. \tag{20}$$

The acceptable solution to the above equation is $W_0 = 1$; hence $\phi(t) = exp(-3t)$.

By Theorem 2, system (18) with cost functional (17) is ρ-sensitive in the sense of Definition 1 if and only if the error α satisfies

$$\rho W_0 - W_1(\alpha) \geqslant 0 \text{ or } \rho \geqslant W_1(\alpha), \tag{21}$$

where

$$W_1(\alpha) = w_1 + 2\int_{-1}^{0} w_2(s)\ exp(-3s)ds$$
$$+ \int_{-1}^{0}\int_{-1}^{0} w_3(r,s)\ exp[-3(r+s)]\ drds, \tag{22}$$

$$-\ 4w_1 - w_1^2 + 2w_2(0) + 5 = 0, \tag{23}$$

$$\frac{-dw_2(s)}{ds} - (2 + w_1)w_2(s) + w_3(0,s) = 0,\ -1 \leqslant s \leqslant 0, \tag{24}$$

$$\frac{\partial w_3(r,s)}{\partial r} + \frac{\partial w_3(r,s)}{\partial s} + w_2(r)w_2(s) = 0,\ -1 \leqslant r \leqslant 0,\ -1 \leqslant s \leqslant 0, \tag{25}$$

with boundary conditions

$$\alpha w_1 = w_2(-1) \tag{26}$$

and

$$\alpha w_2(s) = w_3(-1,s),\ \ -1 \leqslant s \leqslant 0. \tag{27}$$

This boundary value problem was solved by finite difference method and $W_1(\alpha)$ was calculated for various values of α. (See Figure 1.) Note that for each given value of error α, the system under consideration is ρ-sensitive in the sense of Definition 1 for values of ρ greater than or equal to $W_1(\alpha)$. For example, for $\alpha = 0.2$ the system is ρ-sensitive for $\rho \geqslant 0.5$ as indicated in Figure 1. Also, for a given value of ρ one can determine the class of errors for the system to be ρ-sensitive. For example, using Figure 1, for $\rho = 0.5$ the class of errors is found to be $\xi = \{\alpha\colon \alpha \leqslant 0.2\}$. Note that for the errorless system (α=0), $W_1(\alpha) = 1$ as expected.

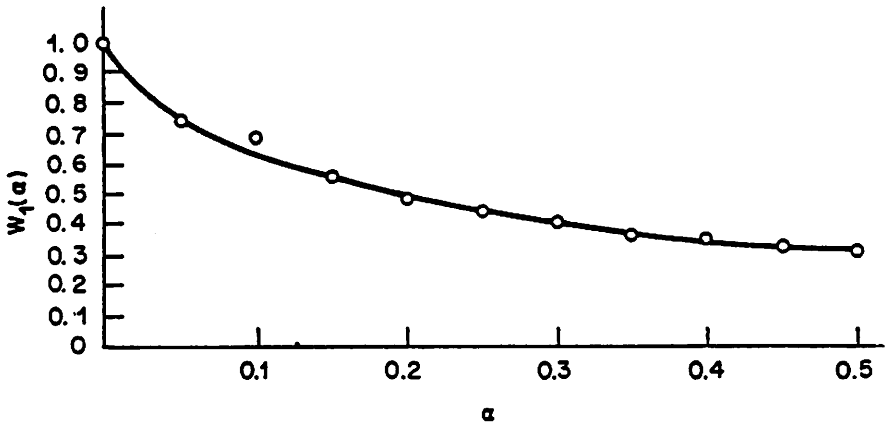

Figure 11. The Plot of $W_1(\alpha)$ Versus α for Example 1

7.6.2 Sensitivity with respect to Approximate Modeling

Uncertainties in the performance of the system may also arise because of the choice of a mathematical model for the system. Often it is computationally convenient to model a linear system with inherent small time delays as a linear nondelay system. More precisely, suppose that the system under consideration is described by state equation (2), and assume that it is approximated by the following state equation:

$$\dot{\mathbf{x}} = (\mathbf{A}+\mathbf{E})\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{x}(0) = \phi(0). \tag{28}$$

In this case, delay h represents the error and again it is desirable to determine the degree of deterioration in the performance of the original system due to this error. Note that matrix $\mathbf{E}$ is known in this case. Assume that delay h belongs to some appropriate error interval $[0,h_j]$. Again the objective is to determine the relation between changes in system (2) and the corresponding changes in the optimal value of cost functional (3).

Since the delay h is small, the initial function (2b) can be reasonably assumed to be a constant, *i.e.*,

$$\mathrm{x}(t) = \phi(0), \quad -h \leqslant t \leqslant 0. \tag{29}$$

Clearly, the degree of validity of this approximation depends on the function $\phi(t)$ and the magnitude of the delay h. In this case ρ-sensitivity is defined as follows.

4. Definition. For some real number ρ and some error interval $[0,h_1]$, system (28) and cost functional (3) are said to be *ρ-sensitive* if for each initial state x_0, we have $J_e \leqslant \rho J_0$ for all $T\epsilon[0,h_1]$ where J_e and J_0 denote the optimal values of cost functional J in (3) evaluated, respectively, along the trajectories of (28) and (2a) with initial function (29). Δ

The following theorem establishes the relation between changes in the optimal cost functional and error α.

5. Theorem. System (28) with cost functional (3) is ρ-sensitive in the sense of Definition 4 if and only if for each delay $T\epsilon[0,h_1]$ with h_1 sufficiently small, the matrix $\mathrm{S}_0 - \rho\mathrm{S}_1(h)$ is positive semi definite, where

$$(\mathrm{A}+\mathrm{E})'\mathrm{S}_0 + \mathrm{S}_0(\mathrm{A}+\mathrm{E}) - \mathrm{S}_0\mathrm{B}\mathrm{R}^{-1}\mathrm{B}'\mathrm{S}_0 + \mathrm{Q} = 0, \tag{30}$$

$$\mathrm{S}_1(h) = \mathrm{w}_1 + 2\int_{-h}^{0}\mathrm{w}_2(s)ds + \int_{-h}^{0}\int_{-h}^{0}\mathrm{w}_3(r,s)drds, \tag{31}$$

and matrices w_1, w_2 and w_3 satisfy (8)-(12).

Proof. When delay h is ignored, the optimal value of cost functional (3) is

$$J_e = \mathrm{x}'\mathrm{S}_0\mathrm{x}_0 \tag{32}$$

(32) where S_0 satisfies (30). Also, similar to the argument used in the proof of Theorem 2, the optimal value of cost functional (3) along the trajectory of system (2a) with initial function (29) is $J_0 = \mathrm{x}'_0\mathrm{S}_1(h)\mathrm{x}_0$ where $\mathrm{S}_1(h)$ is given by (31). Hence $J_e \leqslant \rho J_0$ if and only if matrix $\mathrm{S}_0 - \rho\mathrm{S}_1(h)$ is positive semi definite.

6. Example. Consider the system

$$\dot{x} = -x(t) - x(t-h) + u(t),$$
$$x(t) = e^t, \quad -h \leqslant t \leqslant 0 \tag{33}$$

with cost functional (17). Assume that it is approximated by the *l.t.i.* system

$$\dot{x} = -2x(t) + u(t), \quad x(0) = 1. \tag{34}$$

Here, delay h represents the error. By Theorem 5, system (34) with cost functional (17) is ρ-sensitive in the sense of Definition 4 if and only if delay h is such that

$$S_0 - \rho S_1(h) \geqslant 0 \tag{35a}$$

where

$$-4S_0 - S_0^2 + 5 = 0, \tag{35b}$$

which yields $S_0 = 1$, and $S_1(h)$ is given by (31). The scalar functions w_1, w_2 and w_3 satisfy

$$-2w_1 - w_1^2 + 2w_2(0) + 5 = 0, \tag{36}$$

$$-\frac{\partial w_2(s)}{ds} - (1+w_1)w_2(s) + w_3(0,s) = 0, \quad -h \leqslant s \leqslant 0, \tag{37}$$

$$\frac{\partial w_3(r,s)}{\partial r} + \frac{\partial w_3(r,s)}{\partial s} + w_2(r)w_2(s) = 0, \quad -h \leqslant r \leqslant 0, \quad -h \leqslant s \leqslant 0, \tag{38}$$

with boundary conditions

$$-w_1 = w_2(-h), \quad -w_2(s) = w_3(-h,s), \quad -h \leqslant s \leqslant 0. \tag{39}$$

The above boundary value problem was solved by the finite difference method to find the relationship between ρ and h. The result is shown in Figure 2 where $1/S_1(h)$ is plotted versus h. Note that for each given value of error h, the system under consideration is ρ-sensitive in the sense of Definition 4 for values of ρ less than or equal to $1/S_1(h)$. For example, for $h = 0.5$ the system is ρ-sensitive for $\rho \leqslant 1.1$. Similarly, for a given value of ρ, one can determine the error interval for the system to be ρ-sensitive. For example, for $\rho = 1.1$ the error interval is $[0,0.5]$.

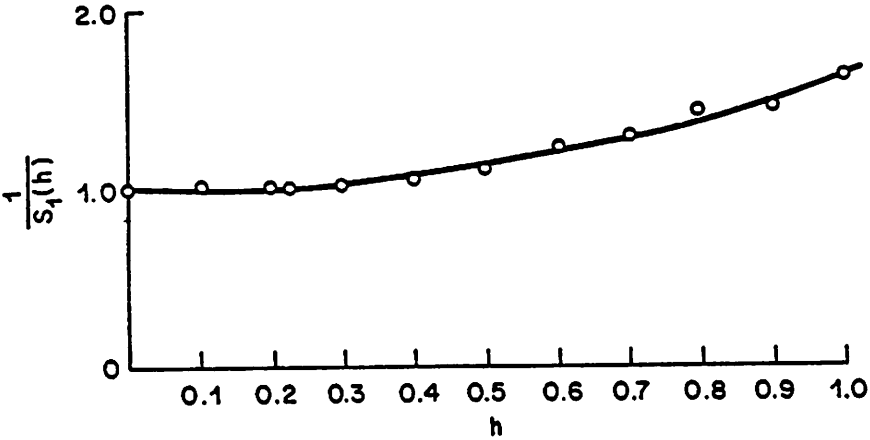

Figure 12. The Plot of $1/S_1(h)$ Versus h for Example 6

7.6.3 Stability Considerations

The plant parameter variations of the type discussed in Section 7.4.1 may cause instability [7.24]. That is, even though system (1) is stable, system (2) may be unstable for certain values of $\mathbf{E}$ and h. It would be desirable to find conditions on error matrix $\mathbf{E}$ and/or delay h such that the perturbed system remains stable. Also for the type of variations discussed in Section 7.4.2, conditions on delay h can be found in order to preserve stability. We will discuss the relation between the stability of the original and the perturbed systems in this section.

Suppose that the uniformly asymptotically stable system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t), \ \ \mathbf{x}(0) = \mathbf{x}_0, \tag{40}$$

changes to

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{E}x(t-h), \tag{41a}$$

$$\mathbf{x}(t) = \phi(t), \ \ -h \leqslant t \leqslant 0, \tag{41b}$$

because of plant parameter variations. The cost functional is given by (3). Consider the Lyapunov equation

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -2\mathbf{H} \tag{42}$$

It has been shown [7.33] that if matrix **H** in (42) is positive definite, and if matrix **P** is such that $\mathbf{H} - \mathbf{P}\mathbf{E}\mathbf{Q}^{-1}\mathbf{E}'\mathbf{P}$ is positive definite, then the solution of the perturbed system is also uniformly asymptotically stable independent of the magnitude of delay h. Application of this result to Example 3 yields the sufficient condition $-2 < \alpha < 2$ for uniform asymptotic stability of the perturbed system. Conditions have also been obtained on both the error matrix **E** and the delay h for stability of the perturbed system [7.34,7.35].

Note that if for some finite value of ρ and for all admissible **E**, the conditions of Theorem 2 are satisfied, then with the usual positive-definiteness assumption on matrices **Q** and **R** in cost functional (3) the system must be stable since J_e will be bounded in this case.

For the type of parameter variations discussed in Section 7.4.2, it has been shown [7.35] that asymptotic stability of system (28) implies asymptotic stability of system (2) for sufficiently small values of delay h.

PROBLEMS

7.1. Consider the system

$$\dot{x} = x(t-1) + u, \quad 0 < t \leqslant 2$$
$$x(t) = 1, \quad -1 \leqslant t \leqslant 0,$$

and the cost functional

$$J = ½ \left[10^5 x^2(2) + \int_0^2 u^2(t)dt\right].$$

Use the method of Section 7.2 to determine the second through fifth-order suboptimal controls for this system and calculate the value of the cost functional in each case. (Note: The exact solution to this problem has been obtained analytically) [7.17].

7.2. For the problem considered in Example 7.2.1, determine the second and third-order suboptimal controls and compare them with the first-order suboptimal control.

7.3. Consider the linear time-varying system with multiple state and control delays, free final state and a quadratic cost functional. Develop the detailed procedure for determining an Nth-order suboptimal control for this system using the sensitivity approach.

7.4. Repeat Problem 7.3 for the case where the final state is specified.

7.5 Consider the second-order linear time-varying multi-delay system described in Example 7.2.2 where $h_{x_2} = 1$ instead of 0.8. Determine the second through tenth-order suboptimal controls and cost functionals and compare them with those in Example 7.2.2.

7.6. Repeat Problems 7.1 and 7.2 using the method of Section 7.3.

7.7. Consider the first-order nonlinear TD system

$$\dot{x}(t) = x(t) + x_1^2(t-1) + 2u(t), \quad 0 < t \leqslant 3$$
$$x(t) = 0.5, \quad -1 \leqslant t < 0$$

and the cost functional

$$J = \int_0^3 [x^2(t)+u^2(t)] \; dt$$

Use the method of Section 7.3 to determine the second through fifth-order suboptimal controls for this system. Develop a table similar to Table 7.3.1 for this problem.

7.8. Repeat Problem 7.5 using the method of Section 7.3.

7.9. Provide detailed justification for the choice of the approximate initial function (7.6.6) for system (7.6.2).

7.10. Show that the system described by state equation (7.6.28) with cost functional (7.6.3) is ρ-sensitive in the sense of Definition 7.6.1 if and only if for each error matrix $\mathbf{E}$, the matrix $\rho\mathbf{W}_0 - \mathbf{w}_1$ is positive semi definite where $\mathbf{W}_0$ satisfies (7.6.5) and $\mathbf{w}_1$ satisfies

$$(\mathbf{A}'+\mathbf{E}')\mathbf{w}_1 + \mathbf{w}_1(\mathbf{A}+\mathbf{E}) - \mathbf{w}_1\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{w}_1 + \mathbf{Q} = 0.$$

(This result is essentially Theorem 1 of [7.27] for nonautonomous *l.t.i.* systems.)

7.11. Consider a system represented by state equations (7.6.1) where

$$\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

with cost functional (7.6.3) where $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = 2\mathbf{I}$. Assume that the state equation of this system changes to (7.6.2) with $h = 0.5$. Find conditions on error matrix $\mathbf{E}$ for this system to be ρ-sensitive. Discuss the stability of the perturbed system.

7.12. Consider the system described by state equation (7.6.2) where

$$\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \phi(t) = \begin{bmatrix} e^t \\ 1 \end{bmatrix}, \quad h = 0.5$$

with cost functional (7.6.3) where $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = 2\mathbf{I}$. Assume that this system is approximately modeled with $h = 0$. Find conditions on h for ρ-sensitivity of this system.

7.13. Consider a scalar nonlinear TD system

$$\dot{x} = -x^2 + x(t-0.1) + x(t)u(t) + u(t-0.1)$$

with initial functions $x(t) = 1$, $u(t) = 0$ $-0.1 \leq t \leq 0$. Use the method of Algorithm 7.4.2 to find a near-optimum control for a cost functional $J = 1/2 \int_0^1 (u^2(t) + 1)dt$.

7.14. Repeat example 7.6.6 for system of Problem 7.7 and a cost functional $J = 1/2 \int_0^2 u^2(t)dt$.

7.15. For system of Problem 7.1 find a singularly-perturbed equivalent model using $m=2$ and $m=5$ subdivisions for the delay $h=1$.

7.16. Using the sensitivity method of Section 7.5 find a near-optimum control for system of Problem 7.15 and $m=2$.

REFERENCES

[7.1] M. Jamshidi and M. Malek-Zavarei, "Suboptimal design of linear control systems with time delay," *Proc. of the IEE*, Vol. 119, pp. 1743-1746, 1972

[7.2] H. C. Chan and W. R. Perkins, "Optimization of time delay systems using parameter imbedding," *IFAC J. Automatica*, Vol. 9, pp. 257-261, 1973

[7.3] M. Malek-Zavarei, "Near optimum design of nonstationary linear systems with state and control delays," *Journal of Optimization Theory and Applications*, Vol. 30, p. 73-88, 1980

[7.4] S. A. Gracovetsky and M. Vidyasagar, "Suboptimal control of neutral systems," *Int. J. of Control*, Vol. 18, pp. 121-128, 1973

[7.5] M. Malek-Zavarei, "Suboptimal control of systems with multiple delays," *J. of Optimization Theory and Applications*, Vol. 30, pp. 621-633, 1980

[7.6] R. R. Bate, "The optimal control of systems with transport lag," *Advances in Control Systems*, Vol. 7, Edited by C. T. Leondes, Academic Press, 1969

[7.7] G. L. Slater and W. R. Wells, "On the reduction of optimal time delay systems to ordinary ones," *IEEE Trans. on Aut. Control*, Vol. AC-17, pp. 154-155, 1972

[7.8] F. Ficken, "The continuation method for functional equations," *Communications on Pure and Applied Mathematics*, Vol. 4, pp. 435-436, 1951

[7.9] S. Roberts and J. Shipman, "Continuation in shooting methods for two-point boundary-value problems," *J. of Mathematical Analysis and Applications*, Vol. 18, pp. 45-58, 1967

[7.10] K. H. Kagiwada and R. F. Kalaba, "Derivation and validation of an initial-value method for certain nonlinear two-point boundary-value problems," *J. of Optimization Theory and Applications*, Vol. 2, pp. 378-385, 1968

[7.11] D. Glasser, "Numerical solution of two-point boundary-value problems on total difference equations," *SIAM J. on Numerical Analysis,* Vol. 6, pp. 591-597, 1969

[7.12] R. A. Werner and J. B. Cruz, "Feedback control which preserves optimally for systems with unknown parameters," *IEEE Trans. on Aut. Control,* Vol. AC-13, pp. 621-629, 1968

[7.13] P. Sannuti and P. Kokotovic', "Near-optimum design of linear systems by a singular perturbation method," *IEEE Trans. on Aut. Control,* Vol. AC-14, pp. 15-21, 1969

[7.14] K. Inoue, H. Akashi, K. Ogino, and Y. Sawaragi, "Sensitivity approaches to optimization of linear systems with time delay," *IFAC J. Automatica,* Vol. 7, pp. 671-679, 1971

[7.15] R. E. Kalman, "Contribution to the theory of optimal control," *Bol. Soc. Mat. Mexicana,* pp. 102-119, 1960

[7.16] D. H. Eller, J. K. Aggarwal, and H. T. Banks, "Optimal control of linear time delay systems," *IEEE Trans. on Aut. Control,* Vol. AC-14, pp. 678-687, 1969

[7.17] T. E. Mueller, "Optimal control of linear systems with time lag," University of Illinois, Urbana, Illinois, Coordinated Science Laboratory, *Report No. R-254,* 1965

[7.18] M. Jamshidi, "An overview on the algebraic matrix Riccati equation and related problems," *J. Large - Scale Systems, Vol. 1, pp. 167-192, 1980*

[7.19] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach,* Pergamon Press, 1986

[7.20] A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis,* 2nd Ed., John Wiley and Sons, Inc., New York, 1980

[7.21] M. Jamshidi and M. Malek-Zavarei, "A hierarchical control for large-scale time-delay systems," *J. Large Scale Systems,* Vol. 4, pp. 149-163, 1983

[7.22] M. Jamshidi, "A three - stage design of non-linear control system with time-delay," *Int. J. Control* Vol. 21, pp. 753 - 762, 1975

[7.23] P. Sannuti, "Near - optimum design of time-lag systems by singular perturbation method," *Preprints 11th JACC,* pp. No. 20-A, Atlanta, GA, June, 1970

[7.24] M. Jamshidi, "A near-optimum controller for cold - rolling mills," *Int. J. Control,* Vol. 16, No. 6, pp. 1137-1154, 1972

[7.25] V. A. Tupciev, "Asymptotic behavior of the solution of a boundary value problem for systems of differential equations of first order with a small parameter in the derivative," *Dokl. Akad Nauk SSSR,* Vol. 143, pp. 1296-1299, 1962

[7.26] M. Jamshidi, *Large-Scale Systems: Modeling and Control,* Elsevier North Holland Publishing Co., pp. 39-40, 1983

[7.27] N. H. McClamroch, L. G. Clark and J. K. Aggarwal, "Sensitivity of linear control systems to large parameter variations," *IFAC J. Automatica,* Vol. 5, pp. 257-263, 1969

[7.28] P. Dorato, "On sensitivity in optimal control systems," *IEEE Trans. Aut. Control,* Vol. AC-8, pp. 256-257, 1963

[7.29] J. B. Cruz, Jr. and W. R. Perkins, "Criteria for system sensitivity to parameter variations," *Proc. 3rd IFAC Cong.,* pp. 18C.1-18C.7, London, 1966

[7.30] A. S. I. Zinober and A. T. Fuller, "The sensitivity of nominally time-optimal control systems to parameter variations," *Int. J. Control,* Vol. 17, pp. 673-703, 1973

[7.31] M. Malek-Zavarei and M. Jamshidi, "Sensitivity of linear time-delay systems to parameter variations," *IFAC J. Automatica,* Vol. 11, pp. 315-319, 1975

[7.32] G. J. Nazaroff, "Stability and stabilization of linear differential delay systems," *IEEE Trans. Aut. Control,* Vol. AC-18, pp. 317-318, 1973

[7.33] R. A. Baker and A. R. Bergen, "Lyapunov stability and Lyapunov functions of infinite dimensional systems," *IEEE Trans. Aut. Control*, Vol. AC-14, pp. 325-334, 1969

[7.34] F. L. N-Nagy and M. N. Al-Tikriti, "Stability criterion of linear control systems with time delay," *Meas. Control*, Vol. 3, March 1970

[7.35] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags*, Academic Press, 1966

CHAPTER 8

NEAR-OPTIMUM DESIGN OF LARGE-SCALE TIME-DELAY SYSTEMS

8.1 INTRODUCTION

In Chapters 6 and 7 we introduced the optimal control and optimization of TD systems in detail. The history of this development can be found in references [8.1-8.18].

Large-scale systems involving time delays have been the subject of recent investigation by some authors [8.19-8.22]. Jamshidi and Malek-Zavarei [8.22,8.23] considered the suboptimal control of a large-scale TD system. They decomposed the system into several subsystems and applied a procedure similar to that of Malek-Zavarei's [8.18] to determine the suboptimal control. (It is noted that in most large-scale systems a natural decomposition exists which can be applied.) This method produces a suboptimal control which is partly closed-loop and partly open-loop.

In this chapter, three approaches for near-optimum control of large-scale TD systems will be introduced. The first is based on the decomposition of linear and nonlinear systems with delay based on the approach due to Jamshidi [8.10]. The next approach deals with the hierarchical control of a serial TD system in which the delayed state of one subsystem acts as an input to the next one down the line and so on [8.20]. The last approach is the application of a hierarchical control algorithms (goal coordination, interaction prediction and costate prediction) to large-scale TD systems due in part to Jamshidi and Malek-Zavarei [8.23], Jamshidi and Wang [8.24,8.25], Jamshidi and Merryman [8.26], Breadeau and Jamshidi [8.27].

8.2 NEAR-OPTIMUM CONTROL OF COUPLED TIME-DELAY SYSTEMS

Consider a large-scale TD systems consisting of a set of N coupled subsystems:

$$
\begin{aligned}
\dot{\mathbf{z}}_1 &= \mathbf{h}_1(\mathbf{z}_1,\mathbf{w}_1,\epsilon c_1(\mathbf{z},\mathbf{z}(t-h),\mathbf{w},\mathbf{w}(t-h),t),t)\\
\dot{\mathbf{z}}_2 &= \mathbf{h}_2\ (\mathbf{z}_2,\mathbf{w}_2,\epsilon c_2(\mathbf{z},\mathbf{z}(t-h),\mathbf{w},\mathbf{w}(t-h),t),t)\\
&\ \ \vdots\\
\dot{\mathbf{z}}_N &= \mathbf{h}_N\ (\mathbf{z}_N,\mathbf{w}_N,\epsilon c_N(\mathbf{z},\mathbf{z}(t-h),\mathbf{w},\mathbf{w}(t-h),t),t)
\end{aligned}
\tag{1}
$$

where $\mathbf{z} = [\mathbf{z}'_1\ \mathbf{z}'_2 \cdots \mathbf{z}'_N]'$ and $\mathbf{w} = [\mathbf{w}'_1\mathbf{w}'_2 \cdots \mathbf{w}'_N]'$ are state and control vectors, respectively, ϵ is a scalar coupling parameter, and h is the delay, not necessarily small. The problem is to find a control function $\mathbf{w}$ composed of $\mathbf{w}_i$, $i=1,2,\ldots,N$ which satisfies (1) with appropriate boundary conditions and minimizes a separable cost functional

$$J = \int_{t_0}^{t_f} \sum_{i=1}^{N} L_i\ (\mathbf{z}_i,\mathbf{w}_i,t)\ dt \tag{2}$$

Assuming sufficient continuity and differentiability of $\mathbf{w}$ *w.r.t.* ϵ, the MacLaurin series expansion for $\mathbf{w}_i$ has the form

$$\mathbf{w}_1 \simeq \sum_{j=1}^{K} \frac{\epsilon^j}{j!}\ \mathbf{w}_i^{(j)}\ ,\ i = 1,2,\ldots,N \tag{3}$$

where $\mathbf{w}_i^{(j)} \triangleq \lim_{\epsilon\to 0} \partial^j\mathbf{w}_i/\partial\epsilon^j$. The computation of coefficients of series, as in the nondelayed case [8.11] involves uncoupled and, for this case unretarded subsystems solutions.

To illustrate the near-optimum control for this class of large-scale systems it is assumed, without any loss of generality, that there are two subsystems involved in (1), *i.e.*, $N = 2$. Letting $\mathbf{z}_1 = \mathbf{x}$, $\mathbf{z}_2 = \mathbf{y}$, $\mathbf{w}_1 = \mathbf{u}$, $\mathbf{w}_2 = \mathbf{v}$, $\mathbf{c}_1(.) = \mathbf{a}(.)$, $\mathbf{c}_2(.) = \mathbf{b}$, $\mathbf{h}_1 = \mathbf{f}$, and $\mathbf{h}_2 = \mathbf{g}$, the state models for the subsystem x and subsystem y become

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},\alpha) = \mathbf{f}(\mathbf{x},\mathbf{u},\epsilon a(\mathbf{x},\bar{\mathbf{x}},\mathbf{y},\bar{\mathbf{y}},\mathbf{u},\bar{\mathbf{u}},\mathbf{v},\bar{\mathbf{v}})) \tag{4a}$$

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y},\mathbf{v},\beta) = \mathbf{g}(\mathbf{y},\mathbf{v},\epsilon b(\mathbf{x},\bar{\mathbf{x}},\mathbf{y},\bar{\mathbf{y}},\mathbf{u},\bar{\mathbf{u}},\mathbf{v},\bar{\mathbf{v}})) \tag{4b}$$

where dimensions of $\mathbf{x}$ and $\mathbf{y}$ are, respectively, n_1 and n_2 and the bar on the variables defines delayed quantities. The initial functions for (4) are

$$x(\tau) = \phi(\tau), \ t_0 - h \leqslant \tau \leqslant t_0 \tag{5a}$$

$$y(\tau) = \psi(\tau), \ t_0 - h \leqslant \tau \leqslant t_0 \tag{5b}$$

Assuming that $L_1(.) = F(.)$ and $L_2(.) = G(.)$ in (2), the necessary conditions for the optimality of the coupled TD systems (4)-(5) reduces to the following two TD and advance sets of TPBV problems [8.1].

$$\dot{x} = H_p' = f \tag{6a}$$

$$\dot{p} = -H_x' - H_x'(s) = F_x' - (f_x + \epsilon f_x a_x)'p - \epsilon(g_\beta b_x)'q$$

$$-\epsilon\left\{(f_\alpha a_x)'p + (g_\beta b_x)'q\right\}\Big|_s, \quad t \in I_1 \tag{6b}$$

$$= -H'_x = F'_x - (f_x + \epsilon f_\alpha a_x)'p - \epsilon(g_\beta b_x)'q\,, \quad t \in I_2 \tag{6c}$$

$$0 = H'_u + H_u'(s) = -F'_u + (f_u + \epsilon f_\alpha a_u)'p + \epsilon(g_\beta b_u)'q$$

$$+ \epsilon\left\{(f_\alpha a_u)'p + (g_\beta b_u)'\, q\right\}\Big|_s, \quad t \in I_1 \tag{6d}$$

$$= H'_u = -F'_u + (f_u + \epsilon f_\alpha a_u)'p + \epsilon(g_\beta b_u)'\, q, \quad t \in I_2 \tag{6e}$$

$$\dot{y} = H'_q = g \tag{7a}$$

$$\dot{q} = -H'_y - H'_y(s) = G'_y - (g_y + \epsilon g_\beta b_y)'q - \epsilon(f_\alpha a_y)'p$$

$$-\epsilon\left\{(g_\beta b_y)'q + (f_\alpha a_y)'p\right.\Big|_s, \quad t \in I_1 \tag{7b}$$

$$= -H'_y = G'_y - (g_y + \epsilon g_\beta b_y)'q - \epsilon(f_\alpha a_y)'p\,, \quad t \in I_2 \tag{7c}$$

$$0 = H_v' + H_v'(s) = -G_v' + (g_v + \epsilon g_\beta b_v)'q + \epsilon(f_\alpha a_v)'p$$

$$+ \epsilon\left\{(g_\beta b_{\bar{v}})'q + (f_\alpha a_{\bar{v}})'p\right\}\Big|_s, \quad t \in I_1 \tag{7d}$$

$$= H'_v = -G'_v + (g_v + \epsilon g_\beta b_v)'\, q + \epsilon(f_\alpha a_v)'\, p, \ t \in I_2 \tag{7e}$$

where $I_1 : t_0 \leqslant t \leqslant t_f - h$, $I_2 : t_f - h \leqslant t \leqslant t_f$, $s = t + h$, $H = -F - G + p'f + q'g$ is the Hamiltonian function, and p and q are, respectively, costates of x and y. The final values of costates $p(t_f)$ and $q(t_f)$ are zero vectors.

Jamshidi [8.10] has proved four assertions regarding the reduction of the linear and nonlinear time-delay, time-advanced TPBV problems as well as their respective sensitivity functions TPBV problems to two independent uncoupled unretarded TPBV problems as $\epsilon \to 0$.

1. Assertion. At $\epsilon = 0$ the necessary conditions (6) and (7) reduce to two independent TPBV problem for the two uncoupled unretarded systems.

Proof. At $\epsilon = 0$, none of the functionals $\mathbf{f}$, $\mathbf{f}_x$, and $\mathbf{f}_v$ depend on $\bar{\mathbf{x}}$, $\mathbf{y}$, $\bar{\mathbf{y}}$, $\bar{\mathbf{u}}$, $\mathbf{v}$, and $\bar{\mathbf{v}}$. Similarly, none of the functions $\mathbf{g}$, $\mathbf{g}_y$, and $\mathbf{g}_v$ depend on $\bar{\mathbf{y}}$, $\mathbf{x}$, $\bar{\mathbf{x}}$, $\bar{\mathbf{v}}$, $\mathbf{v}$ and $\mathbf{u}$. Thus,

$$\dot{\mathbf{x}} = \mathbf{f}, \quad \mathbf{x}(t_0) = \phi(t_0) \tag{8a}$$

$$\dot{\mathbf{p}} = \mathbf{f}_x' - \mathbf{f}_x'\mathbf{p}, \quad \mathbf{p}(t_f) = 0 \tag{8b}$$

$$0 = -\mathbf{F}_u' + \mathbf{f}_u'\mathbf{p} \tag{8c}$$

$$\dot{\mathbf{y}} = \mathbf{g}, \quad \mathbf{y}(t_0) = \psi(t_0) \tag{9a}$$

$$\dot{\mathbf{q}} = \mathbf{G}_y' - \mathbf{g}_y'\mathbf{q}, \quad \mathbf{q}(t_f) = 0 \tag{9b}$$

$$0 = -\mathbf{G}_v' + \mathbf{g}_v'\mathbf{q} \tag{9c}$$

The solutions of TPBV problems (8) and (9) provide $\mathbf{u}^{(0)}$ and $\mathbf{v}^{(0)}$, which are the zeroth-order terms of series in (3).

2. Assertion. If the solution from the necessary conditions for terms $j = 1,2,\ldots,i-1$ is known, the necessary conditions for the ith term reduces to two independent unretarded TPBV problems. Furthermore, the homogeneous parts of these problems remains unchanged for $i \geqslant 1$.

Proof. Differentiate (6) i times with respect to ϵ and let $\epsilon \to 0$; then the following TPBV problem is obtained for subsystem x:

$$\dot{\mathbf{x}}^{(i)} = \mathbf{f}_x\mathbf{x}^{(i)} + \mathbf{f}_u\mathbf{u}^{(i)} + \mathbf{X}^i \tag{10a}$$

$$\dot{\mathbf{p}}^{(i)} = (\mathbf{F}_{xx} - \mathbf{f}_x^1)\mathbf{x}^{(i)} - \mathbf{f}_x'\mathbf{p}^{(i)} + (\mathbf{F}_{xu} - \mathbf{f}_u^1)\,\mathbf{u}^{(i)} + \mathbf{P}_1^i,\ t \in I_1 \tag{10b}$$

$$= (\mathbf{F}_{xx} - \mathbf{f}_x^1)\mathbf{x}^{(i)} - \mathbf{f}_x'\mathbf{p}^{(i)} + (\mathbf{F}_{xu} - \mathbf{f}_u^1)\mathbf{u}^{(i)} + \mathbf{P}_2^i,\ t \in I_2 \tag{10c}$$

$$0 = (\mathbf{f}_x^2 - \mathbf{F}_{ux})\mathbf{x}^{(i)} + \mathbf{f}_u'\mathbf{p}^{(i)} + (\mathbf{f}_u^2 - \mathbf{F}_{uu})\,\mathbf{u}^{(i)} + \mathbf{U}_1^i,\ t \in I_1 \tag{10d}$$

$$= (\mathbf{f}_x^2 - \mathbf{F}_{ux})\mathbf{x}^{(i)} + \mathbf{f}_u'\mathbf{p}^{(i)} + (\mathbf{f}_u^2 - \mathbf{F}_{uu})\,\mathbf{u}^{(i)} + \mathbf{U}_2^i,\ t \in I_2 \tag{10e}$$

where $\mathbf{x}^{(i)}(0) = 0$ and $\mathbf{p}^{(i)}(t_f) = 0$, $\mathbf{f}^1 = \mathbf{f}^1(\mathbf{x},\mathbf{p},\mathbf{v})$, and $\mathbf{f}^2 = \mathbf{f}^2(\mathbf{x},\mathbf{p},\mathbf{v})$, respectively, denote vectors $f_x{}'p$ and $f_u{}'p$. It must also be noted that the nonhomogeneous parts X^i, P^i, and U^i, do not depend on $x^{(i)}$, $p^{(i)}$, $u^{(i)}$, $y^{(i)}$, $q^{(i)}$, or $v^{(i)}$. A similar procedure will lead to a TPBV problem for subsystem y:

$$\dot{y}^{(i)} = g_y y^{(i)} + g_v v^{(i)} + Y^i \tag{11a}$$
$$\dot{q}^{(i)} = (G_{yy} - g_y^1)y^{(i)} - g_y' q^{(i)} + (G_{yv} - g_v^1)v^{(i)} + Q_1^i \,,\; t \in I_1 \tag{11b}$$
$$= (G_{yy} - g_y^1)y^{(i)} - g_y' q^{(i)} + (G_{yv} - g_v^1)\, v^{(i)} + Q_2^i \,,\; t \in I_2 \tag{11c}$$
$$0 = (g_y^2 - G_{vy})y^{(i)} + g_v' q^{(i)} + (g_v^2 - G_{vv})\, v^{(i)} + V_1^i \,,\; t \in I_1 \tag{11d}$$
$$= (g_y^2 - G_{vy})y^{(i)} + g_v' q^{(i)} + (g_v^2 - G_{vv})v^{(i)} + V_2^i \,,\; t \in I_2 \tag{11e}$$

Similar comments apply to Y^i, Q^i, and V^i.

It is noted that the only unknown in TPBV problems (10) and (11) are $x^{(i)}$, $p^{(i)}$, $u^{(i)}$ and $y^{(i)}$, $q^{(i)}$, $v^{(i)}$, respectively. Thus, the solution of the necessary conditions for the ith terms of the control truncated series (3) reduces to two independent TPBV problems. Furthermore, due to the characteristic of sensitivity equations, the homogeneous parts of (10) and (11) remain the same for all $i=1,2,\ldots,k$. *q.e.d.*

For the case when (4) is linear, one has

$$\dot{x} = Ax + Bu + \epsilon a(.)$$
$$= Ax + Bu + \epsilon\,(A_1 y + A_2\bar{y} + A_3\bar{x} + B_1 v + B_2\bar{u} + B_3\bar{v}) \tag{12a}$$
$$\dot{y} = Dy + Ev + \epsilon b(.)$$
$$= Dy + Ev + \epsilon(D_1 x + D_2\bar{x} + D_3\bar{y} + E_1 u + E_2\bar{u} + E_3\bar{v}) \tag{12b}$$

with the initial functions given by (5). The cost function is the usual quadratic form

$$J = \tfrac{1}{2}\int_{t_0}^{t_f} \left[x'Q_1x + y'Q_2y + u'R_1u + v'R_2v\right]dt \tag{13}$$

where Q_i, R_i, $i = 1,2$, have the usual regulating conditions. The necessary conditions for optimality are similar to those discussed for nonlinear system:

$$\dot{\mathbf{x}} = \mathbf{H}_P' = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \epsilon\mathbf{a}(\cdot), \quad \mathbf{x}(\tau) = \phi(\tau) \tag{14a}$$

$$\dot{\mathbf{p}} = \mathbf{Q}_1\mathbf{x} - \mathbf{A}'\mathbf{P} - \epsilon\mathbf{D}'_1\mathbf{q} - \epsilon\mathbf{A}'_3\mathbf{p}(t+h) - \epsilon\mathbf{D}'_2\mathbf{q}(t+h)\ , t \in I_1 \tag{14b}$$

$$= \mathbf{Q}_1\mathbf{x} - \mathbf{A}'\mathbf{P} - \epsilon\mathbf{D}'_1\mathbf{q},\ \mathbf{p}(t_f) = 0\ , t \in I_2$$

$$0 = -\mathbf{R}_1\mathbf{u} + \mathbf{B}'\mathbf{P} + \epsilon\mathbf{E}'_1\mathbf{q} + \epsilon\mathbf{B}'_2\mathbf{p}(t+h) + \epsilon\mathbf{E}'_2\mathbf{q}(t+h)\ , t \in I_1 \tag{14d}$$

$$= -\mathbf{R}_1\mathbf{u} + \mathbf{B}'\mathbf{P} + \epsilon\mathbf{E}'_1\mathbf{q}\ , t \in I_2 \tag{14e}$$

Similarly, for subsystem y:

$$\dot{\mathbf{y}} = \mathbf{H}'_q = \mathbf{D}\mathbf{y} + \mathbf{E}\nu + \epsilon\mathbf{b}(.),\ \mathbf{y}(\tau) = \psi(\tau) \tag{15a}$$

$$\dot{\mathbf{q}} = \mathbf{Q}_2\mathbf{y} - \mathbf{D}'\mathbf{q} - \epsilon\mathbf{A}'_1\mathbf{p} - \epsilon\mathbf{D}'_3\mathbf{q}\ (t+h) - \epsilon\mathbf{A}'_2\mathbf{p}(t+h)\ , t \in I_1 \tag{15b}$$

$$= \mathbf{Q}_2\mathbf{y} - \mathbf{D}'\mathbf{q} - \epsilon\mathbf{A}'_1\mathbf{p},\ \mathbf{q}(t_f) = 0\ , t \in I_2 \tag{15c}$$

$$0 = -\mathbf{R}_2\mathbf{v} + \mathbf{E}'\mathbf{q} + \epsilon\mathbf{B}'_1\mathbf{p} + \epsilon\mathbf{E}'_2\mathbf{q}(t+h) + \epsilon\mathbf{B}'_2\mathbf{p}(t+h)\ , t \in I_1 \tag{15d}$$

$$= -\mathbf{R}_2\mathbf{v} + \mathbf{E}'\mathbf{q} + \epsilon\mathbf{B}'_1\mathbf{p}\ , t \in I_2 \tag{15e}$$

In a similar fashion, the following two assertions are given for the linear system case.

3. Assertion. At $\epsilon = 0$, the necessary conditions (14) and (15) reduce to two independent linear TPBV problems for the two subsystems.

Proof. At $\epsilon = 0$, it follows from (14) to (15) that for subsystem x

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{16a}$$

$$\dot{\mathbf{p}} = \mathbf{Q}_1\mathbf{x} - \mathbf{A}'\mathbf{p} \tag{16b}$$

$$0 = -\mathbf{R}_1\mathbf{u} + \mathbf{B}'\mathbf{p} \tag{16c}$$

and for subsystem y

$$\dot{\mathbf{y}} = \mathbf{D}\mathbf{v} + \mathbf{E}\mathbf{v} \tag{17a}$$

$$\dot{\mathbf{q}} = \mathbf{Q}_2\mathbf{y} + \mathbf{D}'\mathbf{q} \tag{17b}$$

$$0 = -\mathbf{R}_2\mathbf{v} + \mathbf{E}'\mathbf{q}. \tag{17c}$$

It is clear that the above problems are independent of each other, and furthermore their solutions constitute the zeroth-order terms $\mathbf{u}^{(0)}$ and $\mathbf{v}^{(0)}$ given as

$$\mathbf{u}^{(0)} = -\mathbf{R}_1^{-1}\mathbf{B}'\mathbf{K}\mathbf{x}^{(0)} \tag{18a}$$

$$\mathbf{v}^{(0)} = -\mathbf{R}_2^{-1}\mathbf{E}'\mathbf{P}\mathbf{y}^{(0)} \tag{18b}$$

where **K** and **P** are the solutions of the following two matrix Riccati equations:

$$\dot{\mathbf{K}} = -\mathbf{A}'\mathbf{K} - \mathbf{K}\mathbf{A} + \mathbf{K}\mathbf{S}\mathbf{K} - \mathbf{Q}_1, \ \ \mathbf{K}(t_f) = 0 \tag{19}$$

$$\dot{\mathbf{P}} = -\mathbf{D}'\mathbf{P} - \mathbf{P}\mathbf{D} + \mathbf{P}\mathbf{T}\mathbf{P} - \mathbf{Q}_2, \ \ \mathbf{P}(t_f) = 0 \tag{20}$$

where $\mathbf{S} = \mathbf{B}\mathbf{R}_1^{-1}\mathbf{B}'$ and $\mathbf{T} = \mathbf{E}\mathbf{R}_2^{-1}\mathbf{E}'$.

4. Assertion. If the solutions from the necessary conditions for the jth terms, $j = 1,2,\ldots,i-1$, are known, the necessary conditions for the ith term reduces to two independent linear unretarded TPBV problems. Furthermore the homogeneous parts of all solutions remain the same throughout.

Proof. Differentiate (14) i times with respect to ϵ and let $\epsilon \to 0$; then the following TPBV problem is obtained for subsystem x:

$$\dot{\mathbf{x}}^{(i)} = \mathbf{A}\mathbf{x}^{(i)} + \mathbf{B}\mathbf{u}^{(i)} + \mathbf{X}^i \tag{21a}$$

$$\dot{\mathbf{p}}^{(i)} = \mathbf{Q}_1\mathbf{x}^{(i)} + \mathbf{A}'\mathbf{u}^{(i)} + \mathbf{P}_1^i \ , t \in I_1 \tag{21b}$$

$$= \mathbf{Q}_1\mathbf{x}^{(i)} + \mathbf{A}'\mathbf{u}^{(i)} + \mathbf{P}_2^i \ , t \in I_2 \tag{21c}$$

$$0 = -\mathbf{R}_1\mathbf{u}^{(i)} + \mathbf{B}'\mathbf{p}^{(i)} + \mathbf{U}_1^i \ , t \in I_1 \tag{21d}$$

$$= -\mathbf{R}_1\mathbf{u}^{(i)} + \mathbf{B}'\mathbf{p}^{(i)} + \mathbf{U}_2^i \ , t \in I_2 \tag{21e}$$

where $\mathbf{x}^{(i)}(\mathbf{t}_0) = \mathbf{p}^{(i)}(\mathbf{t}_f) = 0$ and nonhomogeneous parts $\mathbf{X}^i$, $\mathbf{P}^i$, and $\mathbf{U}^i$ do not depend on the ith terms. A similar procedure will result in a TPBV problem for subsystem y:

$$\dot{\mathbf{y}}^{(i)} = \mathbf{D}\mathbf{y}^{(i)} + \mathbf{E}\mathbf{v}^{(i)} + \mathbf{Y}^i \tag{22a}$$

$$\dot{\mathbf{q}}^{(i)} = \mathbf{Q}_2\mathbf{y}^{(i)} + \mathbf{D}'\mathbf{q}^{(i)} + \mathbf{Q}_1^i \ , t \in \mathrm{I}_1 \tag{22b}$$

$$= \mathbf{Q}_2\mathbf{y}^{(i)} + \mathbf{D}'\mathbf{q}^{(i)} + \mathbf{Q}_2^i \ , t \in \mathrm{I}_2 \tag{22c}$$

$$0 = -\mathbf{R}_2\mathbf{v}^{(i)} + \mathbf{E}'\mathbf{q}^{(i)} + \mathbf{V}_1^i \ , t \in \mathrm{I}_1 \tag{22d}$$

$$= -\mathbf{R}_2\mathbf{v}^{(i)} + \mathbf{E}'\mathbf{q}^{(i)} + \mathbf{V}_2^i \ , t \in \mathrm{I}_2 \tag{22e}$$

This completes the proof. Δ

In sequel, two algorithms are given for near-optimum control of linear and nonlinear coupled TD systems.

5. Algorithm. - Nonlinear Case

Step 1: Calculate the zeroth-order (unretarded) terms from (8) and (9).

Step 2: a. Solve two matrix Riccati equations:

$$\dot{\mathbf{K}}_x = -\mathbf{K}_x\mathbf{A}_x - \mathbf{A}_x'\mathbf{K}_x + \mathbf{K}_x\mathbf{S}_x\mathbf{K}_x - \mathbf{Q}_x, \;\; \mathbf{K}_x(t_f) = 0 \tag{23}$$

$$\dot{\mathbf{K}}_y = \mathbf{K}_y\mathbf{A}_y - \mathbf{A}'_y\mathbf{K}_y + \mathbf{K}_y\mathbf{S}_y\mathbf{K}_y - \mathbf{Q}_y, \;\; \mathbf{K}_y(t_f) = 0 \tag{24}$$

b. Solve the linear adjoint equations

$$\dot{\mathbf{g}}_x^1 = -(\mathbf{A}_x - \mathbf{S}_x\mathbf{K}_x)'\mathbf{g}_x^1 - \mathbf{K}_x\mathbf{L}_x^1 + \mathbf{l}_x^1, \; t \in \mathbf{I}_1 \tag{25a}$$

$$= -(\mathbf{A}_x - \mathbf{S}_x\mathbf{K}_x)'\mathbf{g}_x^1 - \mathbf{K}_x\mathbf{L}_x^2 + \mathbf{l}_x^2, \; t \in \mathbf{I}_2$$

$$\dot{\mathbf{g}}_y^1 = -(\mathbf{A}_y - \mathbf{S}_y\mathbf{K}_y)'\mathbf{g}_y^1 - \mathbf{K}_y\mathbf{L}_y^1 + \mathbf{l}_y^1, \; t \in \mathbf{I}_1 \tag{25b}$$

$$= -(\mathbf{A}_y - \mathbf{S}_y\mathbf{K}_y)'\mathbf{g}_y^1 - \mathbf{K}_y\mathbf{L}_y^2 + \mathbf{l}_y^2, \; t \in \mathbf{I}_2 \tag{26b}$$

$$\mathbf{g}_x^1(t_f) = 0, \;\; \mathbf{g}_y^1(t_f) = 0$$

c. Obtain $\mathbf{x}^{(1)}$ and $\mathbf{y}^{(1)}$ from

$$\dot{\mathbf{x}}^{(1)} = (\mathbf{A}_x - \mathbf{S}_x\mathbf{K}_x)\mathbf{x}^{(1)} - \mathbf{S}_x\mathbf{g}_x^1 + \mathbf{L}_x^1 \tag{27}$$

$$\dot{\mathbf{y}}^{(1)} = (\mathbf{A}_y - \mathbf{S}_y\mathbf{K}_y)\mathbf{y}^{(1)} - \mathbf{S}_y\mathbf{g}_y^1 + \mathbf{L}_y^1 \tag{28}$$

with $\mathbf{x}^{(1)}(t_0) = 0$, $\mathbf{y}^{(1)}(t_0) = 0$ and $\mathbf{u}^{(1)}$ obtained from

$$\mathbf{u}^{(1)} = -(\Gamma_x + \Pi_x\mathbf{K}_x)\mathbf{x}^{(1)} - \beta_x^1 - \Pi_x\mathbf{g}_x^1, \; t \in \mathbf{I}_1 \tag{29a}$$

$$= -(\Gamma_x + \Pi_x\mathbf{K}_x)\mathbf{x}^{(1)} - \beta_x^2, \; t \in I_2 \tag{29b}$$

$$\mathbf{v}^{(1)} = -(\Gamma_x + \Pi_y\mathbf{K}_y)\mathbf{y}^{(1)} - \beta_y^2 - \Pi_y g_y^1, \; t \in I_1 \tag{30a}$$

$$= -(\Gamma_y + \Pi_y\mathbf{K}_y)\mathbf{y}^{(1)} - \beta_y^2, \; t \in I_2 \tag{30b}$$

Expressions for $\mathbf{A}_x$, $\mathbf{S}_x$, $\mathbf{Q}_x$, Γ_x, Π_x, $\mathbf{L}_x^1$, $\mathbf{L}_x^2$, $\mathbf{l}_x^1$, $\mathbf{l}_x^2$, β_x^1, β_x^2, etc. are [8.10]:

$$\mathbf{A}_x = \mathbf{f}_x - \mathbf{f}_u(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}(\mathbf{f}_x^2-\mathbf{F}_{ux}) \tag{31a}$$
$$\mathbf{S}_x = \mathbf{f}_u(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{f}'_u \tag{31b}$$
$$\mathbf{Q}_x = (\mathbf{F}_{xx}-\mathbf{f}_x^1) - (\mathbf{F}_{xu}-\mathbf{f}_u^1)(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}(\mathbf{f}_x^2-\mathbf{F}_{ux}) \tag{31c}$$
$$\Gamma_x = (\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}(\mathbf{f}_x^2-\mathbf{F}_{ux}) \tag{31d}$$
$$\Pi_x = (\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{f}'_u \tag{31e}$$
$$\mathbf{L}_x^1 = \mathbf{f}_\alpha a - \mathbf{f}_u(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_1^1 \tag{31f}$$
$$\mathbf{L}_x^2 = \mathbf{f}_\alpha a - \mathbf{f}_u(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_2^1 \tag{31g}$$
$$\mathbf{l}_x^1 = \mathbf{Q}_c - (\mathbf{F}_{xu}-\mathbf{f}_u^1)(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_1^1 \tag{31h}$$
$$\mathbf{l}_x^2 = \mathbf{Q}_c - (\mathbf{F}_{xu}-\mathbf{f}_u^1)(\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_2^1 \tag{31i}$$
$$\beta_x^1 = (\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_1^1 \tag{31j}$$
$$\beta_x^2 = (\mathbf{f}_u^2-\mathbf{F}_{uu})^{-1}\mathbf{U}_2^1 \tag{31k}$$

where $\mathbf{Q}_c = -(\mathbf{f}_\alpha^1 a + a'_x\mathbf{f}'_\alpha\mathbf{p} + \mathbf{b}'_x\mathbf{g}'_\beta\mathbf{q})$.

Step j ($j>2$): Solve the linear adjoint equations similar to (25) and (26) for $\mathbf{g}_x^j$ and $\mathbf{g}_y^j$ using the same Riccati solutions as in (23) and (24):

$$\dot{\mathbf{x}}^{(j)} = (\mathbf{A}_x - \mathbf{S}_x\mathbf{K}_x)\mathbf{x}^{(j)} - \mathbf{S}_x\mathbf{g}_x^j + \mathbf{L}_x^j \tag{32}$$
$$\dot{\mathbf{y}}^{(j)} = (\mathbf{A}_y - \mathbf{S}_y\mathbf{K}_y)\mathbf{y}^{(j)} - \mathbf{S}_y\mathbf{g}_y^j + \mathbf{L}_y^j \tag{33}$$

with a similar expression for $\mathbf{v}^{(j)}$ and $\mathbf{u}^{(j)}$ as in (29) and (30).

Step $j+1$: The control vectors are given by expressions (40)-(44) below.

Stop.

6. Algorithm. - Linear Case

Step 1: Using (16) and (17) and Riccati equations (19) and (20), obtain the zeroth-order terms.

Step 2: a. Obtain the adjoint equations

$$\dot{\mathbf{g}}_x^1 = -(\mathbf{A}-\mathbf{SK})'\mathbf{g}_x^1 - \mathbf{KL}_x^1 + \mathbf{l}_x^1 \,,\ t \in \mathbf{I}_1 \tag{34a}$$
$$= -(\mathbf{A}-\mathbf{SK})'\,\mathbf{g}_x^1 - \mathbf{KL}_x^2 + \mathbf{l}_x^2 \,,\ t \in \mathbf{I}_2 \tag{34b}$$
$$\dot{\mathbf{g}}_y^1 = -\,(\mathbf{D}-\mathbf{TP})'\,\mathbf{g}_y^1 - \mathbf{PL}_y^1 + \mathbf{l}_y^1 \,,\ t \in \mathbf{I}_1 \tag{35a}$$
$$= -\,(\mathbf{D}-\mathbf{TP})'\,\mathbf{g}_y^1 - \mathbf{PL}y^2 + \mathbf{l}_y^2 \,,\ t \in \mathbf{I}_2 \tag{35b}$$

where $\mathbf{L}_x^i = \mathbf{X}^i - \mathbf{BR}_1^{-1}\mathbf{U}_i^1$, $\mathbf{l}_x^i = \mathbf{Q}^i$, $\mathbf{L}_y^i = \mathbf{Y}^i - \mathbf{ER}_2^{-1}\mathbf{V}_i^1$, $i=1,2$.

b. Then $\mathbf{x}^{(1)}$, $\mathbf{y}^{(1)}$, $\mathbf{u}^{(1)}$, and $\mathbf{v}^{(1)}$ are obtained from

$$\dot{\mathbf{x}}^{(1)} = (\mathbf{A}-\mathbf{SK})'\,\mathbf{x}^{(1)} + \mathbf{S}\mathbf{g}_x^1 + \mathbf{L}_x^1 \tag{36}$$

$$\dot{\mathbf{y}}^{(1)} = (\mathbf{D}-\mathbf{TP})'\,\mathbf{y}^{(1)} + \mathbf{T}\mathbf{g}_y^1 + \mathbf{L}_y^1 \tag{37}$$

$$\mathbf{u}^{(1)} = -\mathbf{R}_1^{-1}\mathbf{B}'\mathbf{K}\mathbf{x}^{(1)} + \mathbf{R}_l^{-1}\mathbf{U}_1^1 + \mathbf{R}_1^{-1}\mathbf{B}'\mathbf{g}_x^1 \,,\; t \in \mathbf{I}_1 \tag{38a}$$

$$= -\mathbf{R}_1^{-1}\mathbf{B}'\mathbf{K}\mathbf{x}^{(1)} + \mathbf{R}_1^{-1}\mathbf{U}_1^2 \,,\; t \in \mathbf{I}_2 \tag{38b}$$

$$\mathbf{v}^{(1)} = -\mathbf{R}_2^{-1}\mathbf{E}'\mathbf{P}_y^{(1)} + \mathbf{R}_2^{-1}\mathbf{V}_1^1 + \mathbf{R}_2^{-1}\mathbf{E}^T\mathbf{g}_y^1 \,,\; t \in \mathbf{I}_1 \tag{39a}$$

$$= -\mathbf{R}_2^{-1}\mathbf{E}'\mathbf{P}_y^{(1)} + \mathbf{R}_2^{-1}\mathbf{V}_1^2 \,,\; t \in \mathbf{I}_2 \tag{39b}$$

Step $j\,(j > 2)$: This step is similar to the corresponding step of the nonlinear case.

Note that based on the expression for $\mathbf{u}^{(i)}$ and $\mathbf{v}^{(i)}$ in both linear and nonlinear system (29), (30), (38), (39), the control structure has the following general forms:

$$\mathbf{u} = -\alpha_x\mathbf{x} + \sum_{i=1}^{N} \frac{\epsilon^i}{i!}\,\beta_{xi} \,,\; t \in \mathbf{I}_1 \tag{40a}$$

$$= -\alpha_x\mathbf{x} + \sum_{i=1}^{N} \frac{\epsilon^i}{i!}\,\gamma_{xi} \,,\; t \in I_2 \tag{40b}$$

$$\mathbf{v} = -\alpha_y\mathbf{y} + \sum_{i=1}^{N} \frac{\epsilon^i}{i!}\,\beta_{yi} \,,\; t \in I_1 \tag{41a}$$

$$= -\alpha_y\mathbf{y} + \sum_{i=1}^{N} \frac{\epsilon^i}{i!}\,\gamma_{yi} \,,\; t \in I_2 \tag{41b}$$

where

$$\alpha_x = (\Gamma_x + \Pi_x\mathbf{K}_x),\ \text{nonlinear} \tag{42a}$$

$$= \mathbf{R}_1^{-1}\mathbf{B}'\mathbf{K} \,,\ \text{linear} \tag{42b}$$

$$\beta_{xi} = -\beta_{xi}^1 + \Pi_x\mathbf{g}_x^i \,,\ \text{nonlinear} \tag{43a}$$

$$= -\mathbf{R}_1^{-1}\mathbf{U}_1^i + \mathbf{B}_1^{-1}\mathbf{B}'\mathbf{g}_x^i \,,\ \text{linear} \tag{43b}$$

$$\gamma_{xi} = -\beta_{xi}^2 \,,\ \text{nonlinear} \tag{44a}$$

$$= -\mathbf{R}_1^{-1}\mathbf{U}_2^i \,,\ \text{linear} \tag{44b}$$

Similar terms can be obtained for subsystem y.

The following two examples illustrate the two algorithms.

7. Example. Consider a simple nonlinear coupled time-delay system

$$\dot{x}_1 = x_2 + w_1,\ \dot{x}_2 = -x_1 - 0.8x_2 - x_2(t-0.1) - x_1^3 + w \tag{45}$$

with

$$x_1(\tau) = 1 + \tau,\ x_2(\tau) = -1 - \tau\ ,\ -0.1 \leqslant 0.0 \tag{46}$$

In order to decouple the state model (45), let control $w = u + \epsilon v = \epsilon u + v$ and $x = x_1$ and $y = y_1$; then

$$\dot{x} = f(x,u,\epsilon a(.)) = u + \epsilon(y+v) \tag{47a}$$

$$\dot{y} = g(y,v,\epsilon b(.)) = -0.8y + v + \epsilon\left[u-x-x^3-y(t-0.1)\right] \tag{47b}$$

Note that (47) is now in the general couple time-delay systems of (4). The cost function is chosen to be quadratic, with $L_1 = F = ½(x^2+u^2)$, $L_2 = G = ½(y_2+v^2)$, and $t_f = 2.0$ sec.

The result of the zeroth-order terms can be obtained from a pair of linear TPBV problems

$$\dot{x}^{(0)} = -p^{(0)},\ \dot{y}^{(0)} = -0.8y^{(0)} + q^{(0)} \tag{48a}$$

$$\dot{p}^{(0)} = x^{(0)},\ \dot{q}^{(0)} = y^{(0)} - 0.8q^{(0)} \tag{48b}$$

based on the solutions of the associated pair of Riccati equations similar to (23) and (24):

$$x^{(0)}(t) = -exp(-t),\ \ p^{(0)}(t) = exp(-t),\ \ u^{(0)}(t) = exp(-t) \tag{49a}$$

$$y^{(0)}(t) = exp(-1.28t),\ \ q^{(0)}(t) = -0.48\ exp(-1.28t),$$

$$v^{(0)}(t) = -0.48\ exp(-1.28t) \tag{49b}$$

The first order adjoint vectors are

$$g_x^1(t) = -0.2533\ exp(-t) + 0.44\ exp(-1.28t) + 0.336\ exp(-3.28t),\ 0 \leqslant t \leqslant 1.9 \tag{50a}$$

$$= -1.2533\ exp(-t) + 0.44\ exp(-1.28t) + 0.336\ exp(-3.28t),\ 1.9 \leqslant t \leqslant 2.0 \tag{50b}$$

$$g_y^1(t) = -0.532\ exp(-1.28t) + 0.88\ exp(-t) - 0.234\ exp(-3t),\ 0 \leqslant t \leqslant 1.9 \tag{51a}$$

$$= -1.532\ exp(-1.28t) + 0.88\ exp(-t) - 0.234\ exp(-3t),\ 1.9 \leqslant t \leqslant 2.0 \tag{51b}$$

The suboptimal controls u and v after truncating the forcing functions after the first term are

$$u \simeq x + \epsilon g_x^1(t) \tag{52a}$$

$$v \simeq -y + \epsilon g_y^1(t) \tag{52b}$$

The state equations are then

$$\dot{x} \simeq -x - \epsilon g_x^1 - \epsilon g_y^1 \tag{53a}$$

$$\dot{y} \simeq -1.8y - \epsilon\left[2x + x^3 + y(t-0.1) - \epsilon g_x^1\right] + \epsilon g_y^1 \tag{53b}$$

with initial functions (46). Figure 1 shows the suboptimal state and control of the nonlinear plant. The near-optimum performance index turned out to be $J^n = 1.0$.

8. Example. Consider the following linear coupled TD system:

$$\dot{x}_1 = -x_1 + x_2 + 0.2x_1(t-0.1) + w \tag{54a}$$

$$\dot{x}_2 = -x_1 - x_2 - 0.1x_2(t-0.1) + w \tag{54b}$$

with initial functions $x_1(\tau) = x_2(\tau) = 1.0, -0.1, \leqslant \tau \leqslant 0$.

In a similar fashion as in the nonlinear plant example, one may let $x = x_1$, $y = x_2$, and $w = u + \epsilon v = \epsilon u + v$ in the two equations of (54). The following coupled linear systems result:

$$\dot{x} = -x + u + \epsilon a(.) = -x + u + \epsilon\left[2x(t-0.1) + y + v\right] \tag{55a}$$

$$\dot{y} = y + v + \epsilon b(.) = -y + v + \epsilon\left[-x - 0.1y(t-0.1) + u\right] \tag{55b}$$

The cost function is

$$J = \tfrac{1}{2} \int_0^1 (x^2 + y^2 + u^2/3 + v^2/3)\, dt \tag{56}$$

The Riccati equations (19) and (20) become

$$\dot{k} = -2k + 3k^2 - 1, \; k(1) = 0 \tag{57a}$$

$$\dot{p} = -2p + 3p^2 - 1, \; p(1) = 0 \tag{57b}$$

Using the solutions of these equations, the responses of the decoupled systems, *i.e.*, zeroth-order terms, are easily obtained:

$$x^{(0)}(t) = y^{(0)}(t) = exp(-4t), \;\; p^{(0)}(t) = q^{(0)}(t) = -exp(-4t), \tag{58a}$$

$$u^{(0)}(t) = v^{(0)}(t) = -3\; exp(-4t) \tag{58b}$$

The first order adjoint vectors follow form (34) and (35) as

$$g_x^1(t) = -0.47 \times 10^{-4}\, exp(4t) + 0.17\, exp\,(-4t), \; 0 \leqslant t \leqslant 0.9 \tag{59a}$$

$$= -0.61 \times 10^{-4}\, exp(4t) + 0.19\, \exp\,(-4t), \; 0.9 \leqslant t \leqslant 1.0 \tag{59b}$$

$$g_y^1(t) = -0.18 \times 10^{-3}\, exp(4t) + 0.565\, exp(-4t), \; 0 \leqslant t \leqslant 0.9 \tag{60a}$$

$$= -0.18 \times 10^{-3}\, exp(4t) + 0.562\, (-4t), \; 0.9 \leqslant t \leqslant 1.0 \tag{60b}$$

The suboptimal controls, when truncating after one term, are

$$u \simeq -3x - 3\epsilon g_x^1(t) \tag{61a}$$

$$v \simeq -3y - 3\epsilon g_y^1(t) \tag{61b}$$

The state model using the control in (61) will become

$$\dot{x} \simeq -4x - 2\epsilon y + 0.2\epsilon x(t-0.1) - 3\epsilon(g_x^1 - \epsilon g_y^1) \tag{62a}$$

$$\dot{y} \simeq -4y - 4\epsilon x - 0.1\epsilon y(t-0.1) - 3\epsilon(\epsilon g_x^1 + g_y^1) \tag{62b}$$

with $x(0) = y(0) = 1$. Figure 2 shows the state and control responses for a one-term truncation series. The near-optimum performance index turned out to be $J = 0.79$. All the computations were done on an IBM/1130 computer using FORTRAN IV.

The coupling procedure described here is computationally attractive, since all the calculations are based on lower-order unretarded and decoupled subsystems. For the two coupled subsystem formulation, it was shown that a near-optimum controller can be obtained which has an exact feedback portion and an approximate forward term. This result is of course expected, since all sensitivity methods give such near-optimum controls. Also, as in the unretarded nonlinear case [8.28], only one low-order Riccati equations must be solved, regardless of the order of approximation. The only extra computation required for each new approximation is the solution of the linear adjoint vector equation.

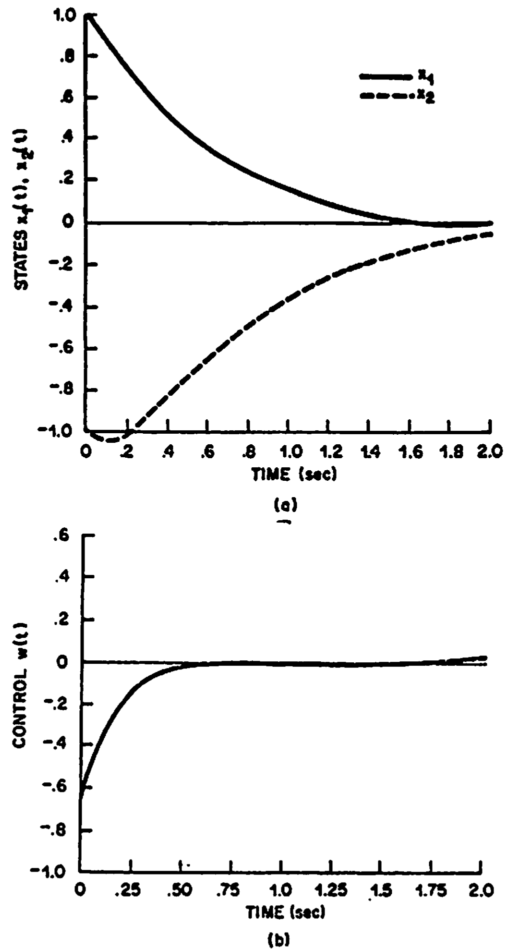

Figure 1. Time Responses for the Nonlinear TD System of Example 7: (a) States, and (b) Control

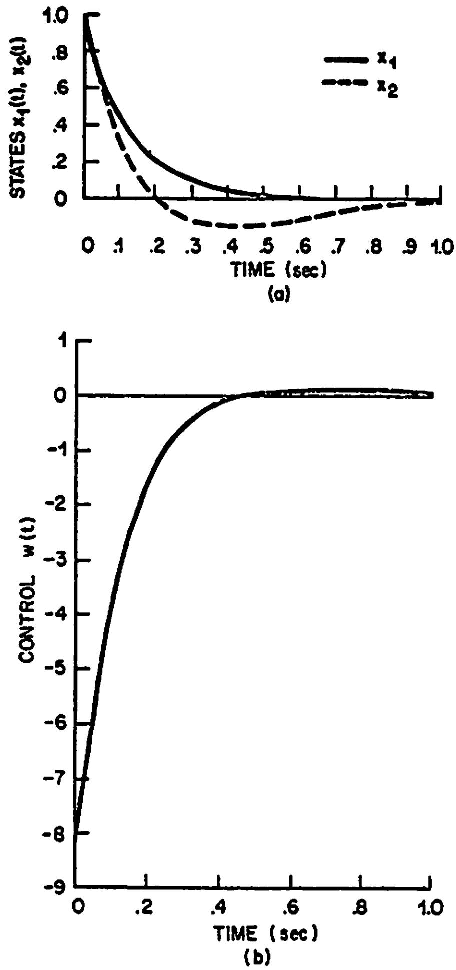

Figure 2. Time Responses for the Linear TD System of Example 8: (a) States, and (b) Control

8.3 HIERARCHICAL CONTROL

In this section three hierarchical control algorithms are briefly introduced to set the tone for their extensions in the next four sections for time-delay systems. The three algorithms are *Goal Coordination, Interaction Prediction* and *Costate Prediction.*

8.3.1 Goal Coordination

Consider a large-scale linear time-invariant system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t),\ \mathbf{x}(0) = \mathbf{x}_o \tag{1}$$

It is assumed that (1) can be decomposed into

$$\dot{\mathbf{x}}_i(t) = \mathbf{A}_i\mathbf{x}_i(t) + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{z}_i(t),\ \mathbf{x}_i(0) = \mathbf{x}_{io} \tag{2}$$

and an interaction term

$$\mathbf{z}_i(t) = \sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j \tag{3}$$

is a linear combination of the states of the other N-1 subsystems and $\mathbf{G}_{ij}$ is an $n_i \times n_j$ matrix. The original system's optimal control problem is reduced to the optimization of N subsystems which collectively satisfy (1) - (3) while minimizing

$$J = \sum_{i=1}^{N}\left[\tfrac{1}{2}\mathbf{x}_i'(t_f)\mathbf{Q}_i\mathbf{x}_i(t_f) + \tfrac{1}{2}\int_0^{t_f}\{\mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \mathbf{u}_i'(t)\mathbf{R}_i\mathbf{u}_i(t) + \mathbf{z}_i'(t)\mathbf{S}_i\mathbf{z}_i(t)\}dt\right] \tag{4}$$

where $\mathbf{Q}_i$ are $n_i \times n_i$ positive semi-definite matrices, $\mathbf{R}_i$ and $\mathbf{S}_i$ are $m_i \times m_i$ and $k_i \times k_i$ positive definite matrices with

$$n = \sum_{i=1}^{N} n_i\ ,\ m = \sum_{i=1}^{N} m_i\ ,\ k = \sum_{i=1}^{N} k_i\ \ ,\ \ k_i \leqslant n_i \tag{5}$$

The physical interpretation of the last term in the integrand of (4) is difficult at this point. In fact the introduction of this term is to avoid singular controls. The "Goal Coordination" or "Interaction Balance" approach of Mesarovic et. al. [8.29] as applied to the "linear-quadratic" problem by Pearson [8.30] and reported by Singh [8.20] is now presented.

In this decomposition of a large interconnected linear system the common coupling factors among its N subsystems are the "interaction" variables $\mathbf{z}_i(t)$ which along with the (2) - (3) constitute the "coupling" constraints. This formulation has been called "global" and is denoted by s_G. The following assumption is considered to hold. The global problem s_G is replaced by a family of N subproblems coupled together through a parameter vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$ and denoted by $s_i(\alpha_i)$, $i=1,2,\ldots,N$. In other words, the global system problem s_G is "imbedded" into a family of subsystem problems $s_i(\alpha)$ through an imbedding parameter α [8.31] in such a way that for a particular value of α^*, the subsystems $s_i(\alpha^*)$ $i=1,2,\ldots,N$ yield the desired solution to s_G. In terms of hierarchical control notation this imbedding concept is nothing but the notion of "coordination", while in mathematical programming problem terminology, it is denoted as the "master" problem [8.32]. Figure 1 shows a two-level control structure of a large-scale system. Under this strategy each local controller i receives α_i^ℓ from the coordinator (second level hierarchy), solves $s_i(\alpha_i^\ell)$ and transmits (reports) some function y_i^ℓ of its solution to the coordinator. The coordinator, in turn, evaluates the next updated value of α, *i.e.*

$$\alpha^{\ell+1} = \alpha^\ell + \epsilon^\ell \mathbf{d}^\ell \tag{6}$$

where ϵ^ℓ is the ℓth iteration step size and the update term $\mathbf{d}^\ell$, as will be shortly seen, is commonly taken as a function of "interaction error,"

$$\mathbf{e}_i(\alpha(t),t) = \mathbf{z}_i(\alpha(t),t) - \sum_{j=1, j\neq i}^{N} \mathbf{G}_{ij}\mathbf{x}_j(\alpha(t),t) \tag{7}$$

The imbedded interaction variable $\mathbf{z}_i(\cdot)$ in (7) can be considered as part of the control variable available to controller i in which case, the parameter vector $\alpha(t)$ serves as a set of "dual" variables or Lagrange multipliers corresponding to interaction equality constraints (3). The fundamental concept behind this approach is to convert the original system's minimization problem into an easier maximization one whose solution can be obtained in a two-level interactive scheme discussed above.

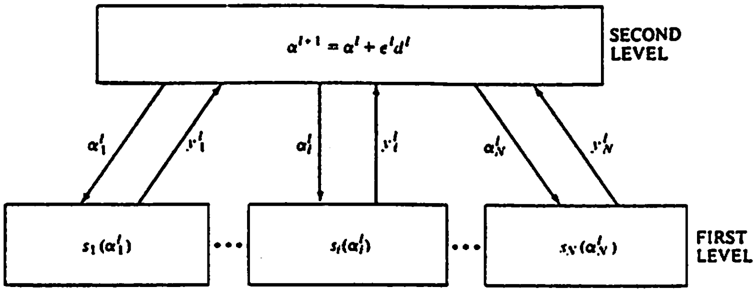

Figure 1. The Two Level Goal-Coordination Structure for Dynamic Systems

$$q(\alpha) = \underset{\mathbf{x,u,z}}{Min} \{L(\mathbf{x,u,z},\alpha)\} \tag{8}$$

subject to (2), where the Lagrangian L(·) is defined by

$$L(\mathbf{x,u,z},\alpha) = \sum_{i=1}^{N} \{\tfrac{1}{2}\mathbf{x}_i'(t_f)\mathbf{Q}_i\mathbf{x}_i(t_f) + \tfrac{1}{2}\int_0^{t_f} [\mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \mathbf{u}_i'(t)\mathbf{R}_i\mathbf{u}_i(t) +$$
$$\mathbf{z}_i'(t)\mathbf{S}_i\mathbf{z}_i(t) + 2\alpha_i'(\mathbf{z}_i(t) - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j(t))]dt\} \tag{9}$$

where the parameter vector α consists of k Lagrange multipliers. In this way the original constrained (subsystem interactions) optimization problem is changed to an unconstrained one. In other words, the constraint (3) is satisfied by determining a set of Lagrange multipliers α_i, $i = 1,2,...,k$. Under the cases like here where the constraint are convex, Geoffrion [8.33] and Singh [8.20] have shown that

$$\underset{\alpha}{Maximize}\ q(\alpha) \equiv \underset{\mathbf{u}}{Minimize}\ J \tag{10}$$

which indicates that minimization of J in (4) subject to (2) - (3) is equivalent to maximizing the dual function $q(\alpha)$ in (8) which respect to α. To facilitate the solution of this problem, it is observed that for a given set of Lagrange multipliers $\alpha = \alpha^*$, the Lagrangian (9) can be rewritten as

$$L(\mathbf{x},\mathbf{u},\mathbf{z},\alpha^*) = \sum_{i=1}^{N}\{1/2\mathbf{x}_i'(t_f)\mathbf{Q}_i(t_f)\ \mathbf{x}_i(t_f) + 1/2\int_0^{t_f}[\mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \mathbf{u}_i'(t)\mathbf{R}_i\mathbf{u}_i(t)$$
$$+ \mathbf{z}_i'S_i\mathbf{z}_i(t) + 2\alpha_i^{*\prime}\mathbf{z}_i\ (t) - 2\sum_{j\neq 1}^{N}\alpha_j^{*\prime}\mathbf{G}_{ji}\mathbf{x}_i(t)]dt\} \triangleq \sum_{i=1}^{N}L_i(\cdot) \tag{11}$$

which reveals that the decomposition is carried on to the Lagrangian in such a way that a sub-Lagrangian exists for each subsystem. Each subsystem would intend to minimize its own sub-Lagrangian L_i as defined by (11) subject to (2) and using the Lagrangian multipliers α^* which are treated as known functions at the first level of hierarchy. The result of each such minimization would allow one to determine the dual function $q(\alpha)^*$ in (8). At the second level, where the solutions of all first-level subsystems are known, the value of $q(\alpha)^*$ would be improved by a typical unconstrained optimization such as the Newton's method, the gradient method or the conjugate gradient method. The reason for a gradient-type method is due to the fact that the gradient of $q(\alpha)$ defined by

$$\nabla_\alpha q(\alpha)|_{\alpha_i = \alpha_i}^{*} = \mathbf{z}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N}\mathbf{G}_{ij}\mathbf{x}_j \triangleq \mathbf{e}_i,\ i=1,2,... \tag{12}$$

is nothing but the subsystems' interaction errors which are known through first-level solutions and $\nabla_{\mathbf{x}}\mathbf{f}$ defines the gradient of $\mathbf{f}$ with respect to $\mathbf{x}$. At the second-level the vector α is updated as indicated by (6) and Figure 1. If a gradient (steepest descent) method is employed the vector $\mathbf{d}^\ell$ in (6) is simply the ℓth iteration interaction error $\mathbf{e}^\ell(t)$. However, a more superior technique from a computational point of view is the conjugate gradient defined by

$$\mathbf{d}^{\ell+1}(t) = \mathbf{e}^{\ell+1}(t) + \gamma^{\ell+1}\mathbf{d}^\ell(t),\ \ 0\leqslant t\leqslant t_f \tag{13}$$

where

$$\gamma^{\ell+1} = \frac{\int_0^{t_f}(\mathbf{e}^{\ell+1}\ (t))'\mathbf{e}^{\ell+1}\ (t)dt}{\int_0^{t_f}(\mathbf{e}^\ell)'\mathbf{e}^\ell dt} \tag{14}$$

and $\mathbf{d}^0 = \mathbf{e}^0$. Once the error vector $\mathbf{e}(t)$ approaches zero the optimum hierarchical control is resulted. Below, a step-by-step computational procedure for the goal coordination method of hierarchal control is given [8.31]:

1. Algorithm - *Goal-Coordination Method*

Step 1: For each first-level subsystem, minimize each sub-Lagrangian L_i using a known Lagrange multiplier $\alpha_i = \alpha_i^*$. Since the subsystems are linear, a Riccati equation formulation can be used here. Store solutions.

Step 2: At second level, a conjugate gradient iterative method similar to (13) - (14) is used to update $\alpha^*_i(t)$ trajectories like (6). Once the total system interaction error in normalized form,

$$Error = \left[\sum_{i=1}^{N} \int_0^{t_f} \{\mathbf{z}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j\}'\{\mathbf{z}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j\} dt \right] / \Delta t$$

is sufficiently small an optimum solution has been obtained for the system. Here Δt is the step size of integration.

8.3.2 Interaction Prediction

An alternative approach in optimal control of hierarchical systems which has both open- and closed-loop forms is the "interaction prediction method" based on the initial work of Takahara [8.34] which avoids second-level gradient-type iterations. Consider a large-scale linear interconnected system which is decomposed into N subsystems each of which is described by

$$\dot{\mathbf{x}}_i(t) = \mathbf{A}_i\mathbf{x}_i(t) + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{z}_i(t),\ \mathbf{x}_i(0) = \mathbf{x}_{io},\ i = 1,2,\ldots,N \tag{15}$$

where the interaction vector $\mathbf{z}_i$ is

$$\mathbf{z}_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j(t) \ . \tag{16}$$

The optimal control problem at the first level is to find a control $\mathbf{u}_i(t)$ which satisfies (15) - (16) while minimizing a usual quadratic cost function

$$J_i = \tfrac{1}{2}\, \mathbf{x}_i'(t_f)\mathbf{Q}_i\mathbf{x}_i(t_f) + \tfrac{1}{2} \int_0^{t_f} \{\mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \mathbf{u}_i'(t)\mathbf{R}_i(t)\mathbf{u}_i(t)\} dt. \tag{17}$$

This problem can be solved by first introducing a set of Lagrange multipliers $\alpha_i(t)$ and costate (adjoint) vectors $\mathbf{p}_i(t)$ to augment the "interaction" equality constraint (16) and subsystem dynamic constraint (15) to the cost function's integrand, *i.e.* the ith subsystem Hamiltonian is defined by

$$H_i = \tfrac{1}{2}\, \mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \tfrac{1}{2}\, \mathbf{u}_i'(t)\mathbf{R}_i\mathbf{u}_i(t) + \alpha_i'\mathbf{z}_i - \sum_{\substack{j=1\\ j\neq i}}^{N} \alpha_j'\mathbf{G}_{ji}\mathbf{x}_i + \mathbf{P}_i'(\mathbf{A}_i\mathbf{x}_i + \mathbf{B}_i\mathbf{u}_i + \mathbf{z}_i) \tag{18}$$

Then the following set of necessary conditions can be written:

$$\dot{\mathbf{p}}_i = -\partial H_i/\partial \mathbf{x}_i = -\mathbf{Q}_i\mathbf{x}_i - \mathbf{A}_i'\mathbf{p}_i + \sum_{j=1}^{N} \mathbf{G}_{ji}'\alpha_i(t) \tag{19}$$

$$\mathbf{p}_i(t_f) = \partial(\tfrac{1}{2}\mathbf{x}_i'(t_f)\mathbf{Q}_i\mathbf{x}_i(t_f))/\partial\mathbf{x}_i(t_f) = \mathbf{Q}_i\mathbf{x}_i(t_f) \tag{20}$$

$$\Omega = \partial H_i/\partial_i = \mathbf{A}_i\mathbf{x}_i(t) + \mathbf{B}_i u_i(t) + \mathbf{z}_i(t),\ \mathbf{x}_i(0) = \mathbf{x}_{io} \tag{21}$$

$$0 = \partial H_i/\partial \mathbf{u}_i = \mathbf{R}_i\mathbf{u}_i^{(t)} + \mathbf{B}_i'\mathbf{P}_i^{(t)} \tag{22}$$

where the vectors $\alpha_i(t)$ and $\mathbf{z}_i(t)$ are no longer considered as unknows at the second level and in fact ω is augmented with $\alpha_i(t)$ to constitute a higher-dimensional coordination vector to be obtained shortly. For the purpose of solving the first-level problem it suffices to assume $(\alpha_i'(t) \mid \mathbf{z}_i'(t))$ as known.

Note that $\mathbf{u}_i$ (t) can be eliminated from (22)

$$\mathbf{u}_i(t) = -\mathbf{R}_i^{-1}\mathbf{B}_i'\mathbf{p}_i(t) \tag{23}$$

and substituted into (19) - (21) to obtain,

$$\dot{\mathbf{x}}_i(t) = \mathbf{A}_i\mathbf{x}_i(t) - S_i\mathbf{P}_i(t) + \mathbf{z}_i(t),\ \mathbf{x}_i(0) = \mathbf{x}_{io} \tag{24}$$

$$\dot{\mathbf{p}}_i(t) = -\mathbf{Q}_i\mathbf{x}_i(t) - \mathbf{A}_i'\mathbf{p}_i(t) + \sum_{j=1}^{N} \mathbf{G}_{ji}\alpha_i(t),\ \mathbf{p}_i(t_f) = \mathbf{Q}_i\mathbf{x}_i(t_f) \tag{25}$$

which constitute a two-point boundary-value (TPBV) problem and $\mathbf{S}_i \triangleq \mathbf{B}_i\mathbf{R}_i^{-1}\ \mathbf{B}_i'$. Using the usual Riccati formulation, this linear TPBV problem can be decoupled by introducing the following formulation. Here it is assumed that

$$\mathbf{p}_i(t) = \mathbf{K}_i(t)\mathbf{x}_i(t) + \mathbf{g}_i(t) \tag{26}$$

where $\mathbf{g}_i(t)$ is an n_i-dimensional open-loop "adjoint" or "compensation" vector. The resulting relations are given by

$$\dot{\mathbf{K}}_i(t) = -\mathbf{K}_i(t)\mathbf{A}_i - \mathbf{A}_i'\mathbf{K}_i(t) + \mathbf{K}_i(t)\mathbf{S}_i\mathbf{K}_i(t) - \mathbf{Q}_i \tag{27}$$

$$\dot{\mathbf{g}}_i(t) = -(\mathbf{A}_i - \mathbf{S}_i\mathbf{K}_i(t))'\mathbf{g}_i(t) + \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{G}_{ji}'\alpha_j'(t) \tag{28}$$

whose final conditions $\mathbf{K}_i(t_f)$ and $\mathbf{g}_i(t_f)$ and

$$\mathbf{K}_i(t_f) = \mathbf{Q}_i \ , \ \mathbf{g}_i(t_f) = \mathbf{0} \ . \tag{29}$$

Following this formulation, the first-level optimal control (23) becomes,

$$\mathbf{u}_i(t) = -\mathbf{R}_i^{-1}\mathbf{B}_i'\mathbf{K}_i(t)\mathbf{x}_i(t) - \mathbf{R}_i^{-1}\mathbf{B}_i'\mathbf{g}_i(t) \tag{30}$$

which has a partial feedback (closed-loop) term and a feed forward (open-loop) term.

The second-level problem is essentially updating the new coordination vector $(\alpha_i'(t) \mid z_i'(t))'$. For this purpose, define the additively separable Lagrangian

$$\begin{aligned} L = \sum_{i=1}^{N} L_i = \sum_{i=1}^{N} \Bigg[& \tfrac{1}{2}\, \mathbf{x}_i'(t_f)\mathbf{Q}_i\mathbf{x}_i(t_f) + \int_0^{t_f} \{\tfrac{1}{2}\mathbf{x}_i'(t)\mathbf{Q}_i\mathbf{x}_i(t) + \\ & \tfrac{1}{2}\mathbf{u}_i'(t)\mathbf{R}_i\mathbf{u}_i(t) + \alpha_i'(t)\mathbf{z}_i(t) - \sum_{\substack{j=1 \\ j\neq i}}^{N} \alpha_j'(t)\mathbf{G}_{ji}\mathbf{x}_j(t) + \\ & \mathbf{p}_i'(t)[-\dot{\mathbf{x}}_i(t) + \mathbf{A}_i\mathbf{x}_i(t) + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{z}_i(t)]\}dt \Bigg] \end{aligned} \tag{31}$$

The values of $\alpha_i(t)$ and $\mathbf{z}_i(t)$ can be obtained by

$$0 = \partial L_i(\cdot)/\partial \mathbf{z}_i(t) = \alpha_i(t) + \mathbf{p}_i(t) \tag{32}$$

$$0 = \partial L_i(\cdot)/\partial \alpha_i(t) = \mathbf{z}_i(t) - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{G}_{ji}\mathbf{x}_j(t) \tag{33}$$

which provide

$$\alpha_i(t) = \mathbf{p}_i(t)\ ,\ \mathbf{z}_i(t) = \sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{G}_{ji}\mathbf{x}_j(t). \tag{34}$$

The second-level coordination procedure at the $(\ell+1)th$ iteration is simply

$$\begin{bmatrix} \alpha_i(t) \\ \mathbf{z}_i(t) \end{bmatrix}^{\ell+1} = \begin{bmatrix} -\mathbf{p}_i(t) \\ \sum\limits_{\substack{j=1\\ j\neq i}}^{N} \mathbf{G}_{ji}\mathbf{x}_j(t) \end{bmatrix}^{\ell} \tag{35}$$

The interaction prediction method is formulated by the following algorithm.

2. Algorithm - ***Interaction Prediction***

Step 1: Solve N independent differential matrix Riccati equations (27) with final condition (29) and store $\mathbf{K}_i(t), i=1,2,\ldots,N$.

Step 2: For initial $\alpha_i^\ell(t), \mathbf{z}_i^\ell(t)$, solve the equation (28) with final condition (29), evaluate and store $\mathbf{g}_i(t)$, and $\mathbf{p}_i(t)$, $i=1,2,\ldots,N$.

Step 3: Solve the state equation

$$\dot{\mathbf{x}}_i(t) = (\mathbf{A}_i - \mathbf{S}_i\mathbf{K}_i(t))\mathbf{x}_i(t) - \mathbf{S}_i\mathbf{g}_i(t)\mathbf{z}_i(t), \mathbf{x}_i(0) = \mathbf{x}_{io} \tag{36}$$

and store $\mathbf{x}_i$ (t), $i=1,2,\ldots, N$.

Step 4: At the second-level, use the results of Steps 2 and 3 and (35) to update the coordination vector, $[\alpha_i'(t) \mid z_i'(t)]'$.

Step 5: Check for the convergence at the second level by evaluating the overall interaction error,

$$\mathbf{e}(t) = \sum_{i=1}^{N} \int_0^{t_f} \{\mathbf{z}_i(t) - \sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j(t)\}'\{\mathbf{z}_i(t) - \sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{G}_{ij}\mathbf{x}_j\} dt/\Delta t\ . \tag{37}$$

8.3.3 Costate Prediction

In this section a computationally effective approach for hierarchical control of nonlinear discrete-time systems due to Mahmoud et.al. [8.35], Hassan and Singh [8.36] is presented. The scheme is applied to nonlinear systems and its linear extension is rather straightforward.

Consider a nonlinear discrete-time system described by

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x},\mathbf{u},(k+1,k)),\ \mathbf{x}(0)=\mathbf{x}_o \tag{38}$$

with a quadratic cost function

$$J = \tfrac{1}{2} \sum_{k=0}^{K-1} (\mathbf{x}'(k)\mathbf{Q}\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}\mathbf{u}(k)). \tag{39}$$

The procedure begins by rewriting the nonlinear state equation (38) as

$$\mathbf{x}(k+1)=\mathbf{A}(\mathbf{x},\mathbf{u},(k+1,k))\mathbf{x}(k)+\mathbf{B}(\mathbf{x},\mathbf{u},(k+1,k))\mathbf{u}(k)+\mathbf{C}(\mathbf{x},\mathbf{u},(k+1,k)) \tag{40}$$

where $\mathbf{A}(\cdot)$, $\mathbf{B}(\cdot)$ are block diagonal matrices and $\mathbf{C}(\mathbf{x},\mathbf{u},(k+1,k))= \mathbf{C}_1(\mathbf{x},\mathbf{u},(k+1,k))\mathbf{x}(k)+ \mathbf{C}_2(\mathbf{x},\mathbf{u},(k+1,k))\mathbf{u}(k)$. It is noted that the reformulation (40) of the system (38) is always possible. Moreover, for N blocks in $\mathbf{A}$ and $\mathbf{B}$, it is assumed that matrices $\mathbf{Q}$ and $\mathbf{R}$ also have N blocks. The basic reason for the reformulation (40) is to provide "predicted" state and control vectors $\mathbf{x}^*$ and $\mathbf{u}^*$ to fix the arguments in the nonlinear coefficient matrices $\mathbf{A}(\cdot)$, $\mathbf{B}(\cdot)$, $\mathbf{C}_1(\cdot)$ and $\mathbf{C}_2(\cdot)$. Therefore the problem (38) - (40) can be rewritten as

$$\min \tilde{J} = \tfrac{1}{2} \sum_{k=0}^{K-1} (\mathbf{x}'(k)\mathbf{Q}\mathbf{x}(k)+\mathbf{u}'(k)\mathbf{R}\ \mathbf{u}(k)) \tag{41}$$

subject to

$$\mathbf{x}(k+1)=\mathbf{A}(\mathbf{x}^*,\mathbf{u}^*,(k+1))\mathbf{x}(k)+\mathbf{B}(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\mathbf{u}(k)+\mathbf{C}(\mathbf{x}^*,\mathbf{u}^*,(k+1,k)) \tag{42}$$

$$\mathbf{x}^*(k) = \mathbf{x}(k)\ ,\ \mathbf{u}^*(k) = \mathbf{u}(k). \tag{43}$$

The modified problem can be solved by defining a Hamiltonian

$$\begin{aligned} H(\cdot)=&\tfrac{1}{2}\mathbf{x}'(k)\mathbf{Q}\mathbf{x}(k)+\tfrac{1}{2}\mathbf{u}'(k)\mathbf{R}\mathbf{u}(k)+\mathbf{p}'(k+1)\{\mathbf{A}(\mathbf{x}^*,\ \mathbf{u}^*,(k+1,k))\mathbf{x}(k)+ \\ &\mathbf{B}(\mathbf{x}^*,\ \mathbf{u}^*,(k+1,k))\mathbf{u}(k)+\mathbf{C}(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\} + \\ &\alpha'(k)(\mathbf{x}(k)-\mathbf{x}^*(k)) + \beta'(k)(\mathbf{u}(k)-\mathbf{u}^*(k)). \end{aligned} \tag{44}$$

In view of the assumptions made on matrices and vectors $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{Q}$ and $\mathbf{R}$, it is clear that the Hamiltonian $H(\cdot)$ in (44) is additively separable for given $\mathbf{x}^*$ and $\mathbf{u}^*$, *i.e.*

$$H = \sum_{i=1}^{N} H_i = \sum_{i=1}^{N} \{ \tfrac{1}{2}\mathbf{x}_i'(k)\mathbf{Q}\mathbf{x}_i(k) + \tfrac{1}{2}\mathbf{u}_i'(k)\mathbf{R}\mathbf{u}_i(k) +$$

$$\mathbf{p}_i'(k+1)[\mathbf{A}_i(\cdot)\mathbf{x}_i(k) + \mathbf{B}_i(\cdot)\mathbf{u}_i(k)] + \mathbf{c}_i(\cdot)\} +$$

$$\alpha_i'(k)[\mathbf{x}_i(k) - \mathbf{x}_i^*(k)] + \beta_i'(k)[\mathbf{u}_i(k) - \mathbf{u}_i^*(k)]\} . \tag{45}$$

the necessary conditions for optimality are given by

$$0 = \partial H_i/\partial \mathbf{u}_i \tag{46}$$

$$\mathbf{x}_i(k+1) = \partial H_i/\partial \mathbf{p}_i(k+1) \tag{47}$$

$$\mathbf{p}_i(k) = -\partial H_i/\partial \mathbf{x}_i(k) \tag{48}$$

$$0 = \partial H/\partial \alpha,\ 0 = \partial H/\alpha\beta \tag{49}$$

$$0 = \partial H/\partial \mathbf{x}^*(k),\ 0 = \partial H/\partial \mathbf{u}^*(k). \tag{50}$$

Relation (46) yields an expression for $\mathbf{u}_i(k)$:

$$\mathbf{u}_i(k) = -\mathbf{R}_i^{-1}\{\mathbf{B}_i'(\mathbf{x}^*, \mathbf{u}^*, (k+1,k))\mathbf{p}_i(k+1) + \beta_i(k)\} \tag{51}$$

and substituting $\mathbf{u}_i(k)$ in Equation (47) yields a new expression for $\mathbf{x}_i(k+1)$, *i.e.*

$$\mathbf{x}_i(k+1) = \mathbf{A}_i(\mathbf{x}^*, \mathbf{u}^*,(k+1,k))\mathbf{u}_i(k) - \mathbf{B}_i(\mathbf{x}^*,\mathbf{u}^*,(k+1,k)). \tag{52}$$

$$\mathbf{R}_i^{-1}\{\mathbf{B}_i'(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\mathbf{p}_i'(k+1) + \beta_i(k)\} + \mathbf{c}_i(\mathbf{x}^*, \mathbf{u}^*, (k+1,k))$$

with $\mathbf{x}_i(0) = \mathbf{x}_{io}$. The condition (48) gives the adjoint equation

$$\mathbf{p}_i(k) = \mathbf{Q}_i\mathbf{x}_i(k) + \mathbf{A}_i'(\mathbf{x}^*, \mathbf{u}^*, (k+1,k))\mathbf{p}_i(k+1) + \alpha_i(k),\ \mathbf{p}_i(K) = 0 \tag{53}$$

the necessary condition (49) leads to the equality constraints (43), *i.e.*

$$\mathbf{x}^*(k) = \mathbf{x}(k),\ \mathbf{u}^*(k) = \mathbf{u}(k). \tag{54}$$

The first of the two conditions in (50) leads to an expression for $\alpha(k)$, *i.e.*

$$\alpha(k) = \mathbf{F}_\mathbf{x}'(\mathbf{x}^*,\mathbf{u}^*,\mathbf{x}, (k+1,k)) + \mathbf{D}_\mathbf{x}'(\mathbf{x}^*, \mathbf{u}^*, (k+1,k))\mathbf{p}(k+1) \tag{55}$$

where

$$
\begin{aligned}
\mathbf{F}_{\mathbf{x}}(\cdot) &= \partial \mathbf{z}(\cdot)/\partial \mathbf{x}_{*} = \partial\{\mathbf{A}(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\mathbf{x}(k)\}/\partial \mathbf{x}^* \\
\mathbf{G}_{\mathbf{x}}(\cdot) &= \partial \mathbf{y}(\cdot)/\partial \mathbf{x}^* = \partial\{\beta(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\mathbf{u}(k)\}/\partial \mathbf{x}^* \\
\mathbf{D}_{\mathbf{x}}(\cdot) &= \partial \mathbf{c}(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))/\partial \mathbf{x}^* \; .
\end{aligned} \tag{56}
$$

Finally, in order to obtain an expression for $\beta(k)$, the second condition in (50) yields

$$
\beta(k) = \mathbf{F}_u'(\mathbf{x}^*, \mathbf{u}^*, \mathbf{x}_l, (k+1,k)) + \mathbf{G}_u'(\mathbf{x}^*, \mathbf{u}^*, \mathbf{u}, (k+1, k)) + \mathbf{D}_{\mathbf{u}}'(\mathbf{x}^*,\mathbf{u}^*,(k+1,k)) \tag{57}
$$

where $\mathbf{F}_{\mathbf{u}}(\cdot)$, $\mathbf{G}_{\mathbf{u}}(\cdot)$ and $\mathbf{D}_{\mathbf{u}}(\cdot)$ and derivatives of the expressions in the brackets of (56) with respect to $\mathbf{u}^*$. The following algorithm summarizes the costate prediction method.

3. Algorithm - *Costate Prediction Approach*

Step 1. Set iteration index $\ell=1$ and guess vectors $\mathbf{p}^1,\mathbf{x}^{*1},\mathbf{u}^{*1}$ and β^1.

Step 2. At first-level, substitute $\mathbf{p}^\ell(k)$, $\mathbf{x}^{*\ell}(k)$, $\mathbf{u}^{*\ell}(k)$ and $\beta^\ell(k)$, $k=0,\dots, K-1$ in (51)-(52) to obtain $(\mathbf{u}_i^\ell(k)$ and $\mathbf{x}_i^\ell(k)$ for $i=1,2,\dots, N$. Similarly use (57) to find $\beta^\ell(k)$.

Step 3. At second-level use $\mathbf{x}_i^\ell(k)$, $\mathbf{u}_i^\ell(k)$ and (53)-(57) to update

$$
\begin{aligned}
&\mathbf{q}^\ell = (\mathbf{p}_i^\ell,\mathbf{x}^{*\ell},\mathbf{u}^{*\ell}, \beta^\ell), i.e. \\
&\quad \mathbf{p}_i^{\ell+1}(k) = \mathbf{Q}_i\mathbf{x}_i^\ell(k) + \mathbf{A}_i'(\mathbf{x}^*,\mathbf{u}^*,(k+1,k))\mathbf{p}_i^{\ell+1}(k+1)+\alpha_i^\ell(k) \\
&\quad \mathbf{x}^{*\ell+1}(k) = \mathbf{x}^\ell(k) \\
&\quad \mathbf{u}^{*\ell+1}(k) = \mathbf{u}^\ell(k) \\
&\quad \beta^{\ell+1}(k) = \{F_{\mathbf{u}}'(\cdot)\,|_{(\mathbf{x}^{*\ell},\mathbf{u}^{*\ell},\mathbf{x}^\ell)} + G_{\mathbf{u}}'(\cdot)\,|_{x^{*\ell},u^{*\ell},u^\ell} + D_{\mathbf{u}}'(\cdot)\,|_{(\mathbf{x}^{*\ell},\mathbf{u}^{*\ell})}\}\mathbf{p}(k+1)
\end{aligned} \tag{59}
$$

Step 4. If $\mathbf{q}^{\ell+1}(k) = \mathbf{q}^\ell(k)$ for $k=0,1,\dots,K-1$, stop and $\mathbf{u}^{\ell+1}(k)$ is the optimal control. Otherwise go to Step 2.

4. Example

In this example the interaction prediction algorithm is used to hierarchically optimize a two-subsystem system.

Consider a fourth-order system

$$\dot{\mathbf{x}} = \left[\begin{array}{cc|cc} 2 & 0.1 & 0.01 & 0 \\ 0.2 & -1 & 0.10 & -0.5 \\ \hline 0.05 & 0.15 & 1 & 0.05 \\ 0 & -0.2 & -0.25 & -1.2 \end{array}\right]\mathbf{x} + \left[\begin{array}{c|c} 1 & 0 \\ 0.1 & 0 \\ \hline 0 & 0.5 \\ 0 & 0.25 \end{array}\right]\mathbf{u}$$

with $\mathbf{x}(0) = (-1,0.1,1.0,-0.5)'$ and a quadratic cost function with $\mathbf{Q} = diag(2,1,1,2), \mathbf{R} = diag(1,2)$ and no terminal penalty. It is desired to use the interaction prediction method to find an optimal control for $t_f = 1$.

Solution: The system was divided into two 2nd-order subsystems and steps outlined in Algorithm 2 were applied. At the first step, two independent differential matrix Riccati equations were solved by using the standard Runge-Kutta methods. The elements of the Riccati matrix were fitted in by quadratic polynomial in Chebyschev sense [8.37] for computational convenience:

$$\mathbf{K}_1(t) = \begin{bmatrix} 4.44+0.32t+1.26t^2 & 0.09+0.007t-0.027t^2 \\ 0.09+0.007t-0.027t^2 & 0.5+0.034t-0.141t^2 \end{bmatrix}$$

$$\mathbf{K}_2(t) = \begin{bmatrix} 2.87-5.26t+2.42t^2 & -0.1+0.16t-0.054t^2 \\ -0.1+0.16t-0.054t^2 & 0.73+0.118t-0.83t^2 \end{bmatrix}$$

At the first level, a set of two second-order adjoint equations of the form and two subsystem state equations as in Step 3 of Algorithm 2 using 4th-order Runge-Kuta method and initial values

$$\alpha_1(t) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \ \mathbf{x}_1(0) = \begin{bmatrix} -1 \\ 0.1 \end{bmatrix}, \ \mathbf{z}_1(t) = \mathbf{G}_{12}\mathbf{x}_2(0) = \begin{bmatrix} 0.01 \\ 0.35 \end{bmatrix}$$
$$\alpha_2(t) = \begin{bmatrix} 0.75 \\ 0.75 \end{bmatrix}, \ \mathbf{x}_2(0) = \begin{bmatrix} 1.0 \\ 0.5 \end{bmatrix}, \ \mathbf{z}_2(t) = \mathbf{G}_{21}\mathbf{x}_1(0) = \begin{bmatrix} -0.035 \\ -0.02 \end{bmatrix} .$$

were solved at the second level, the interaction vectors

$$[\alpha_{11}(t),\alpha_{12}(t),\mathbf{z}_{11}(t),\mathbf{z}_{12}(t)]' \ and \ [\alpha_{21}(t),\alpha_{22}(t),\mathbf{z}_{21}(t),\mathbf{z}_{22}(t)]'$$

were predicted using the recursive relations (35) and at each information exchange iteration the total interaction error (37) was evaluated for $\Delta t = 0.1$ and a cubic spline interpolator program. The interaction error was reduced to 3.5113456×10^{-6} in six interactions as shown in Figure 2.

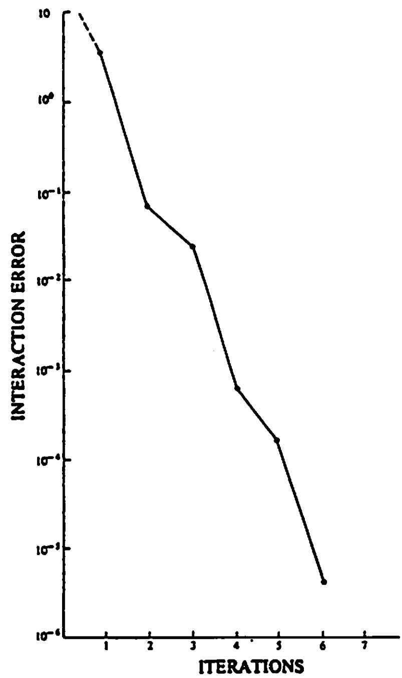

Figure 2. Interaction Error vs. Iterations For the Interaction Prediction Example 4

5. A CAD Example

Below, the same problem is solved using a computer-aided design package - LSSPAK/PC using an IBM PC [8.38]

```
******************************************************************************
**   INTRPRD solves a two subsystem hierarchical control system using the method   **
**   of interaction prediction. The algorithm may be found in Reference [8.31].     **
**   Note: This program is not good for systems with more than                      **
**            two subsystems. Intended for use with DOS 2.1.                        **
**                                                                                  **
******************************************************************************
```

Optimization via the interaction prediction method.

Total No. of 2nd level interactions = 4

Error tolerance for multi-level iterations = .001

Order of lst subsystem n1 = 2

Order of 2nd subsystem n2 = 2

Order of lst subsystem control vector r1 = 1

Order of 2nd subsystem control vector r2 = 1

Initial time (to): 0

Final time (tf): 1.

Step size (Dt): .2

Matrix A1

0.200D+01 0.100D+00

0.200D+00 -.100D+01

Matrix B1

0.100D+01

0.100D+00

Matrix R1

0.100D+01

Matrix G12

0.100D-01 0.000D+00

0.100D+00 -.500D+00

Lagrange multiplier Alpha-1

All (1): .5 All (2): .5

Initial conditions × 2

x2(1): 1 x2(2): -.5

Initial interaction vector Z1 for subsystem 1

0.0100

0.3500

Matrix K1

0.444D+01 0.320D+00 -.126D+01

0.900D-01 0.700D-02 -.270D-01

0.500D+00 0.340D-01 -.141D+00

Matrix K2

0.287D+01 -.526D+01 0.242D+01

-.100D+00 0.160D+00 -.540D-01

0.730D+00 0.118D+00 -.830D+00

Matrix A2

0.100D+01 0.500D-01

-.250D+00 0.200D+00

Matrix B2

0.500D+00

0.250D+00

Matrix R2

0.200D+01

Matrix G21

0.500D-01 0.150D+00

0.000D+00 -.200D+00

Lagrange multiplier Alpha-2

Al2 (1): .75

Al2 (2): .75

Initial conditions × 1

x1(1): -1

x1(2): .1

At second level iteration No. 1 interaction error = 0.403D+00

At second level iteration No. 2 interaction error = 0.230D-02

At second level iteration No. 3 interaction error = 0.977D-03

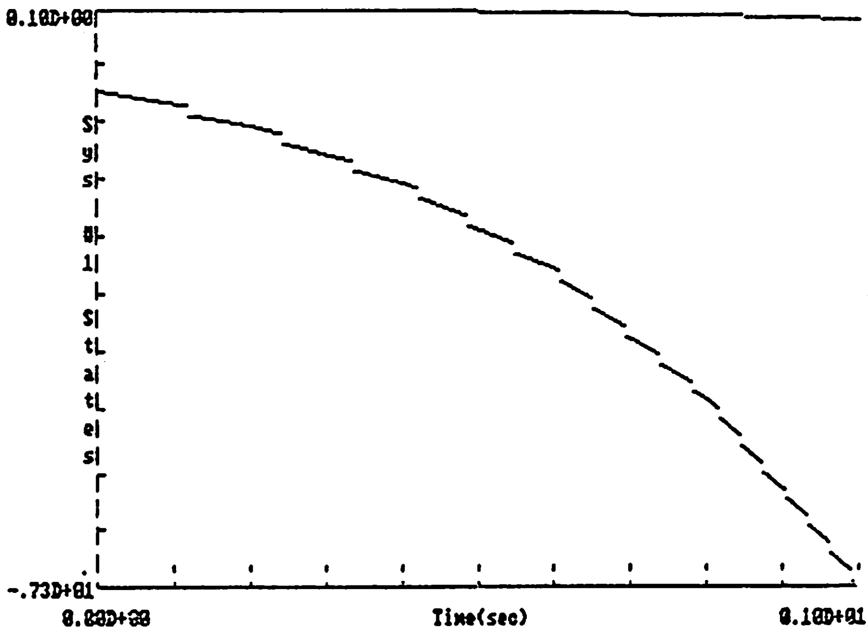

Figure 3a. Optimal Trajectories of Example 4 on LSSPAK/PC
Using Interaction Prediction Program INTRPRD:
Part (a) - Subsystem No. 1 Optimal State Trajectories

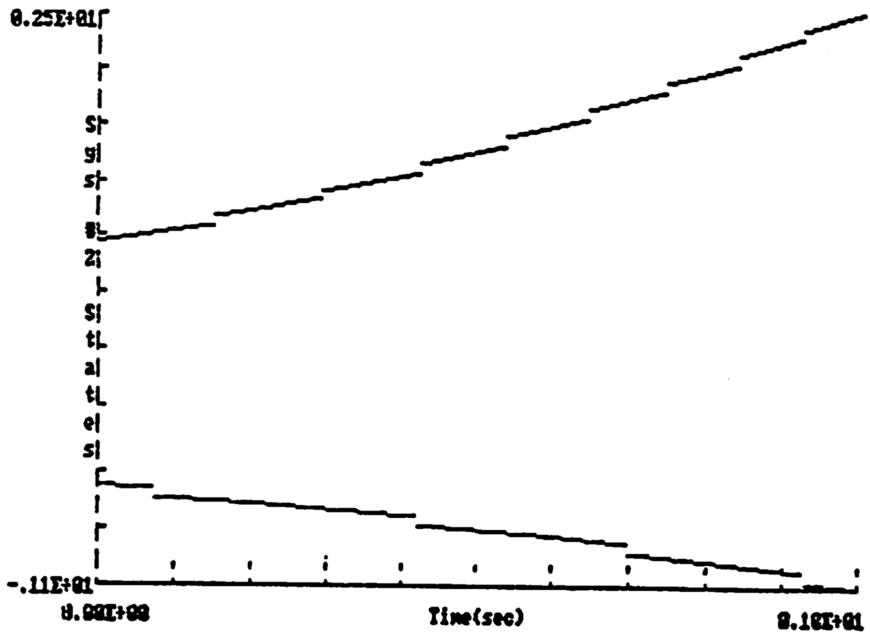

Figure 3b. Part (b) - Subsystem No. 1 Optimal Control Trajectory

Figure 3c. Part (c) - Subsystem No. 2 Optimal State Trajectories

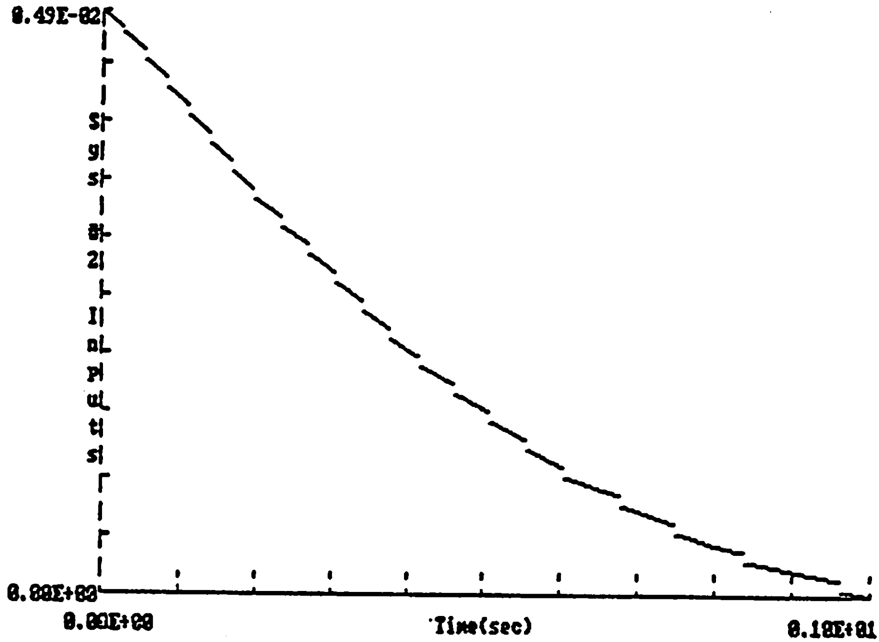

Figure 3d. Part (d) - Subsystem No. 2 Optimal Control Trajectory

8.4 HIERARCHICAL CONTROL OF SERIAL TIME-DELAY SYSTEMS

In this section the hierarchical control of a linear large-scale system with a serial structure and time-delay linear interaction shown in Figure 1 is considered. The state and output equations of this system are represented by

$$\dot{\mathbf{x}}_1(t) = \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{B}_1\mathbf{u}_1(t) + \mathbf{v}_1(t) \tag{1a}$$

$$\mathbf{y}_1(t) = \mathbf{C}_1\mathbf{x}_1(t) \tag{1b}$$

$$\dot{\mathbf{x}}_2(t) = \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{B}_2\mathbf{u}_2(t) + \mathbf{D}_2\mathbf{x}_1(t-h_1) + \mathbf{v}_2(t) \tag{1c}$$

$$\mathbf{y}_2(t) = \mathbf{C}_2\mathbf{x}_2(t) \tag{1d}$$

$$\vdots \tag{1e}$$

$$\dot{\mathbf{x}}_N(t) = \mathbf{A}_N\mathbf{x}_N(t) + \mathbf{B}_N\mathbf{u}_N(t) + \mathbf{D}_N\mathbf{x}_{N-1}(t-h_{N-1}) + \mathbf{v}_N(t) \tag{1f}$$

$$\mathbf{y}_N(t) = \mathbf{C}_N\mathbf{x}_N(t) \tag{1g}$$

where $\mathbf{x}_i(t)$, $\mathbf{y}_i(t)$, and $\mathbf{u}_i(t)$ are n_i, m_i, and r_i-dimensional state, output, and control vectors, respectively, $\mathbf{v}_i(t)$ is the ith external input, and h_i, $i=1,2,\ldots,N-1$, are the delays between the N subsystems. The cost function to be minimized is

$$J = \tfrac{1}{2}\sum_{i=1}^{N}\left\{\mathbf{y}_i(t_f) - \mathbf{y}_i^*(t_f)\right\}'\mathbf{F}_i\left\{\mathbf{y}_i(t_f) - \mathbf{y}_i^*(t_f)\right\} + \tfrac{1}{2}\int_0^{t_f}\sum_{i=1}^{N}\left\{(\mathbf{y}_i-\mathbf{y}_i^*)'\,\mathbf{Q}_i\,(\mathbf{y}_i-\mathbf{y}_i^*) + \mathbf{u}_i'\mathbf{R}_i u_i\right\}dt \tag{2}$$

where $\mathbf{y}_i^* = \mathbf{y}_i^*(t)$, $i=1,2,\ldots,N$, are the desired outputs of the N subsystems and the bracketed term outside the integral sign corresponds to the terminal output penalty function. It is assumed that $t_f > h_1+h_2+\ldots+h_{N-1}$ The application of the maximum principle of time-delay systems, as demonstrated earlier, results in a set of $2n$ TPBV problems with delay and advance terms which is clearly impractical. Here $n = n_1 + \cdots + n_N$. Any other approach, such as dynamic programming, as demonstrated by Singh [8.20], is still formidable and computationally infeasible.

The approach which seems most suitable is an approximate optimum solution based on hierarchical control. Under a hierarchical structure, the first-level problem would be to minimize

$$J_i = \tfrac{1}{2}\left[\mathbf{y}_i(t_f) - \mathbf{y}_i^*(t_f)\right]' F_i \left[\mathbf{y}_i(t_f) - \mathbf{y}_i^*(t_f)\right] + \tfrac{1}{2}\int_0^{t_f} \left\{ (\mathbf{y}_i - \mathbf{y}_i^*)' Q_i (\mathbf{y}_i - \mathbf{y}_i^*) + u_i' R_i u_i \right\} dt \tag{3}$$

subject to

$$\dot{\mathbf{x}}_i(t) = A_i \mathbf{x}_i(t) + B_i \mathbf{u}_i(t) + D_i \mathbf{x}_{i-1}(t - h_{i-1}) + \mathbf{v}_i(t) \tag{4}$$

$$\mathbf{y}_i(t) = C_i \mathbf{x}_i(t) \tag{5}$$

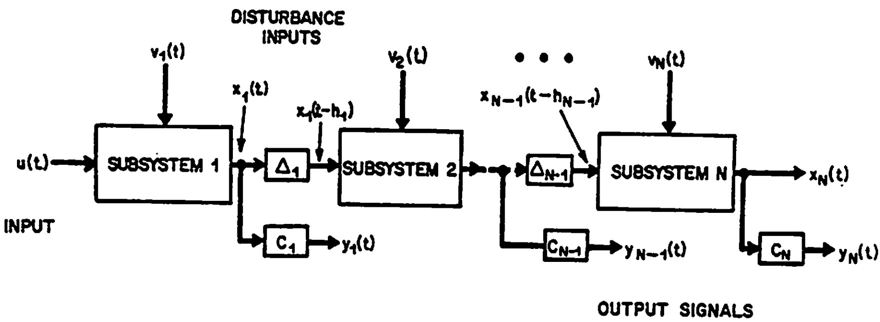

Figure 1. A Block Diagram Representation of a Serial Large-Scale TD System

Then assuming that the disturbance $\mathbf{v}_i(t)$ and an estimate $\hat{\mathbf{x}}_{i-1}(t-h_{i-1})$ of $\mathbf{x}_{i-1}(t-h_{i-1})$ is transferred down from the second level, the first-level problem (3)-(5) can be readily solved through a Riccati formulation. Under such conditions, the second-level problem would be an iterative search to reduce the interaction error between $\mathbf{x}_1(t-h_j)$ and $\hat{\mathbf{x}}_j(t-h_j)$ for $j=1,2,\ldots,N-1$ to a small value. Such approaches as "goal coordination" or "interaction prediction" to be discussed in the next sections may well be used. The near-optimum scheme suggested by Singh [8.20] is to stop the second-level iteration once the state of any given subsystem $\mathbf{x}_i(t)$ becomes known and simply transfer down an estimate $\hat{\mathbf{x}}_i(t-h_j)$as an extra known external input to be incorporated in the optimization of the $(i+l)$th problem. Thus, the ith subsystem optimization at the first level given by (3)-(5) after relacing $\mathbf{x}_{i-1}(t-h_{-1})$ by $\hat{\mathbf{x}}_{i-1}(t-h_{i-1})$ would be summarized as

$$\mathbf{u}_i(t) = -\mathbf{R}_i^{-1}\mathbf{B}'_i \left\{ \mathbf{K}_i(t)\mathbf{x}_i(t) + \mathbf{g}_i(t) \right\} \tag{6}$$

where Riccati matrix $\mathbf{K}_i(t)$ and adjoint vector $\mathbf{g}_i(t)$ are obtained from

$$\dot{\mathbf{K}}_i(t) = \mathbf{K}_i(t)\mathbf{A}_i - \mathbf{A}'_i(t) + \mathbf{K}_i(t)S_i\mathbf{K}_i(t) - \mathbf{C}'_i\mathbf{Q}_i\mathbf{C}_i \tag{7a}$$

$$\mathbf{K}_i(t_f) = \mathbf{C}'_iF_i\mathbf{C}_i \tag{7b}$$

$$\dot{\mathbf{g}}_i(t) = -\left[\mathbf{A}_i - S_i\mathbf{K}_i(t) \right]' \mathbf{g}_i(t) + \mathbf{K}_i(t)\ \mathbf{w}_i(t) - \mathbf{C}_i\mathbf{Q}_i\mathbf{y}_i^*(t) \tag{8a}$$

$$\mathbf{g}_i(t_f) = \mathbf{C}'_i\mathbf{F}_i\mathbf{y}_i^*(t_f) \tag{8b}$$

where $\mathbf{S}_i = \mathbf{B}_i\mathbf{R}_i^{-1}\mathbf{B}'_i$ and

$$\begin{aligned} \mathbf{w}_i(t) &= \mathbf{v}_i(t) + \mathbf{D}_i\hat{\mathbf{x}}_{i-1}\ (t-h_{i-1}),\ \text{for } i=2,3,\ldots,N \\ &= \mathbf{v}_i(t),\ \text{for } i=1 \end{aligned} \tag{9}$$

The closed-loop system becomes

$$\dot{\mathbf{x}}_i(t) = \left[\mathbf{A}_i - \mathbf{S}_i\mathbf{K}_i\ (t) \right] \mathbf{x}_i(t) + \mathbf{D}_i\dot{\mathbf{x}}_{i-1}(t-h_{i-1}) + \mathbf{v}_i(t) \tag{10}$$

The near-optimum control of the serial TD system is summarized by the following algorithm.

1. Algorithm.

Step 1: Set $i=1$ and input t_f, $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{v}_i$, $\cdots$

Step 2: Solve $\mathbf{K}_i(t)$ and $\mathbf{g}_i(t)$ using (7) and (8) and store.

Step 3: Integrate (10) to compute $\mathbf{x}_i(t)$ and store $\hat{\mathbf{x}}_i(t-h_i) = \mathbf{x}_i(t-h_i)$

Step 4: If $N < i = i + 1$, go to Step 2.

Step 5: Stop.

The following example illustrates Algorithm 1.

2. Example. Consider a three-subsystem serial TD system defined by the following state and output equations:

Subsystem 1:

$$\dot{\mathbf{x}}_1 = \mathbf{A}_1\mathbf{x}_1 + \mathbf{B}_1\mathbf{u}_1 + \mathbf{v}_1, \quad \mathbf{y}_1 = \mathbf{C}_1\mathbf{x}_1 \tag{11}$$

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0.1 \\ -1 & -1 \end{bmatrix},$$

$$\mathbf{B}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} 1.5 \\ 0.2 \end{bmatrix}, \quad \mathbf{C}_1 = [1 \quad 0.1]$$

$$Q_1 = R_1 = F_1 = [1], \quad t_0 = 0, t_f = 4, \quad \mathbf{x}_1(0) = [1 \quad -1]'$$

Subsystem 2:

$$\dot{\mathbf{x}}_2 = \mathbf{A}_2\mathbf{x}_2 + \mathbf{B}_2\mathbf{u}_2 + \mathbf{D}_2\mathbf{x}_1(t-h_1) + \mathbf{v}_2,\ \ \mathbf{y}_2 = \mathbf{C}_2\mathbf{x}_2 \tag{12}$$

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 0.1 \\ 0.1 & -1 & 0.2 \\ 1 & 0.85 & -0.9 \end{bmatrix},\ B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix},\ D_2 = \begin{bmatrix} 0.15 & -0.2 \\ 0.25 & 0.8 \\ 0.1 & -0.1 \end{bmatrix}$$

$$\mathbf{Q}_2 = \mathbf{F}_2 = [2]\ \ \mathbf{R}_2 = 2I_2,\ \ \mathbf{C}_2 = [10.5\ \ 0.1]\ ,\ h_1 = 0.05$$

$$\mathbf{x}_2(0) = [1\ \ 0.5\ \ -1]'$$

Subsystem 3:

$$\dot{\mathbf{x}}_3 = \mathbf{A}_3\mathbf{x}_3 + \mathbf{B}_3\mathbf{u}_3 + \mathbf{D}_3\mathbf{x}_2(t-h_2) + \mathbf{v}_2,\ \ \mathbf{y}_3 = \mathbf{C}_3\mathbf{x}_3 \tag{13}$$

$$\mathbf{A}_3 = \begin{bmatrix} 0.1 & 0.5 & -0.25 \\ 0.5 & 0 & 0.8 \\ -1 & 1 & -2 \end{bmatrix},\ \mathbf{B}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\ \mathbf{v}_3 = \begin{bmatrix} 1.2 \\ 1 \\ -1 \end{bmatrix},$$

$$\mathbf{D}_3 = \begin{bmatrix} 1 & 0.2 & 0.5 \\ -0.5 & 0.1 & 0.8 \\ 0.75 & 1 & -1 \end{bmatrix}$$

$$Q_3 = R_3 = [2],\ F_3 = [0.5],\ \mathbf{C}_3 = [2\ \ 1\ \ 1]$$
$$h_2 = 0.1,\ \mathbf{x}_3(0) = [1\ \ -0.5\ \ 1]'$$

It is desired to find the optimal controls for the three subsystems using Algorithm 1.

As the first step, one should solve three differential matrix Riccati equations of types described by (7). This was done on an HP-9845 computer using a fourth-order Runge-Kutta method, and the solutions were fit in second-order polynomials. The step size of 0.05 was chosen to conveniently recover the delayed quantities defined in (12) and (13). Next the adjoint equation (10) was solved for $i=1,2$ and 3. The resulting optimal controls and outputs for the three subsystems are shown in Figure 2.

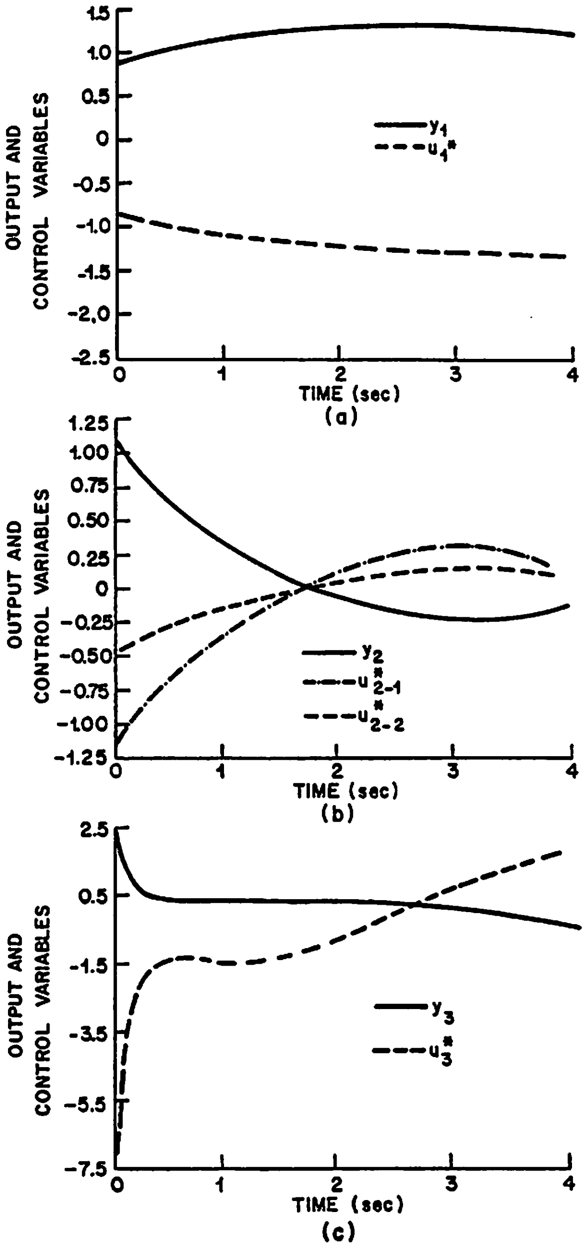

Figure 2. Output and Control Time Responses of the Serially Time-Delayed Example 2: (a) Subsystem 1, (b) Subsystem 2 and (c) Subsystem 3

8.5 HIERARCHICAL CONTROL OF NONSERIAL TIME-DELAY SYSTEMS

Consider a large-scale TD system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{C}(t)\mathbf{x}(t-h_x) + \mathbf{D}(t)\mathbf{u}(t-h_u) \ , \ t \geq t_0 \tag{1a}$$

$$\mathbf{x}(t) = \beta(t) \ , \ t_0 - h_x \leq t \leq t_0, \ \mathbf{u}(t) = \alpha(t), \ t_0 - h_u \leq t \leq t_0 \tag{1b}$$

where $\mathbf{x}(t) \in R^n$ and $\mathbf{u}(t) \in R^m$ are respectively, state and control vectors; $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$ and $\mathbf{D}(t)$ are real, piecewise-continuous matrices of appropriate dimensions defined on the appropriate intervals; t_0 is the initial process time; $\beta(t)$ and $\alpha(t)$ are specified initial functions; and h_x and h_u are constant positive scalars. The cost functional to be minimized is

$$J = \tfrac{1}{2}\, \mathbf{x}'(t_f)\mathbf{F}\mathbf{x}(t_f) + \tfrac{1}{2} \int_{t_0}^{t_f} \Big[\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t) \Big]\, dt \tag{2}$$

where matrix $\mathbf{F}$ is symmetric positive-semidefinite; the matrix $\mathbf{Q}(t)$ is symmetric positive-semidefinite and piecewise-continuous; the matrix $\mathbf{R}(t)$ is symmetric, piecewise-continuous. The problem is to find a control $\mathbf{u}(t)$, $t_0 \leq t \leq t_f$, which for fixed final time t_f and free final state $\mathbf{x}(t_f)$ minimizes the cost functional J in (2).

Assume that the state vector $\mathbf{x}$ and the control vector $\mathbf{u}$ are decomposed as follows [8.22,8.23]:

$$\mathbf{x} = \Big[\mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_N \Big]' , \tag{3}$$

$$\mathbf{u} = \Big[\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_N \Big]' \tag{4}$$

where $\mathbf{x}_i \in R^{n_i}$, $\mathbf{u}_i \in R^{r_i}$, $i=1,2,\ldots,N$, and

$$\sum_{i=1}^{N} n_i = n, \ \sum_{i=1}^{N} r_i = r \tag{5}$$

The initial vector functions $\beta(t)$ and $\alpha(t)$ will be partitioned accordingly. The matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, $\mathbf{F}$, $\mathbf{Q}$ and $\mathbf{R}$ will also be partitioned accordingly. We denote the ith block row of matrix $\mathbf{A}$ by $\mathbf{A}_i$ and the ij block of matrix $\mathbf{A}$ by $\hat{\mathbf{A}}_{ij}$, $i\ j=1,2,\ldots,N$. Also let $\hat{\mathbf{A}}_i = \mathbf{A}_i$ with $\mathbf{A}_{ii}$ set to zero, *i.e.*,

$$\hat{\mathbf{A}}_i = \Big[\mathbf{A}_{i1}, \mathbf{A}_{i2}, \ldots, \mathbf{A}_{i,i-1}, \mathbf{0}, \mathbf{A}_{i,i+1}, \ldots, \mathbf{A}_{iN} \Big] \tag{6}$$

Similar notation applies to matrix $\hat{\mathbf{B}}_i$.

With the above notation, the ith subsystem, $i=1,2,\ldots,N$, becomes

$$\dot{\mathbf{x}}_i(t) = \mathbf{A}_{ii}(t)\mathbf{x}_i(t) + \mathbf{B}_{ii}(t)\mathbf{u}_i(t) + \mathbf{z}_i(t), \quad t \geqslant t_0 \tag{7a}$$

$$\mathbf{x}_i(t_0) = \beta_i(t_0) \tag{7b}$$

where

$$\mathbf{z}_i(t) = \hat{\mathbf{A}}_i(t)\mathbf{x}(t) + \hat{\mathbf{B}}_i(t)\mathbf{u}(t) + \mathbf{C}_i(t)\mathbf{x}\ (t-h_x) + \mathbf{D}_i(t)\mathbf{u}(t-h_u) \tag{8}$$

Also the cost functional (2) is assumed

$$J = \sum_{i=1}^{N} J_i \tag{9}$$

where

$$J_i = \tfrac{1}{2}\, \mathbf{x}'_i(t_f) \left[\mathbf{F}_{ii}\mathbf{x}_i(t_f) + 2\mathbf{f}_i(t_f) \right] + \tfrac{1}{2} \int_{t_0}^{t_f} \left\{ \mathbf{x}'_i(t) \left[\mathbf{Q}_{ii}(t)\ \mathbf{x}_i(t) + 2\mathbf{q}_i(\mathbf{x}(t)) \right] \right.$$
$$\left. + \mathbf{u}'_i(t) \left[\mathbf{R}_{ii}(t)\ \mathbf{u}_i(t) + 2\mathbf{r}_i\ (\mathbf{u}(t)) \right] \right\} dt \tag{10}$$

and

$$\mathbf{f}_i\left[\hat{\mathbf{x}}_i(\mathrm{T})\right] = \tfrac{1}{2}\, \mathbf{F}_i\hat{\mathbf{x}}_i(T),\ \mathbf{q}_i\left[\mathbf{x}_i(t)\right] = \tfrac{1}{2}\, \mathbf{Q}_i\hat{\mathbf{x}}_i(t),\ \mathbf{r}_i\left[\hat{\mathbf{u}}_i(t)\right] = \tfrac{1}{2}\, \mathbf{R}_i\hat{\mathbf{u}}_i(t) \tag{11}$$

No constraints need to be put on the method of system decomposition. However, often a natural decomposition exists for any given large-scale TD system. If this is not the case, it is partitioned in such a way so that on the average, a maximum number of zeros occur in the off-diagonal blocks of the matrices **A**, **B**, **C**, **D**, **F**, **Q** and **R** or the system would be weakly coupled within the context of nondelay systems [8.31].

In sequel, an iterative procedure is proposed to determine a near-optimum control for the problem defined by (1)-(2). In each iteration the delay terms and the terms due to interdependence between subsystems will appear as known, forcing functions obtained in the previous iteration. More precisely, in iteration k the following nondelay optimal control problem is solved:

$$\dot{\mathbf{x}}_i^k = A_{ii}\mathbf{x}_i^k + \mathbf{B}_{ii}\mathbf{u}_i^k + \mathbf{z}_i^{k-1},\ t \geqslant t_0 \tag{12a}$$

$$\mathbf{x}_i^k(t_0) = \beta_i(t_0) \tag{12b}$$

$$J_i^k = \tfrac{1}{2}\,\mathbf{x}_i^{k'}(t_f)\left[\mathbf{F}_{ii}\mathbf{x}_i^k\,(t_f) + 2\mathbf{f}_i^{k-1}\left(\mathbf{x}_i^k\,(t_f)\right)\right] + \tfrac{1}{2}\int_{t_0}^{t_f}\left\{\mathbf{x}_i^{k'}\left[\mathbf{Q}_{ii}\mathbf{x}_i^k + 2\mathbf{q}_i^{k-1}\,(\mathbf{x}^k)\right]\right.$$

$$\left. + \mathbf{u}_i^{k'}\left[\mathbf{R}_{ii}\mathbf{u}_i^k + 2\mathbf{r}_i^{k-1}\,(\mathbf{u}^k)\right]\right\}dt \tag{13}$$

where

$$\mathbf{z}_i^{k-1} = \hat{\mathbf{A}}_i\mathbf{x}^{k-1} + \hat{\mathbf{B}}_i\mathbf{u}^{k-1} + \mathbf{C}_i\mathbf{x}^{k-1}(t-h_x) + \mathbf{D}_i\mathbf{u}^{k-1}\,(t-h_u) \tag{14a}$$

and

$$\mathbf{x}_i^k(t) = \beta_i(t),\ t_0-h_x \leqslant t \leqslant t_0,\ k=1,2,\cdots \tag{14b}$$

$$\mathbf{u}_i^k(t) = \alpha_i(t),\ t_0-h_u \leqslant t \leqslant t_0,\ k=1,2,\cdots \tag{14c}$$

Note that dependence of t has been omitted whenever it does not cause ambiguity. In order to initiate the iterations, we let

$$\mathbf{x}_i^0(t) = \psi_{ii}(t,\,t_0)\beta_i(t_0),\ \mathbf{u}_i^0(t) = \mathbf{h}_i(t),\ t \geqslant t_0 \tag{15}$$

where $\psi(t,t_0)$ is the state transition matrix corresponding to system matrix $\mathbf{A}_{ii}(t) - \mathbf{B}_{ii}\mathbf{R}_{ii}^{-1}\mathbf{B}_{ii}\mathbf{K}_{ii}(t)$, $\mathbf{h}_i(t)$ is an arbitrary continuous function, and $\mathbf{K}_{ii}(t)$ is the symmetric positive-definite solution to the following DMRE:

$$\dot{\mathbf{K}}_{ii} + \mathbf{K}_{ii}\mathbf{A}_{ii} + \mathbf{A}'_{ii}\mathbf{K}_{ii} - \mathbf{K}_{ii}\mathbf{S}_{ii}\mathbf{K}_{ii} + \mathbf{Q}_{ii} = \mathbf{0},\ \mathbf{K}_{ii}(t_f) = \mathbf{F}_{ii} \tag{16}$$

where $\mathbf{S}_{ii} = \mathbf{B}_{ii}\mathbf{R}_{ii}^{-1}\mathbf{B}'_{ii}$. The following theorem justifies the above procedure.

1. Theorem. Consider the sequence of linear state equations (12) with (14) and (15) and the associated cost function (10). Suppose that for the $k$the iteration the optimal state trajectory $\mathbf{x}_i^{*k}$ and the optimal control $\mathbf{u}_i^{*k}$ are obtained. Then the sequence of n-dimensional state vectors $[\mathbf{x}_1^{*k},\mathbf{x}_2^{*k},\ldots\mathbf{x}_N^{*k}]'$ and m-dimensional control vectors $[\mathbf{u}_1^{*k},\mathbf{u}_2^{*k},\ldots,\mathbf{u}_N^{*k}]'$ uniformly converge, respectively to $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$, the optimal state trajectory and the optimal control for the optimization problem defined by (1) and (2).

The proof of this theorem, given after the following lemma is established.

2. Lemma. Consider the system of delay-differential equations

$$\dot{x}_i(t) = A_{ii}(t)\, x_i(t) + \hat{A}_i(t)\, x(t) + C_i(t)\, x(t-h_x),\ t \geqslant t_0 \tag{17a}$$

$$x_i(t) = \beta_i(t),\ t_0 - h_x \leqslant t \leqslant t_0,\ i=1,2,\ldots,N \tag{17b}$$

where $A_{ii}(t)$, $\hat{A}_i(t)$ and $C_i(t)$, $i=1,2,\ldots,N$, are real piecewise-continuous matrices defined previously; vectors x and x_i are defined in (3) and h_x is a constant real positive scalar. Then the sequence of vector functions $\{x_i^k(t)\}$, $i=1,2,\ldots,N$, defined by

$$x_i^0(t) = \phi_{ii}(t,t_0)\beta_i(t_0).\ t_0 \leqslant t \leqslant t_f \tag{18a}$$

$$x_i^k(t) = \phi_{ii}(t,t_0)\beta_i(t_0) + \int_{t_0}^{t_f} \phi_{ii}\ (t,\tau) \Big[\ \hat{A}_i(\tau)\ x^{k-1}$$

$$+ C_i\ (\tau)x^{k-1}(\tau-h_x)\ ]d\tau,\ t_0 \leqslant t \leqslant t_f,\ k=1,2,\cdots \tag{18b}$$

$$x_i^k(t) = \beta_i(t),\ \ t_0-h_x \leqslant t \leqslant t_0,\ \ k=1,2,\cdots \tag{18c}$$

where $\phi_{ii}(t,\tau)$ denotes the state transition matrix corresponding to the system matrix $A_{ii}(t)$, and $\hat{A}_i(t)$, defined in (5), converges uniformly to the solution of (17).

Proof. Equations (17) for $i=1,2,..,N$ imply

$$\dot{x}(t) = A_d(t)x(t) + \hat{A}(t)x(t) + C(t)x(t-h_x),\ t_0 \leqslant t \leqslant t_f \tag{19a}$$

where $A_d(t) = diag(A_{ii}(t)\ )$ and $\hat{A}(t) = A(t)$, where the block diagonals and $A_{ii}(t)$ are zero matrices. Also, equations (17b) for $i=1,2,\ldots,N$ imply

$$x(t) = \beta(t),\ \ t_0 - h_x \leqslant t \leqslant t_0 \tag{19b}$$

In the same manner, using equations (18), the sequence

$$x^k(t) = \Big[\ x_1^{k\prime}(t),\ x_2^{k\prime}(t),\ \ldots,\ x_N^{k\prime}(t)\ \Big]' \tag{20}$$

will be defined by

$$\mathbf{x}^0(t) = \phi_d(t,t_0)\beta(t_0),\ t \geqslant t_0 \tag{21a}$$

$$\mathbf{x}^k(t) = \phi_d(t,t_0)\beta(t_0) + \int_{t_0}^{t} \phi_d(t,\tau)\Big[\ \hat{\mathbf{A}}\ (\tau)\ \mathbf{x}^{k-1}(\tau) + \mathbf{C}\ (\tau)\ \mathbf{x}^{k-1}(\tau-h)\ \Big]dt,$$

$$k=1,2,\ldots,\ t \geqslant t_0 \tag{21b}$$

$$\mathbf{x}^k(t) = \beta(t),\quad t_0-h_x \leqslant t \leqslant t_0,\ k=1,2,\ldots,\ \ t \geqslant t_0 \tag{21c}$$

where

$$\phi_d(t,\tau) = diag\Big[\ \phi_{ii}(t,\tau)\ \Big] \tag{21d}$$

It will be shown that the sequence of vector functions $\{\mathbf{x}^k(t)\}$ defined above uniformly converges to the solution of (19). Equations (21b) yields

$$\mathbf{x}^2(t) - \mathbf{x}^1(t) = \int_{t_0}^{t} \phi_d(t,\tau) \Big\{\ \hat{\mathbf{A}}(\tau)\ \Big[\ \mathbf{x}^1(\tau) - \mathbf{x}^0(\tau)\ \Big]$$

$$+\ \mathbf{C}\ (\tau)\ \Big[\ \mathbf{x}^1(\tau-h_x) - \mathbf{x}^0(t-h_x)\ \Big]\ \Big\}\tau\ dt \tag{22a}$$

which implies

$$||\mathbf{x}^2(t) - \mathbf{x}^1(t)\ || \leqslant M \int_{t_0}^{t} \Big\{\ N_\alpha\ ||\ \mathbf{x}^1(\tau) - \mathbf{x}^0(\tau)\ ||$$

$$+\ N_c\ ||\ \mathbf{x}^1(\tau-h_x) - \mathbf{x}^0(\tau-h_x)\ ||\ \Big\}\ d\tau \tag{22b}$$

where

$$M = \operatorname*{Sup}_{t,\tau\ \in\ [t_0 t_f]} ||\ \phi_d(t,\tau)\ ||,\ N_\alpha = \operatorname*{Sup}_{t,\tau\ \in\ [t_0 t_f]} ||\ \hat{\mathbf{A}}(\tau)\ ||,\ N_c = \operatorname*{Sup}_{t,\tau\ \in\ [t_0 t_f]} ||\ \mathbf{C}(\tau)\ || \tag{22c}$$

It is convenient to choose a norm $||\ .||$ such that $||\ \phi_d(t_0,t_0)\ || = ||\ \mathbf{I}\ || = 1$. This then guarantees that $M \geqslant 1$. Let

$$L = \sup_{t \in [t_0-h_x, t_0]} \| \beta(t) \| \tag{23}$$

Then from (21) we have

$$\begin{aligned}
\|x^1(t) - x^0(t)\| &\leq M \left[N_a \int_{t_0}^{t} \|x^0(\tau)\| \, d\tau + N_c \int_{t_0}^{t} \|x^0(\tau-h_x)\| \, d\tau \right] \\
&= M \Bigg\{ N_a \int_{t_0}^{t} \|\phi_d(\tau,t_0) \, \beta(t_0)\| \, d\tau \\
&\quad + N_c \left[\int_{t_0}^{t_0+h_x} \| \beta(\tau-h_x) \| \, d\tau + \int_{t_0+h_x} \| \phi(\tau-h_x, t_0) \, \beta(t_0) \| \, d\tau \right] \Bigg\} \\
&\leq M \left\{ N_a \left[ML(t-t_0) \right] + N_c \left[Lh_x + ML(t-t_0-h_x) \right] \right\} \\
&\leq M^2 L \, (N_a+N_c) \, (t-t_0), \quad t_0 \leq t \leq t_f
\end{aligned} \tag{24}$$

Using (24), (22b) implies that

$$\|x^2(t) - x^1(t)\| \leq M^3 L (N_a+N_c)^2 (t-t_0)^2/2!, \quad t_0 \leq t \leq t_f \tag{25}$$

and, by induction,

$$\|x^k(t) - x^{k-1}(t)\| \leq M^{k+1} L (N_a+N_c)^k (t-t_0)^k/k!, \quad t_0 \leq t \leq t_f \tag{26}$$

Applying the triangle inequality one has, for any r,

$$||\mathbf{x}^{k+r}(t)-\mathbf{x}^k(t)|| \leqslant \sum_{i=k+1}^{k+r} \left[M^{i+1} L (N_a+N_c)^i \frac{(t-t_0)^i}{i!} \right.$$

$$\leqslant \frac{ML\left[M(N_a+N_c)(t-t_0)\right]^{k+1}}{(k+1)!} \times$$

$$\left\{1 + \frac{M(N_a+N_c)(t-t_0)}{1!} + \left[M(N_a+N_c)(t-t_0)\right]^2 + \cdots \right\}$$

$$\leqslant \frac{ML\left[M(N_a+N_c)(t-t_0)\right]^{k+1}}{(k+1)!} \; exp\left[M(N_a+N_c)(t-t_0)\right],$$

$$t_0 \leqslant t \leqslant t_f \tag{27}$$

Therefore, the sequence $\{\mathbf{x}^k(t)\}$, $k=0,1,2,\cdots$ is a Cauchy Sequence in $C^n[t_0-h_x,t_f]$, and it follows that this sequence converges [8.39]. The limit of this sequence is the solution to (19), and the lemma is proved.

Proof of Theorem 1. Consider the ith subproblem defined by (12)-(13). Note that $\mathbf{z}_i^{k-1}$ is known in the kth iteration and acts as an extra perturbing input in the kth state equation. The Hamiltonian function for this problem is

$$H_i^k = \tfrac{1}{2}\, \mathbf{x}'_i \left[\mathbf{Q}_{ii}\mathbf{x}_i^k + 2\mathbf{q}_i^{k-1}\right] + \tfrac{1}{2}\, \mathbf{u}'_i \left[\mathbf{R}_{ii}\mathbf{u}_i^k + 2\mathbf{r}_i^{k-1}\right]$$
$$+ \mathbf{p}_i^{k'}\left[\mathbf{A}_{ii}\mathbf{x}_i^k + \mathbf{B}_{ii}\mathbf{u}_i^k + \mathbf{z}_i^{k-1}\right]$$

where $\mathbf{p}_i^k(t)$ is the corresponding costate vector and $\mathbf{q}_i^k$ and $\mathbf{r}_i^k$ are defined in (13). The necessary and sufficient conditions for optimality are

$$\dot{\mathbf{p}}_i^k = -\frac{\partial H_i^k}{\partial \mathbf{x}_i^k} \; -\mathbf{Q}_{ii}\mathbf{x}_i^k - \mathbf{q}_i^{k-1} - \mathbf{A}'_{ii}\mathbf{p}_i^k,\; t \geqslant t_0 \tag{28a}$$

$$\mathbf{p}_i^k(t_f) = \frac{\partial}{\partial \mathbf{x}_i^k(t_f)} \left\{\tfrac{1}{2}\, \mathbf{x}_i^{k'}(t_f)\left[\mathbf{F}_{ii}\mathbf{x}_i^k(t_f) + 2\mathbf{f}_i(t_f)\right\}\right. = \mathbf{F}_{ii}\mathbf{x}_i^k(t_f) + \mathbf{f}_i(t_f) \tag{28b}$$

$$0 = \frac{\partial H_i^k}{\partial \mathbf{u}_i^k} = \mathbf{R}_{ii}\mathbf{u}_i^k + \mathbf{r}_i^{k-1} + \mathbf{B}'_{ii}\mathbf{p}_i^k,\; t \geqslant t_0 \tag{28c}$$

Equations (12a) and (28a) can be decoupled by defining the adjoint vectors $\mathbf{g}_i^k(t)$, $k=1,2,...$ as follows:

$$\mathbf{p}_i^k = \mathbf{K}_{ii}\mathbf{x}_i^k + \mathbf{g}_i^k, \quad t_0 \leqslant t \leqslant t_f \tag{29}$$

where $\mathbf{K}_{ii}$ is the symmetric positive-definite solution to the DMRE (16). Equations (28) and (29) imply that

$$\dot{\mathbf{g}}_i^k = -(\mathbf{A}_{ii} - \mathbf{S}_{ii}\mathbf{K}_{ii})'\mathbf{g}_i^k - \mathbf{K}_{ii}\gamma_i^{k-1} - \mathbf{q}_i^{k-1} \tag{30a}$$

where

$$\gamma_i^{k-1} = -\mathbf{D}_{ii}\mathbf{R}_{ii}^{-1}\mathbf{r}_i^{k-1} + \mathbf{z}_i^{k-1} \tag{31}$$

The boundary condition for Equation (30a) can be determined by comparison of (29) for $t = t_f$ with (28b) which yields

$$\mathbf{g}_i^k(t_f) = \mathbf{f}_i(t_f), \quad \mathbf{K}_{ii}(t_f) = \mathbf{F}_{ii} \tag{30b}$$

Note that the second condition of (30b) conforms with that of (16). From (28c) and (29), the optimal control for the kth optimization problem can be written as

$$\mathbf{u}_i^{*k} = -\mathbf{R}_{ii}^{-1}(\mathbf{B}'_{ii}\mathbf{p}_i^k + \mathbf{r}_i^{k-1}) = -\mathbf{R}_{ii}^{-1}\mathbf{B}'_{ii}\mathbf{K}_{ii}\mathbf{x}_i^{*k} - \mathbf{R}_{ii}^{-1}\mathbf{B}'_{ii}\mathbf{g}_i^k - \mathbf{R}_{ii}^{-1}\mathbf{r}_i^{k-1} \tag{32}$$

Hence, from (12a) and (32), the optimal state trajectory $\mathbf{x}_i^{*k}(t)$ is the solution to

$$\dot{\mathbf{x}}_i^{*k} = \left[\mathbf{A}_{ii} - \mathbf{S}_{ii}\mathbf{K}_{ii}\right]\mathbf{x}_i^{*k} - \mathbf{S}_{ii}\mathbf{g}_i^k + \gamma_i^{k-1} \quad t_0 \leqslant t \leqslant t_f \tag{33}$$

From (30a), note that $g_i^k(t)$ depends on known functions and $\mathbf{z}_i^{k-1}$. The solution to (33) with boundary condition (12b) is

$$\mathbf{x}_i^{*k}(t) = \psi_{ii}(t,t_0)\beta_i(t_0) + \int_{t_0}^{t} \psi_{ii}(t,\tau)\left[-\mathbf{S}_{ii}(\tau)\mathbf{g}_i^k(\tau) + \gamma_i^{k-1}(\tau)\right]d\tau,$$

$$t_0 \leqslant t \leqslant t_f, \; k=1,2,\cdots \tag{34}$$

Comparison of (15) with (18a) and (34) with (18b) shows that the sequence $\{\mathbf{x}_i^{*k}(t)\}$ converges uniformly. Also, the sequences $\{\mathbf{g}_i^k(t)\}$ and $\{\mathbf{u}_i^{*k}(t)\}$ converge because, from (30a) and (32), these sequences are related to $\{\mathbf{x}_i^{*k}(t)\}$ by continuous transformations. From Lemma 2, the limit of the sequence $\{\mathbf{x}_i^{*k}(t)\}$ is the solution to

$$\dot{\mathbf{x}}_i^* = \left[\mathbf{A}_{ii} - \mathbf{S}_{ii}\mathbf{K}_{ii}\right] \mathbf{x}_i^* - \mathbf{S}_{ii}\mathbf{g}_i - \mathbf{B}_{ii}\mathbf{R}_{ii}^{-1}\mathbf{r}_i$$
$$+ \mathbf{A}_i\mathbf{x}^*(t) + \mathbf{B}_i\mathbf{u}^* + \mathbf{C}_i\mathbf{x}^*(t-h_x) + \mathbf{D}_i\mathbf{u}^*(t-h_u), t_0 \leqslant t \leqslant t_f \tag{35a}$$

$$\mathbf{x}_i^*(t) = \beta_i(t), \ t_0 - h_x \leqslant t \leqslant t_0 \tag{35b}$$
$$\mathbf{u}_i^*() = \alpha_i(t), \ t_0 - h_u \leqslant t \leqslant t_0$$

where $\mathbf{x}_i^*(t)$, $\mathbf{u}_i^*(t)$, and $\mathbf{g}_i(t)$ are, respectively, the limits of the sequences $\{\mathbf{x}'^{*k}_i(t)\}$, $\{\mathbf{u}_i^{*k}(t)\}$ and $\{\mathbf{g}_i^k(t)\}$. From (32)

$$\mathbf{u}_i^* = -\mathbf{R}_{ii}^{-1}\mathbf{B}_{ii}'\mathbf{K}_{ii}\mathbf{x}_i^* - \mathbf{R}_{ii}^{-1}\mathbf{B}_{ii}'\mathbf{g}_i - \mathbf{R}_{ii}^{-1}\mathbf{r}_i \tag{36}$$

substituting $\mathbf{u}_i^*(t)$ from (36) for $\mathbf{u}_i(t)$ in (6) and comparing the result with (35a) shows that $\mathbf{x}_i^*(t)$ and $\mathbf{u}_i^*(t)$ are, respectively, the optimal state trajectory and the optimal control for the ith optimization subproblem defined by (6)-(7) and (10) for $i=1,2,\ldots,N$. Or equivalently,

$$\mathbf{x}^*(t) = \left\{ \mathbf{x}'_1{}^*(t), \ldots, \mathbf{x}'_N{}^*(t) \right\} \tag{37a}$$

and

$$\mathbf{u}^*(t) = \left\{ \mathbf{u}'_1{}^*(t), \ldots, \mathbf{u}'_N{}^*(t) \right\} \tag{37b}$$

are, respectively, the optimal state trajectory and the optimal control for the optimization problem defined by (1) and (2). Thus the theorem is proved.

The following algorithm summarizes the hierarchical control of a large-scale TD system.

4. Algorithm.

Step 1: Solve (16) to obtain $\mathbf{K}_{ii}(t)$ and store.

Step 2: Determine $\mathbf{x}_i^0(t)$ and $\mathbf{u}_i^0(t)$ from (15). Set $i=1$, $k=1$.

Step 3: Use $\mathbf{x}_i^{k-1}$, $\mathbf{u}_i^{k-1}$ and (14b), (14c) in (14a) to obtain $\mathbf{z}_i^k(t)$, $t \geqslant t_0$.

Step 4: Solve the ith optimal control subproblem to find $\mathbf{x}_i^{*k}(t)$ and $\mathbf{u}_i^{*k}(t)$, $t \geqslant t_0$, as described by (35) and (36).

Step 5: If $i = i + 1 < N$, go to Step 3.

Step 6: Form $\mathbf{x}^k(t)$ and $\mathbf{u}^k(t)$ as in (3)-(4). If $\max_{t_0 \leqslant t \leqslant t_f} \{ \| \mathbf{x}^k(t) - \mathbf{x}^{k-1}(t) \| \} \leqslant \epsilon$ and $\max_{t_0 \leqslant t \leqslant t_f} \{ \| \mathbf{u}^k(t) - \mathbf{u}^{k-1}(t) \| \} \leqslant \epsilon$, for all $t_0 \leqslant t \leqslant t_f$, where ϵ is a prespecified small constant, stop. Otherwise, set $k = k + 1$ and go to Step 3.

5. Example. To illustrate the above algorithm consider the following linear system,

$$\dot{\mathbf{x}} = \left[\begin{array}{cc|cc} -3 & 2 & 0 & 0 \\ 0 & -4 & 2 & 0 \\ \hline -1 & -1 & -3 & 0 \\ 0 & 1 & 1 & -2 \end{array}\right] \mathbf{x} + \left[\begin{array}{ccccc} -0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & 0 \end{array}\right] \mathbf{x}(t-0.1) +$$

$$\left[\begin{array}{cc|cc} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right] \mathbf{u} + \left[\begin{array}{cccc} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \mathbf{u}(t-0.1) \tag{38}$$

with $\mathbf{x}(t) = 0$. The objective function is assumed to be quadratic, *i.e.*,

$$J = \tfrac{1}{2}\, \mathbf{x}'(2)\, \mathbf{S}\mathbf{x}(2) + \tfrac{1}{2} \int_0^2 (\mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})\, dt \tag{39}$$

where

$$\mathbf{S} = \left[\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \tag{40}$$

$\mathbf{R} = diag\ (1\ 1\ 1\ 1)$ and $\mathbf{Q} = diag\ (1\ 2\ 1\ 2)$.

The system is decomposed into two second-order subsystems and the steps outlined in Algorithm 4 were followed. At Step 1, two differential matrix Riccati equations were solved using a 4th order Runge-Kutta integration method and their solutions were fit in a 4th order polynomials as given below.

$K_1(t) =$

$$\begin{bmatrix} 0.02 - 0.69t + 2.21t^2 - 2.23t^3 + 0.72t^4 & 0.04 - 1.27t + 4.06t^2 - 4.15t^3 + 1.34t^4 \\ 0.04 - 1.27t + 4.06t^2 - 4.15t^3 + 1.34t^4 & 0.07 - 0.35t + 1.21t^2 - 1.31t^3 + 0.46t^4 \end{bmatrix}$$

$K_2(t) =$

$$\begin{bmatrix} 0.005 - 1.42t^2 - 4.53t^3 + 1.44t^4 & 0.02 - 0.69t + 2.19t^2 - 2.21t^3 + 0.71t^4 \\ 0.02 - 0.69t + 2.19t^2 - 2.21t^3 + 0.71t^4 & 0.02 - 0.61t + 1.94t^2 - 1.98t^3 + 0.64t^4 \end{bmatrix} \tag{41}$$

The step size was chosen 0.1 to conveniently recover the delayed quantities defined in equation (38).

The interaction error was reduced to 2.10778×10^{-8} in 12 iterations as shown in Fig. 1. Figures 2 through 5 show the optimal states and inputs of the system.

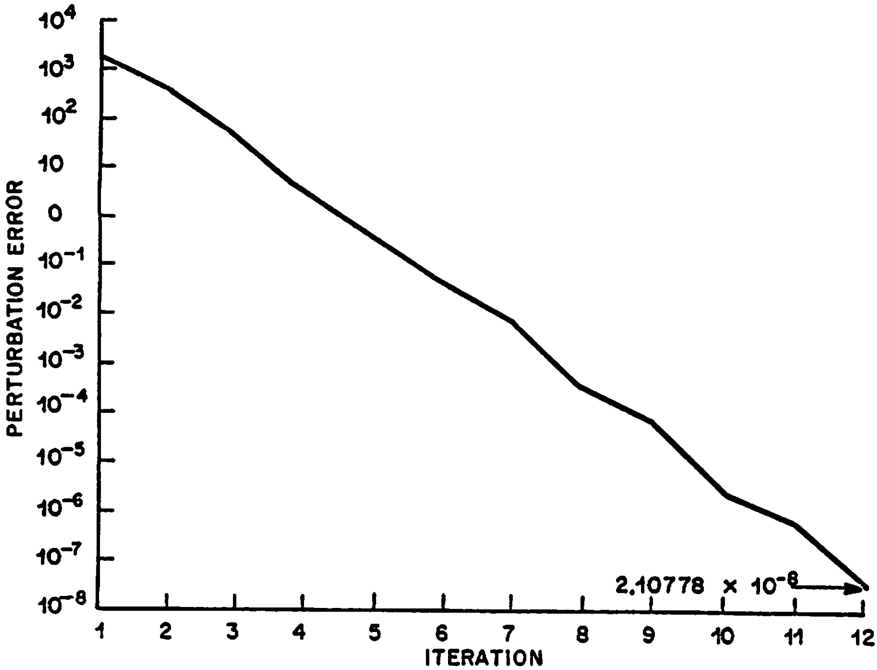

Figure 1. Perturbation Error versus Iteration for Example 5

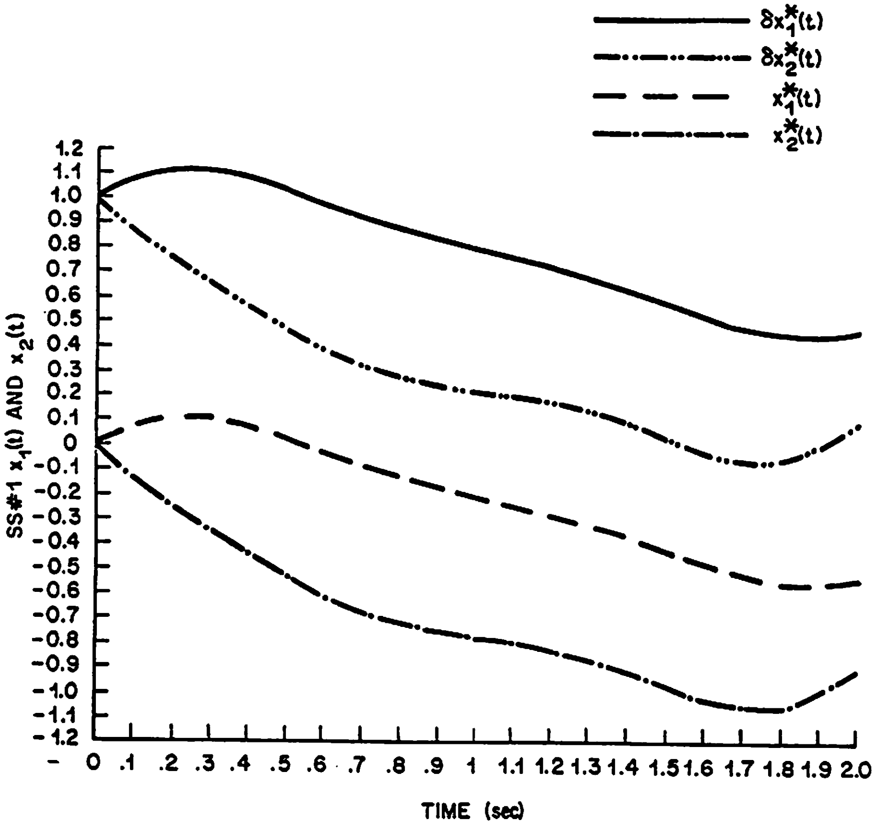

Figure 2. Optimal State Trajectories for Subsystem No. 1 of Example 5

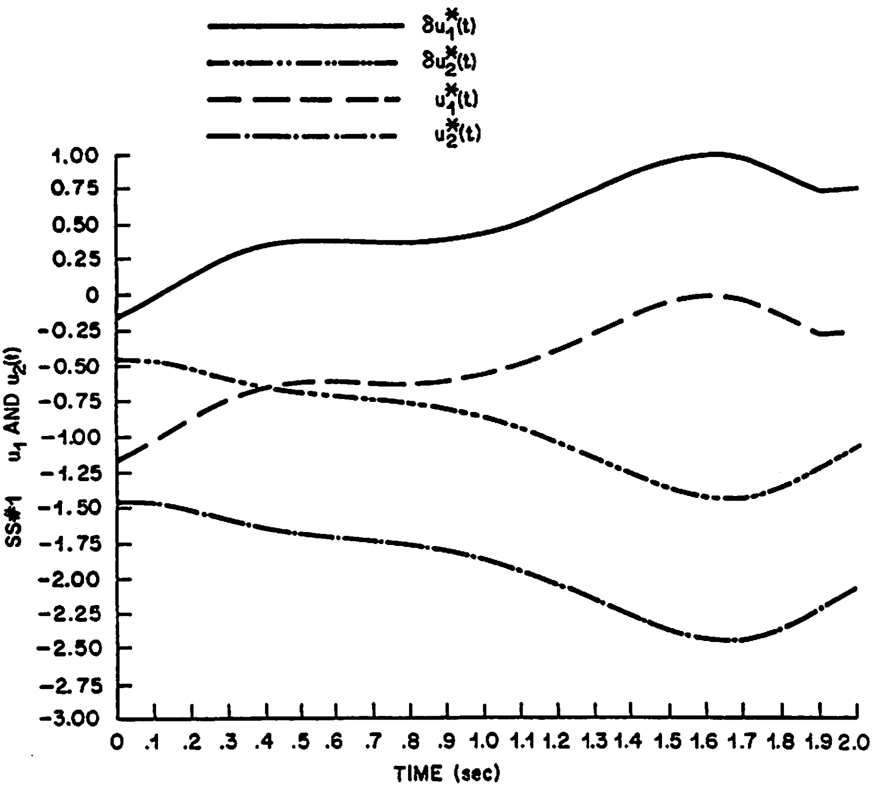

Figure 3. Optimal Control Trajectories for Subsystem No. 1 of Example 5

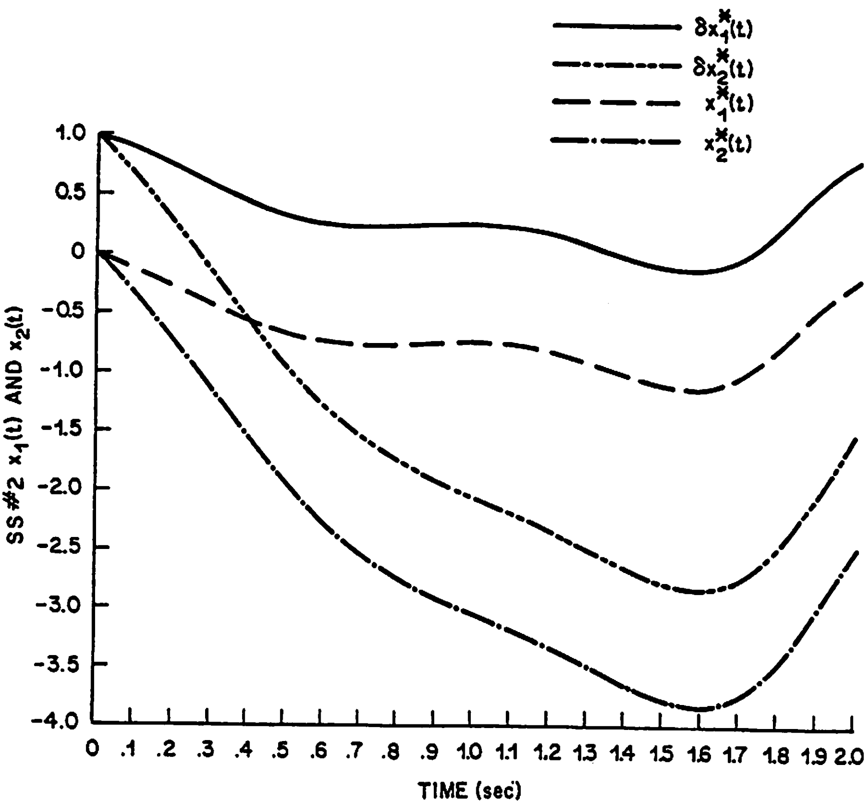

Figure 4. Optimal State Trajectories for Subsystem No. 2 of Example 5

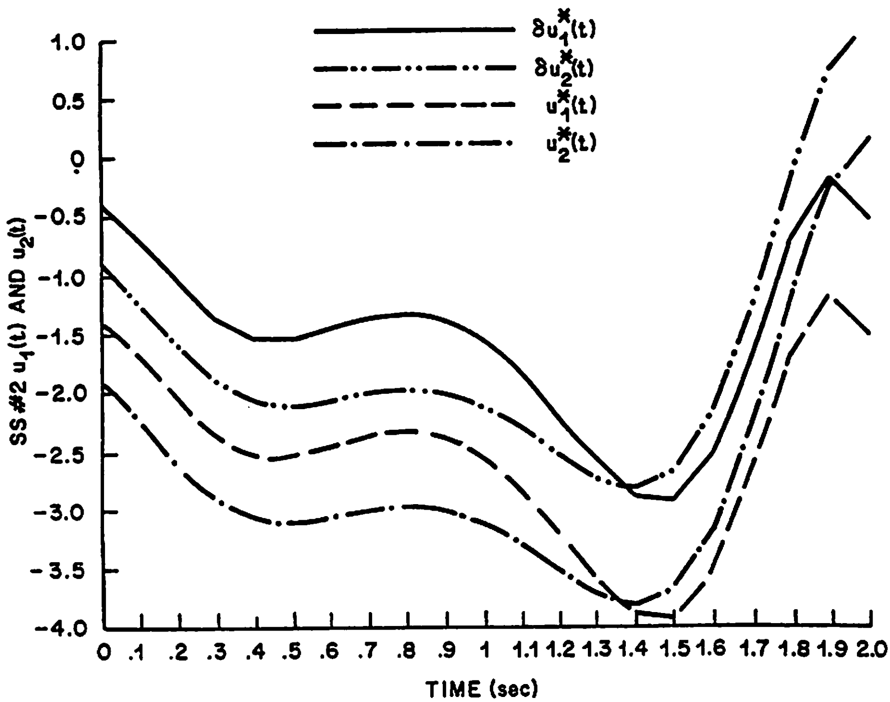

Figure 5. Optimal Control Trajectories for Subsystem No. 2 of Example 5

8.6 HIERARCHICAL CONTROL OF TIME-DELAY SYSTEMS VIA INTERACTION PREDICTION

In this Section the "interaction prediction" method of hierarchical control of large-scale systems, described in Section 8.3.2, is extended to nonlinear TD systems with multiple delays in both the state and the control. A fast converging iterative algorithm is developed to obtain a near-optimum control for such systems. A Taylor series expansion is used to linearize the original system whose near-optimum control results from the optimization of a sequence of nondelay linear time-varying subsystems problems. As in some of the earlier efforts [8.20,8.22,8.23,8.31], the interaction terms among subsystems involve delay quantities. The coordination of subproblems is identical to that of the nondelayed interaction predication. The scheme has been applied to a number of numerical examples one of which is illustrated here.

Consider a nonlinear large-scale TD system

$$\dot{\mathbf{x}} = \mathbf{f}\Big[\mathbf{x}(t),\ \mathbf{x}(t-h_{x1})\ ,\ldots,\ \mathbf{x}(t-h_{xp}),\mathbf{u}(t),\ \mathbf{u}(t-h_{u1})\ ,\ldots,\ \mathbf{u}(t-h_{uq})\Big] \tag{1a}$$

$$t \geqslant t_0,\ h_{xp} \geqslant \cdots \geqslant h_{xi} \geqslant \cdots \geqslant h_{x1}$$
$$h_{uq} \geqslant \cdots \geqslant h_{ui} \geqslant \cdots \geqslant h_{u1}$$

$$\mathbf{x}(t) = \mathbf{x}_0(t),\ t_0 - h_{xp} \leqslant t \leqslant t_0 \tag{1b}$$

$$\mathbf{u}(t) = \mathbf{u}_0(t),\ t_0 - h_{uq} \leqslant t \leqslant t_0 \tag{1c}$$

where $\mathbf{x}(t)\epsilon R^n$ and $\mathbf{u}(t)\epsilon R^m$ are, respectively, state and control vectors, h_{xi}, $i=1,2,\ldots,p$ and h_{ui}, $j=1,2,\ldots,q$ are constant positive scalars representing delays, $f(.)$ is assumed to be a continuously differentiable function of its arguments; t_0 is the initial process time; $\mathbf{x}_0(t)$ and $\mathbf{u}_0(t)$ are specified initial functions. The cost functional, to be minimized, is

$$J = \tfrac{1}{2}\ \mathbf{x}'(t_f)\mathbf{S}\mathbf{x}(t_f) + \tfrac{1}{2}\int_{t_0}^{t_f} \Big[\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t)\Big]dt \tag{2}$$

where without any loss of generality, the matrices $\mathbf{S}$ and $\mathbf{Q}(t)$ are assumed to be diagonal and positive semi-definite; and piecewise continuous, and the matrix $\mathbf{R}(t)$ is diagonal positive definite and piecewise-continuous. The problem is to find a control $\mathbf{u}(t)$, $t_0 \leqslant t \leqslant t_f$ which for fixed final time t_f and free final state $\mathbf{x}(t_f)$ satisfies (1) and approximately minimizes the cost functional J in (2).

8.6.1 System Linearization and Decomposition

For the nonlinear system (1), a linear approximation can be obtained through a Taylor's series expansion about a nominal set of trajectories $(\mathbf{x}_n,\mathbf{u}_n)$. It is assumed that the nominal trajectories satisfy (1), *i.e.*,

$$\dot{\mathbf{x}}_n = \mathbf{f}\Big(\mathbf{x}_n,\mathbf{x}_n(t-h_{x1}) \;,\ldots,\; \mathbf{x}_n(t-h_{xp}),\, \mathbf{u}_n,\mathbf{u}_n(t-h_{u1}) \;,\ldots,\; \mathbf{u}_n(t-h_{uq})\Big) \triangleq \mathbf{f}_n \tag{3}$$

Let $(\delta\mathbf{x},\delta\mathbf{u})$ be the perturbation vectors between $(\mathbf{x},\mathbf{u})$ and nominal trajectories $(\mathbf{x}_n,\mathbf{u}_n)$, *i.e.*,

$$\begin{aligned} \delta\mathbf{x} &= \mathbf{x} - \mathbf{x}_n \\ \delta\mathbf{u} &= \mathbf{u} - \mathbf{u}_n \end{aligned} \tag{4}$$

Expanding (1) about the nominal trajectories will be

$$\begin{aligned} \delta\dot{\mathbf{x}}(t) = \mathbf{f}_n &+ \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n}\right]'\delta\mathbf{x}(t) + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n(t-h_{1x})}\right]'\delta\mathbf{x}(t-h_{1x}) \\ &+ \cdots + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n(t-h_{px})}\right]'\delta\mathbf{x}(t-h_{px}) + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{u}_n}\right]'\delta\mathbf{u} \\ &+ \left[\frac{\partial \mathbf{f}}{\partial \mathbf{u}_n(t-h_{1u})}\right]'\delta\mathbf{u}(t-h_{1u}) + \cdots \\ &+ \left[\frac{\partial \mathbf{f}}{\partial \mathbf{u}_n(t-h_{qu})}\right]'\delta\mathbf{u}(t-h_{qu}) + \text{H.O.T.} \end{aligned} \tag{5}$$

Considering the definition of $\mathbf{f}_n$ in (3), (5) can be rewritten as

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}_0(t)\delta\mathbf{x}(t) + \sum_{k=1}^{p} \mathbf{A}_k\delta\mathbf{x}(t-h_{kx}) + \mathbf{B}_0(t)\delta\mathbf{u}(t) + \sum_{\ell=1}^{q} \mathbf{B}_\ell\delta\mathbf{u}(t-h_{\ell u}) \tag{6a}$$

with initial functions

$$\delta\mathbf{x}(t) = \mathbf{x}_0(t) - \mathbf{x}_n(t) = \delta\mathbf{x}_0(t),\; t_0 - h_{xp} \leqslant t \leqslant t_0 \tag{6b}$$

$$\delta\mathbf{u}(t) = \mathbf{u}_0(t) - \mathbf{u}_n(t) = \delta\mathbf{u}_0(t),\; t_0 - h_{uq} \leqslant t \leqslant t_0 \tag{6c}$$

where

$$\mathbf{A}_0 \triangleq \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n} \right]' \quad \mathbf{A}_k \triangleq \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n(t-h_{kx})} \right]', \; k=1,2,\ldots,p \tag{6d}$$

$$\mathbf{B}_0 \triangleq \left[\frac{\partial \mathbf{f}}{\partial \mathbf{u}_n} \right]' \quad \mathbf{B}_\ell = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}_n(t-h_{u\ell})} \right]', \; \ell=1,2,\ldots,q \tag{6e}$$

and the associated cost functional becomes:

$$\hat{J} = \tfrac{1}{2}\, \delta\mathbf{x}'(t_f) \left[\mathbf{S}\delta\mathbf{x}(t_f) + 2\mathbf{S}\mathbf{x}_n(t_f) \right] + \tfrac{1}{2} \int_{t_0}^{t_f} \left[\delta\mathbf{x}'(t) \left[\mathbf{Q}(t)\delta\mathbf{x}(t) + 2\mathbf{Q}(t)\mathbf{x}_n(t) \right] \right.$$

$$\left. + \, \delta\mathbf{u}'(t) \left[\mathbf{R}(t)\delta\mathbf{u}(t) + 2\mathbf{R}(t)\mathbf{u}_n(t) \right] \right] dt \tag{7}$$

where the effect of the nominal trajectories $(\mathbf{x}_n,\mathbf{u}_n)$ on the original cost function which does not influence the optimization process with respect to $(\mathbf{x},\mathbf{u})$ has been ignored.

We start with the decomposition of the system (6) into N subsystems by partitioning the state vector $\mathbf{x}$ and the control vector $\mathbf{u}$ below:

$$\delta\dot{\mathbf{x}}_1(t) = \mathbf{A}_{0i}(t)\delta\mathbf{x}_i(t) + \mathbf{B}_{0i}(t)\delta\mathbf{u}_i(t) + \mathbf{z}_i(t), \quad t \geqslant t_0 \tag{8a}$$

$$\delta\mathbf{x}_i(t) = \delta\mathbf{x}_{0i}(t), \quad h_{xp} \leqslant t \leqslant t_0 \tag{8b}$$

$$\delta\mathbf{u}_i(t) = \delta\mathbf{u}_{0i}(t), \quad h_{uq} \leqslant t \leqslant t_0 \tag{8c}$$

$$\mathbf{z}_i(t) = \hat{\mathbf{A}}_{0i}(t)\delta\mathbf{x}(t) + \hat{\mathbf{B}}_{0i}(t)\delta\mathbf{u}(t) + \sum_{k=1}^{p} \mathbf{A}_{ki}(t)\delta\mathbf{x}(t-h_k) + \sum_{\ell=1}^{q} \mathbf{B}_{\ell i}(t)\delta\mathbf{u}(t-h_\ell)$$

$$= \mathbf{z}_{ie}(t) \tag{9}$$

where $\mathbf{A}_{0i}$ is the diagonal block of matrix $\mathbf{A}_0$, $\hat{\mathbf{A}}_{0i}$ is the ith block row of matrix $\mathbf{A}_0$ with $\mathbf{A}_{0i}$ set to zero, *i.e.*,

$$\hat{\mathbf{A}}_{0i} = \textit{block-diag} \left[\mathbf{A}_{i1},\mathbf{A}_{i2},\ldots,\mathbf{A}_{i,i-1},\mathbf{0},\mathbf{A}_{i,i+1},\ldots,\mathbf{A}_{iN} \right] \tag{10}$$

A similar notation applies to matrices $\hat{\mathbf{B}}_{0i}$, $\hat{\mathbf{B}}_{0i}$; $\hat{\mathbf{A}}_{ki}$ is the ith block row of matrix $\mathbf{A}_k$, and the same as $\mathbf{B}_{\ell i}$ to $\mathbf{B}_\ell$ for $i=1,2,\ldots,N$, $k=1,2,\ldots,p$, $\ell=1,2,\ldots,q$ and $\delta\mathbf{x}_i \in R^{ni}$,$\delta\mathbf{u}_i \epsilon R^{mi}$, $i=1,2,\ldots,N$ so that

$$\sum_{i=1}^{N} n_i = n, \quad \sum_{i=1}^{N} m_i = m \tag{11}$$

In order to guarantee a satisfactory performance of the system despite the on-off participation of the subsystem, we also assume the cost functional (7), to be represented by

$$\min_{u} \hat{J} = \sum_{i=1}^{N} \min_{u_i} \hat{J}_i \tag{12}$$

where

$$\hat{J}_i = \tfrac{1}{2}\, \delta x'_i(t_f) \left[S_i \delta x_i(t_f) + 2S_i x_{ni}(t_f) \right] + \tfrac{1}{2} \int_{t_0}^{t_f} \left[\delta x'_i(t) \left(Q_i(t)\delta x_i(t) + 2Q_i(t) x_{ni}(t) \right) + \delta u'_i(t) \left(R_i(t)\delta u_i(t) + 2R_i(t) u_{ni}(t) \right) \right] dt \tag{13}$$

However, a natural decomposition often exists for any given large-scale TD system. If this is not the case, we can partition the system such that, on the average, a maximum number of zeroes occur in the off diagonal blocks of the matrices A_0 and B_0 or the system would be weakly-coupled withing the context of nondelay systems [8.31].

8.6.2 Method of Solution

We propose an iterative procedure to determine a suboptimal control for the problem defined by (6) and (7). In each iteration the delay terms and the interacting terms will act as components of the coordination vector, *i.e.*, inputs obtained in the previous iteration.

First, let us introduce a set of Lagrange multipliers $\alpha_i(t)$ and costate (adjoint) vectors $\delta p_i(t)$ to augment the "perturbation" equality constraint (9) and subsystem dynamic constraint (8) to the cost function's integrand, therefore the ith subsystem Hamiltonian in the rth iteration is defined by

$$\begin{aligned} H_i^r = {} & \tfrac{1}{2}\, \delta x'^r(t) \left[Q_i \delta x_i^r(t) + 2Q_i x_{ni}(t) \right] + \tfrac{1}{2}\, \delta u'^r_i(t) \left[R_i \delta u_i^r(t) + 2R_i u_{ni}(t) \right] \\ & + p'^r_i(t) \left[A_{0i} \delta x_i^r(t) + B_{0i} \delta u_i^r(t) + z_i^{r-1}(t) \right] \\ & + \alpha'^r_i(t) \left[z_i^{r-1}(t) - z_{ie}^{r-1}(t) \right] \end{aligned} \tag{14}$$

The necessary conditions for optimality are

$$\delta\dot{\mathbf{p}}_i^r = \frac{\partial H_i^r}{\partial \delta \mathbf{x}_i^r} = -\mathbf{Q}_i\delta\mathbf{x}_i^r - \mathbf{Q}_i\mathbf{x}_{ni} - \mathbf{A}'_{0i}\delta\mathbf{p}_i^r \quad t \geqslant t_0 \tag{15a}$$

$$\delta\mathbf{p}_i^r(t_f) = \mathbf{S}_i\delta\mathbf{x}_i^r(t_f) + \mathbf{S}_i\mathbf{x}_{ni}(t_f) \tag{15b}$$

$$0 = \frac{\partial H_i^r}{\partial \delta \mathbf{u}_i^r} = \mathbf{R}_i\delta\mathbf{u}_i^r + \mathbf{R}_i\mathbf{u}_{ni} + B_{0i}\delta\mathbf{p}_i^r \tag{15c}$$

Equations (8a) and (15a) can be decoupled by defining the adjoint vectors $\mathbf{g}_i^r(t)$, $r=1,2,\cdots$ as follows:

$$\delta\mathbf{p}_i^r = \mathbf{K}_i\delta\mathbf{x}_i^r + \mathbf{g}_i^r \tag{16}$$

where $\mathbf{K}_i$ is the symmetric positive-definite solutions to the DMRE as below:

$$\dot{\mathbf{K}}_i = -\mathbf{K}_i\mathbf{A}_{0i} - \mathbf{A}'_{0i}\mathbf{K}_i + \mathbf{K}_i\mathbf{B}_{0i}\mathbf{R}_i^{-1}\mathbf{B}'_{0i}\mathbf{K}_i - \mathbf{Q}_i \tag{17}$$

Equations (15) and (16) imply that

$$\dot{\mathbf{g}}_i^r = -\left(\mathbf{A}_{0i} - \mathbf{B}_{0i}\mathbf{R}_i^{-1}\mathbf{B}'_{0i}\mathbf{K}_i \right)' \mathbf{g}_i^r - \mathbf{K}_i\mathbf{z}_i^{r-1} + \mathbf{K}_i\mathbf{B}_{0i}\mathbf{u}_{ni} - \mathbf{Q}_i\mathbf{x}_{ni} \tag{18a}$$

The final conditions for (17) and (18a) can be determined by comparison of (16) for $t = t_f$ with (15b), which yields

$$\mathbf{K}_i(t_f) = \mathbf{S}_i \tag{18b}$$

$$\mathbf{g}_i^r(t_f) = \mathbf{S}_i\mathbf{x}_{ni}(t_f) \tag{18c}$$

Following this formulation, the optimal control for the rth optimization problem can be written as

$$\delta\mathbf{u}_i^{*r} = -\mathbf{R}_i^{-1}\left(\mathbf{B}'_{0i}\delta\mathbf{p}_i^r + \mathbf{R}_i\mathbf{u}_{ni} \right) = -\mathbf{R}_i^{-1}\mathbf{B}'_{0i}\mathbf{K}_i\mathbf{x}_i^{*r} - \mathbf{R}_i\mathbf{B}'_{0i}\mathbf{g}_i^r - \mathbf{u}_{ni} \tag{19}$$

Hence, from (8a) and (19), the optimal state trajectory $\delta\mathbf{x}^{*r}(t)$ is the solution to

$$\delta\dot{\mathbf{x}}_i^{*r} = \left[\mathbf{A}_{0i} - \mathbf{B}'_{0i}\mathbf{R}_i^{-1}\mathbf{B}_{0i}\mathbf{K}_i \right] \delta x_i^{*r} - \mathbf{B}_{0i}\mathbf{R}_i^{-1}\mathbf{B}_{0i}\mathbf{R}_i^{-1}\mathbf{B}'_{0i}\mathbf{g}_i^r - \mathbf{B}_{0i}\mathbf{u}_{ni} + \mathbf{z}_i^{r-1} \tag{20a}$$

$$\delta\mathbf{x}_i^*(t) = \delta\mathbf{x}_{i0}(t), \quad t - h_p \leqslant t \leqslant t_0 \tag{20b}$$

$$\delta \mathbf{u}_i^*(t) = \mathbf{u}_{i0}(t), \quad t - h_q \leqslant t \leqslant t_0 \tag{20c}$$

The chosen optimization strategy, however, cannot achieve the optimal performance index

$$\hat{J}^* = \sum_{i=1}^{N} \hat{J}_i^* \tag{21}$$

by using only the local control $\delta \mathbf{u}^*(t)$, unless all the perturbations are absent ($\mathbf{z}_i = 0$, $i=1,2,\ldots,N$). To solve this problem, let's form the second level problem which is essentially updating the new coordination vector $\alpha_i(t)$ and $\mathbf{z}_i(t)$. For this purpose, define the additively separable Lagrangian

$$L = \sum_{i=1}^{N} L_i \tag{22}$$

where

$$L_i = \tfrac{1}{2}\, \delta \mathbf{x}'_i(t_f) \left[\mathbf{S}_i \delta \mathbf{x}_i(t_f) + 2\mathbf{S}_i \mathbf{x}_{ni}(t_f) \right] +$$
$$\tfrac{1}{2} \int_{t_0}^{t_f} \left\{ \left[\delta \mathbf{x}'_i \left(\mathbf{Q}_i \delta \mathbf{x}_i + 2\mathbf{Q}_i \mathbf{x}_{ni} \right) + \mathbf{u}'_i \left(\mathbf{R}_i \delta \mathbf{u}_i + 2\mathbf{R}_{in} \mathbf{u}_{ni} \right) \right] + \right.$$
$$\left. \alpha'_i(\mathbf{z}_i - \mathbf{z}_{ie}) + \delta \mathbf{p}'_i(-\dot{\mathbf{x}}_i + \mathbf{A}_{0i}\mathbf{x}_i + \mathbf{B}_{0i}\mathbf{u}_i + \mathbf{z}_i) \right\} dt$$

The values of $\alpha_i(t)$ and $\mathbf{z}_i(t)$ can be obtained by

$$0 = \frac{\partial L_i}{\partial \alpha_i} = \mathbf{z}_i - \mathbf{z}_{ie} \tag{24a}$$

$$0 = \frac{\partial L_i}{\partial \mathbf{z}_i} = \alpha_i + \delta \mathbf{p}_i \tag{24b}$$

which provides at rth iteration

$$\mathbf{z}_i^r = \mathbf{z}_{ie}^{r-1} = \hat{\mathbf{A}}_{0i}\delta_{\mathbf{x}}^{r-1} + \hat{\mathbf{B}}_{0i}\delta \mathbf{u}^{r-1} + \sum_{k=1}^{p} \mathbf{A}_{ki}\delta \mathbf{x}^{r-1}(t-h_{kx}) + \sum_{\ell=1}^{q} \mathbf{B}_{\ell i}\mathbf{u}^{r-1}(t-h_{\ell u}) \tag{25a}$$

$$\alpha_i^r = -\,\delta \mathbf{p}_i^{r-1} \tag{25b}$$

Once the total system perturbation error, in normalized form,

$$E = \sum_{i=1}^{N} \int_{t_0}^{t_f} \left[(\mathbf{z}_i - \mathbf{z}_{ie}) \right] dt/\Delta t \tag{26}$$

is sufficiently small an optimal solution has been obtained. Here, Δt is the step size of integration.

8.6.3 Computational Algorithm

In this section a computational algorithm is presented for the control of a TD system via the interaction predication scheme.

1. Algorithm.

Step 1: Choose an arbitrary nominal control $\mathbf{u}_n(t)$ which satisfies (1).

Step 2: Solve (1) with the known control vector $\mathbf{u}(t)$ to obtain the nominal state vector $\mathbf{x}_n(t)$.

Step 3: Obtain a linear approximation system of (6a) by substituting the nominal trajectories $(\mathbf{x}_n(t), \mathbf{u}_n(t))$ in (6d), (6e), and decompose the corresponding perturbation system as in (8) and (9).

Step 4: Solve (17) with final condition (18b) to obtain $\mathbf{K}_i(t)$ and store. $i=1,2,\ldots,N$.

Step 5: Determine $\mathbf{x}_{i0}(t)$ by state transition matrix and $\beta_{i0}(0)$ $i=1,2,\ldots,N$.

Step 6: Set $r=1$, use $\delta\mathbf{x}_{i0}(t)$ and $\delta\mathbf{u}_{i0}(t)$, (from (8c)) in (9) to obtain $\mathbf{z}_{i0}(t)$, $t \geqslant t_0$.

Step 7. Use the final condition (18c) and (18a) to obtain $\mathbf{g}_i^r(t)$, and get $\delta\mathbf{p}_i^r(t)$ from (16) thereafter. Store $g_i^r(t)$ and $p_i^r(t)$, $i=1,2,\ldots,N$.

Step 8. Solve the ith suboptimal control problem to find $\delta\mathbf{x}^{*r}$ and $\delta\mathbf{u}^{*r}(t)$, $t \geqslant t_0$ as described by (19) and (20). Store $\delta\mathbf{x}^{*r}(t)$, $\delta\mathbf{u}^{*r}(t)$, $i=1,2,\ldots,N$.

Step 9. Use the stored $\delta\mathbf{u}^{*r}(t)$ and $\delta\mathbf{x}^{*r}(t)$, and $p_i(t)$ in (25) to update $\alpha_i(t)$ and $\mathbf{z}_i(t)$, $i=1,2,\ldots,N$.

Step 10. Check for the convergence at the second level by evaluating the overall perturbation error from (26).

Step 11. If $E < \epsilon$ for a prespecified small constant then

$$\delta \mathbf{u}^*(t) = \delta \mathbf{u}^{*r}(t) + \mathbf{u}_n(t) \tag{27}$$

$$\delta \mathbf{x}^*(t) = \delta \mathbf{x}^{*r}(t) + \mathbf{x}_n(t) \tag{28}$$

and stop, otherwise set $r = r + 1$ and go to Step 7.

A simplified pictorial representation of the above algorithm is shown in Figure 1.

The convergence of this algorithm can be proved similar to Algorithm 8.4.4 and it is left as an exercise (see Problem 8.7). The following example, taken from Jamshidi and Wang [8.25] illustrates the algorithm.

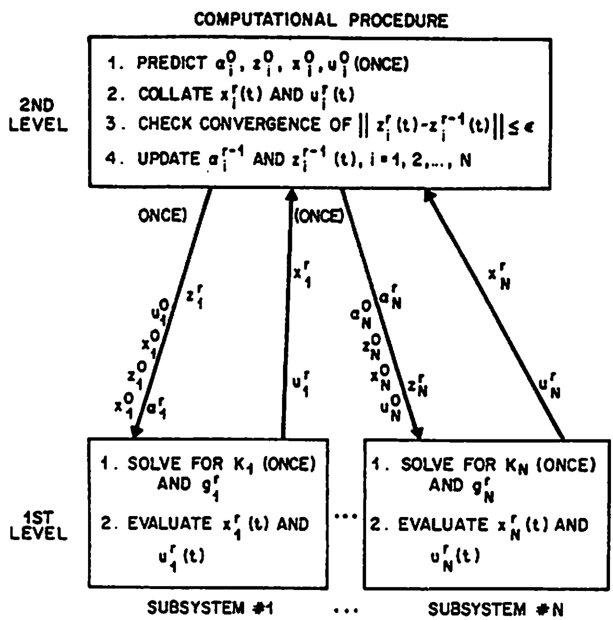

Figure 1. Block Diagram for the Computational Procedure

2. Example. Consider a model of a typical water resource system, shown in Figure 2.

The list of symbols used in the system model is given below:

d = irrigation release ($10^9 ft^3$), α = fraction of water return to main stream

h_i = i=1,2,3 delay time (month or week), Y input water flow ($10^9\ ft^3$)

S = reservoir storage ($10^9\ ft^3$), e evaporation ($10^9\ ft^3$)

u = water release ($10^9\ ft^3$)

Table 1 presents a summary of the state and control variables used in this example.

TABLE 1. State and Control Variables for Example 2

Vector	State	Control
Subsystem No. 1		
Physical	S_1,S_2	U_1,d_1,U_2
Mathematical	x_1,x_2	u_1,u_2,u_3
Subsystem No. 2		
Physical	S_3,S_4	d_2,U_3,d_3,U_4
Mathematical	x_3,x_4	u_4,u_5,u_6,u_7

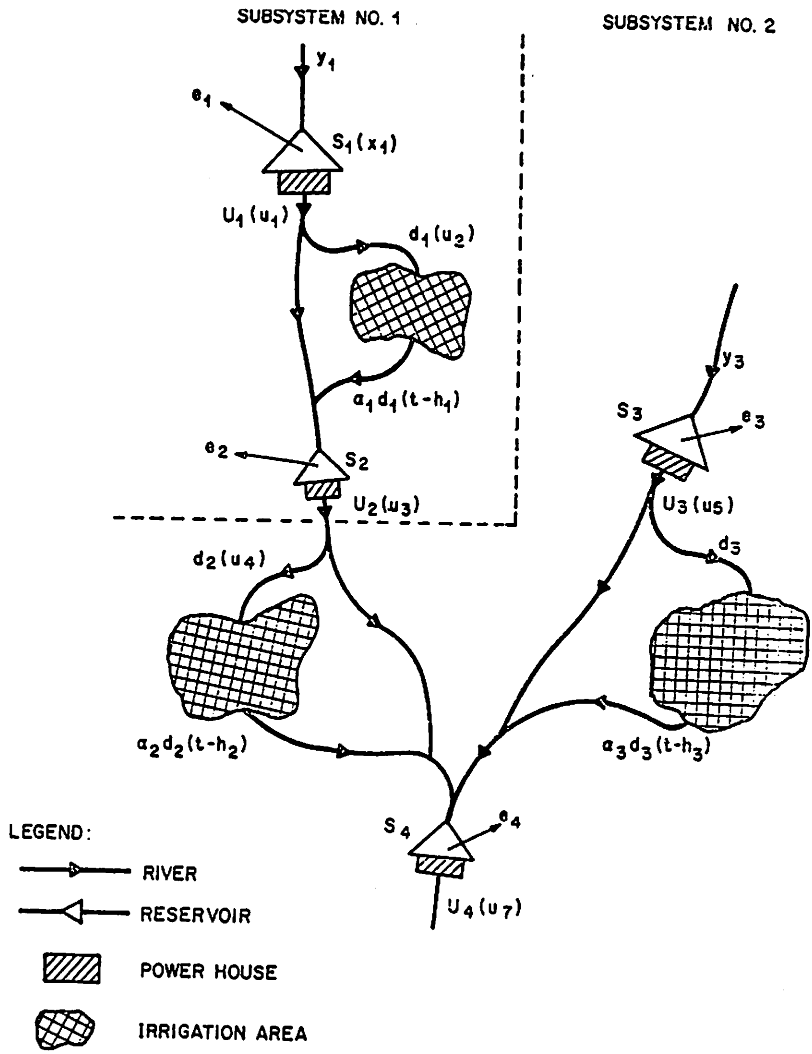

Figure 2. Model of Water Resource System of Example 2

The state equation derived from the system model is both nonlinear and time-delay as shown below:

$$
\begin{aligned}
\dot{x}_1(t) &= y_1(t) - u_1(t) - e_1(t) \\
\dot{x}_2(t) &= u_1(t) - u_2(t) + \alpha_1 u_2(t-h_1) - u_3(t) - e_2(t) \\
\dot{x}_3(t) &= y_3(t) - u_5(t) - e_3(t) \\
\dot{x}_4(t) &= u_3(t) - u_4(t) + \alpha_2 u_4(t-h_2) + u_5(t) - u_6(t) \\
&\quad + \alpha_3 u_6(t-h_3) - u_7(t) - e_4(t)
\end{aligned}
\tag{29}
$$

where the evaporation $e_i(t)$ are normally assumed to be a nonlinear empirical function of the reservoir capacity, *i.e.*.,

$$e_i(t) \triangleq \beta_i x_i^{2/3}(t) \quad , \quad i=1,2,3,4. \tag{30}$$

A suitable cost function for this system is

$$\max_{\mathbf{u},\mathbf{x}} \; J = \tfrac{1}{2}\, \mathbf{x}'(12)\mathbf{S}(12)\mathbf{x}(12) + \tfrac{1}{2} \int_1^{12} \Big[\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}(t)\mathbf{u}(t) \Big]\, dt$$

where

$$
\begin{aligned}
&\mathbf{R} = block\text{-}diag(\mathbf{R}_1, \mathbf{R}_2) \\
&\mathbf{R}_1 = diag(-1,0,-2) \quad , \quad \mathbf{R}_2 = diag(-1,-1,-2,-1) \\
&\mathbf{Q}_1 = \mathbf{S}_1 = diag(-1,-2) \quad , \quad \mathbf{Q}_2 = \mathbf{S}_2 = -\mathbf{I}_2
\end{aligned}
\tag{32}
$$

Here the weighting matrices $\mathbf{R}_i$, $\mathbf{Q}_i$, $\mathbf{S}_i$, $i = 1,2$ are all chosen to be negative definite for the maximization problem. The parameters in the state equation are chosen to be:

$$
\begin{aligned}
&\alpha_1 = 20\% \qquad \alpha_2 = 30\% \qquad\qquad \alpha_3 = 25\% \\
&\beta_i = 15\%, i = 1, \ldots, 4 \qquad h_j = 1 \; month, j = 1,2,3
\end{aligned}
\tag{33}
$$

Simulating from the data given in Peters et al [8.40], the input water flow is chosen to be:

$$
\begin{aligned}
y_1 &= 1 + 0.6 sin(\pi/6t), \qquad t = 1,2,\ldots, 12 \\
y_3 &= 0.8 + 0.6 sin(\pi/6t)
\end{aligned}
\tag{34}
$$

Using the above data, the system can then be rewritten as:

$$\dot{x}_1 = 1 + 0.6sin(\pi/6t) - u_1 - 0.15x_1^{2/3} \tag{35a}$$

$$\dot{x}_2 = u_1 - u_2 - u_3 + 0.2u_2(t-1) - 0.15x_2^{2/3} \tag{35b}$$

$$\dot{x}_3 = 0.8 + 0.6sin(\pi/6t) - u_5 - 0.15x_3^{2/3} \tag{35c}$$

$$\dot{x}_4 = u_3 - u_4 + 0.3u_4(t-1) + u_5 - u_6 + 0.25u_6(t-1) - u_7 - 0.158x_4^{2/3} \tag{35d}$$

$$\text{for } t = 1,2,\ldots, 12 \tag{35e}$$

The initial conditions used here are:

$$x_1 = 0.6, \quad u_1 = 1, \quad u_5 = 0.8 \tag{36a}$$

$$x_2 = 0.7, \quad u_2 = 0.4, \quad u_6 = 0.4 \tag{36b}$$

$$x_3 = 0.8, \quad u_3 = 0.5, \quad u_7 = 0.7 \tag{36c}$$

$$x_4 = 0.7, \quad u_4 = 0.3 \quad (10^9 ft^3) \tag{36d}$$

First, choose the nominal control trajectories to be:

$$u_{u1} = u_1 = 1 \tag{37a}$$

$$u_{n2} = u_2 = 0.4 \quad u_{n5} = u_5 = 0.8 \tag{37b}$$

$$u_{n3} = u_3 = 0.5 \quad u_{n6} = u_6 = 0.4 \tag{37c}$$

$$u_{n4} = u_4 = 0.3 \quad u_{n7} = u_7 = 0.6 \tag{37d}$$

Substitute the nominal control vectors in (35) to find the nominal state trajectories $x_{ni}(t)$ $i = 1,2,3,4$. From (6b) and (6e) using the nominal trajectories values, a linear model can thus be obtained. The system is then decomposed into two subsystems due to its physical form. The Algorithm is applied to solve the problem using an HP-9845 computer. The perturbation error converged to 8.438302×10^{-12} in twelve iterations. Figures 3 to 7 shown the interaction error and trajectories of the state, control and vectors and input water flow.

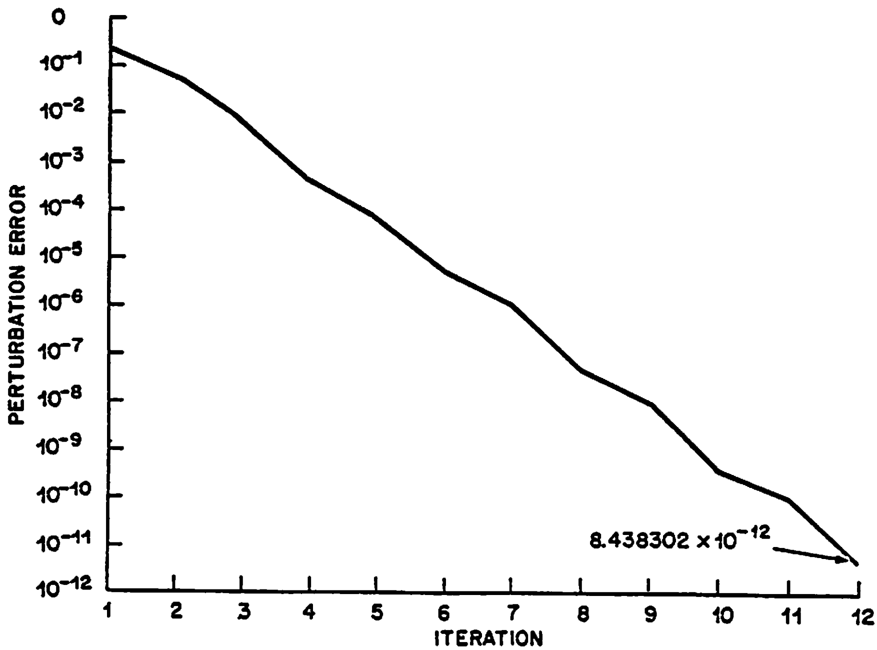

Figure 3. Perturbation Error Versus Iteration for Example 2

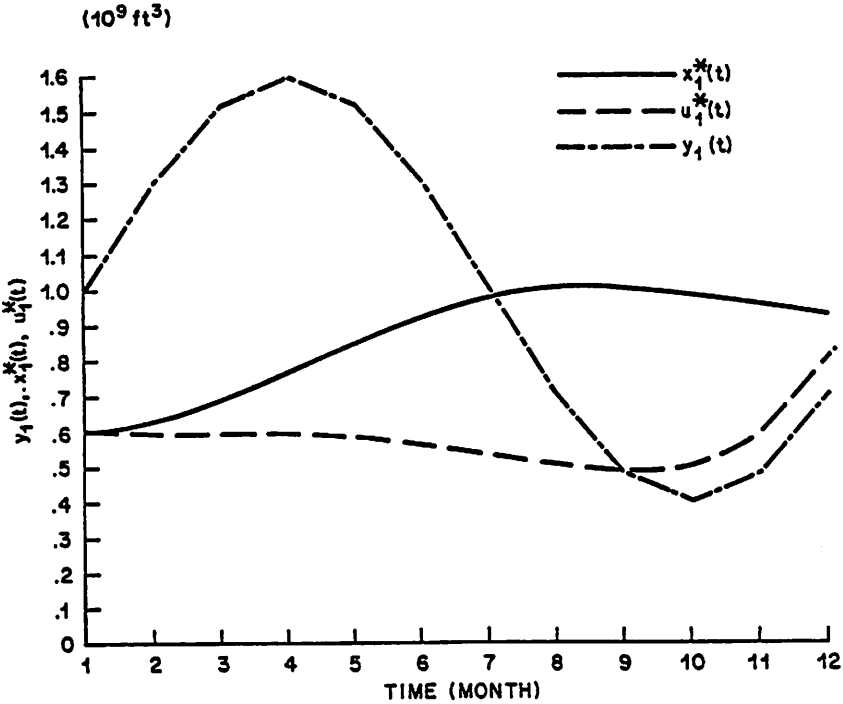

Figure 4. Time Responses of $y_l(t)$, $x_1^*(t)$ and $u_1^*(t)$ for Example 2

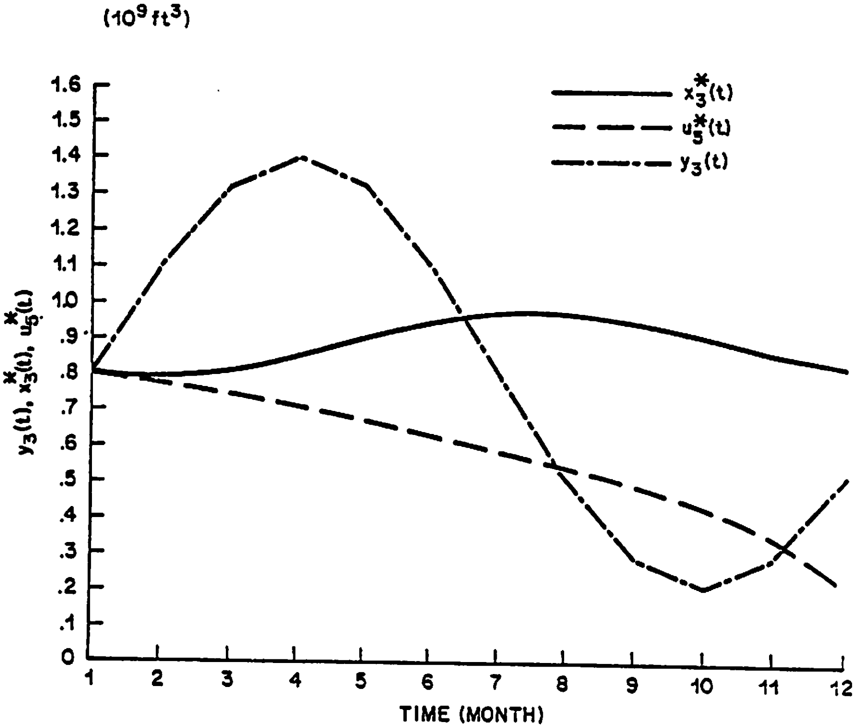

Figure 5. Time Responses of $y_3(t)$, $x_3^*(t)$ and $u_5^*(t)$ for Example 2

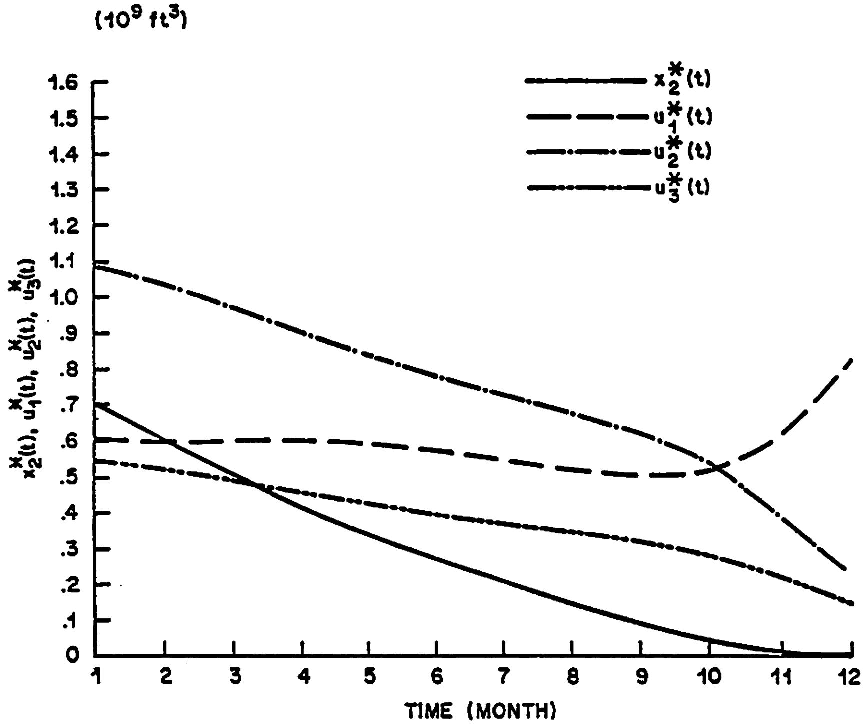

Figure 6. Time Responses of $x_2^*(t)$, $u_1^*(t)$, $u_2^*(t)$ and $u_3^*(t)$ for Example 2

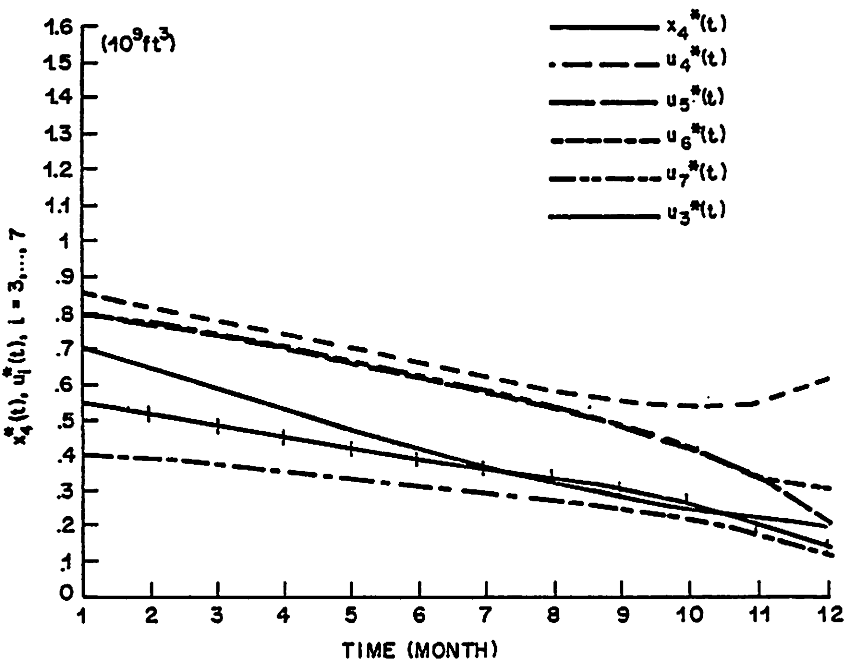

Figure 7. Time Responses of $x_4^*(t)$, $u_3^*(t)$, $u_4^*(t)$, $u_5^*(t)$, $u_6^*(t)$ and $u_7^*(t)$ for Example 2

8.7 HIERARCHICAL CONTROL OF TIME-DELAY SYSTEMS VIA COSTATE PREDICTION

In the previous section the extension of the non-delay systems optimization via interaction prediction method was presented. As it was demonstrated by Algorithm 8.5.1, this demanded the solutions of a set of N Riccati matrix equations and N adjoint vector equations at the first level, while the second level solution contributed a set of simple substitutions. Thus the key variable at the first level, as discussed in Section 8.3.3, is the costate vector which must be obtained either directly (nonlinear case) or indirectly (linear case) each time the hierarchical iteration is switched to the first level. In this section the costate prediction scheme of non-delay discrete-time delay systems of Section 8.3.3 is extended for TD systems.

Consider a nonlinear discrete-time system with time delay described in state form by:

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k),\mathbf{u}(k),\mathbf{x}(k-h),\mathbf{u}(k-d),(k+1),k) \tag{1}$$

with initial functions,

$$\mathbf{x}(\tau) = \phi(\tau), \qquad -h \leqslant \tau \leqslant 0 \tag{2}$$

$$\mathbf{u}(\tau) = \beta(\tau), \qquad -d \leqslant \tau \leqslant 0 \tag{3}$$

In the above relations, $\mathbf{x}(k)$ and $\mathbf{u}(k)$ are n-dimensional and m-dimensional state and control vectors, respectively.

The objective function is

$$J = g(\mathbf{x}(K)) + \sum_{k=0}^{K-1} (L(\mathbf{x}(k),k) + G(\mathbf{u}(k),k)) \tag{4}$$

The problem is to find a sequence of control vectors $\mathbf{u}(k)$, $k=0,1,\ldots,K-1$, which minimizes the cost function J in (4) while satisfying (1) - (3). The term $g(\mathbf{x}(K))$ in (4) represents a penalty term, $L(\cdot)$ and $G(\cdot)$ are scalar functions of their arguments.

Within the usual hierarchical control structure, the system (1)-(3) must be decomposed into N subsystems. In order to undertake the optimization process, consider the discrete-time TD system (1) with initial functions (2)-(3) and a non-separable cost functional

$$J = \sum_{k=0}^{K-1} \Psi(\mathbf{x}(k),\mathbf{u}(k),k) \quad . \tag{5}$$

Assume that the system is asymptotically stable and has an equilibrium point. Now let us rewrite (1) and (5) as

$$\begin{aligned} \mathbf{x}(k+1) = {} & \mathbf{A}(\mathbf{x},\ \mathbf{u},\ \mathbf{x}(k-h),\ \mathbf{u}(k-d),((k+1),k))\mathbf{x}(k) + \\ & \mathbf{B}(\mathbf{x},\mathbf{u},\mathbf{x}(k-h),\mathbf{u}((k-d),(k+1),k))\mathbf{u}(k) \\ & + \mathbf{c}(\mathbf{x},\mathbf{u},\mathbf{x}(k-h),\mathbf{u}(k-d),(k+1),k) \end{aligned} \tag{6}$$

$$J = \sum_{k=0}^{K-1} \tfrac{1}{2}[\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}(k)\mathbf{u}(k) + \psi(\mathbf{x},\mathbf{u},k)] \tag{7}$$

where $\mathbf{A}(\cdot)$ and $\mathbf{B}(\cdot)$ contain the block-diagonal parts of the Jacobian Matrices $\mathbf{f}'_x$ and $\mathbf{f}'_u$. Vector $\mathbf{c}(\cdot)$ represents the off-diagonal parts, other nonlinearities, and time-delay terms involved in the original system's dynamics. The only remaining term is,

$$\psi(\cdot) = \Psi(\mathbf{x},\mathbf{u},k) - \tfrac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x} - \tfrac{1}{2}\mathbf{u}'\mathbf{R}\mathbf{u} \tag{8}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are block-diagonal matrices. Note that the argument k has been dropped for clarity when ambiguity is not caused.

In the two-level costate prediction scheme for the non-delayed case predicted state $\mathbf{x}^*$ and control $\mathbf{u}^*$ vectors fix $\Psi(\cdot)$ in the cost function (7) and $\mathbf{c}(\cdot)$ in the state equation (6) by adding constraints $\mathbf{x}^*(k)=\mathbf{x}(k)$ and $\mathbf{u}^*(k)=\mathbf{u}(k)$. The delay terms can be fixed in a similar fashion. Therefore, since the nonlinear and delayed terms are fixed, the first-level problems represent linear unretarded discrete-time systems. The second-level problem is to satisfy the two assumed conditions relating $\mathbf{x}^*$ to $\mathbf{x}$ and $\mathbf{u}^*$ to $\mathbf{u}$, and to update the value of the costate vector $\mathbf{p}(k)$.

In summary, the problem is reformulated as follows:

$$\underset{\mathbf{x},\mathbf{u}}{Min}\, J = \sum_{k=0}^{K-1} \left\{ \tfrac{1}{2}[\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}(k)\mathbf{u}(k)] + \psi^*(\cdot) \right\} \tag{9}$$

subject to:

$$\mathbf{x}(k+1) = \mathbf{A}^*(\cdot)\mathbf{x}(k) + B^*(\cdot)\mathbf{u}(k) + \mathbf{c}^*(\cdot) \tag{10}$$

$$\mathbf{x}^*(k) = \mathbf{x}(k) \tag{11}$$

$$\mathbf{u}^*(k) = \mathbf{u}(k) \tag{12}$$

where the superscript * denotes the evaluation at the predicted values $\mathbf{x}^*$ and $\mathbf{u}^*$.

Once the system has been reformulated, the next step is to define the Hamiltonian function for the new problem, determine the conditions for optimality, and formulate the costate prediction algorithm for the TD discrete-time system.

First define the Hamiltonian function for the overall system defined by (9) through (12).

$$H(\cdot) = \tfrac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}' + \tfrac{1}{2}\mathbf{u}'\mathbf{R}\mathbf{u} + \psi^*(\cdot) + \mathbf{p}'(k+1)\left[\mathbf{A}^*\mathbf{x}+\mathbf{B}^*\mathbf{u} + \mathbf{c}^*(\cdot)\right]$$

$$+ \alpha'(\mathbf{x}-\mathbf{x}^*) + \beta'(\mathbf{u}-\mathbf{u}^*) \tag{13}$$

where $\mathbf{p}$ is the costate vector and α and β are vectors of Lagrange multipliers. Assume that the Hamiltonian of (13) is additively separable for given $\mathbf{x}^*$ and $\mathbf{u}^*$, *i.e.*,

$$H = \sum_{i=1}^{N} H_i \tag{14}$$

where the index i denotes the ith subsystem and

$$H_i = \tfrac{1}{2}(\mathbf{x}'_i\mathbf{Q}_i\mathbf{x}_i+\mathbf{u}'_i\mathbf{R}_i\mathbf{u}_i) + \psi_i^*(\cdot) + \mathbf{p}'_i(k+1)(\mathbf{A}_i^*(\cdot)\mathbf{x}_i$$
$$+ \mathbf{B}_i^*(\cdot) + \mathbf{c}_i^*(\cdot)) + \alpha_i'(\mathbf{x}_i-\mathbf{x}_i^*) + \beta'_i(\mathbf{u}_i-\mathbf{u}_i^*). \tag{15}$$

Again, the argument k has been dropped for simplicity. The necessary conditions for optimizing the first level's subsystem are [8.26, 8.41]

$$0 = \frac{\partial H_i(\cdot)}{\partial \mathbf{u}_i(k)} \tag{16a}$$

$$\mathbf{x}_i(k+1) = \frac{\partial H_i(\cdot)}{\partial \mathbf{p}_i(k+1)} \tag{16b}$$

$$\mathbf{p}_i(k) = \frac{\partial H_i(\cdot)}{\partial \mathbf{x}_i(k)} \tag{16c}$$

These conditions yield

$$u_i(k) = R_i^{-1}(B_i^{*\prime}p_i(k+1) + \beta_i(k)) \quad (17a)$$

$$u_i(k) = \beta(k), \ -d \leqslant k \leqslant 0 \quad (17b)$$

$$x_i(k+1) = A_i^* x_i(k) - B_i^* R_i^{-1}[B_i^{*\prime}p_i(k+1) + \beta_i(k)] + c_i^*(\cdot) \quad (18a)$$

$$x_i(k) = \phi_i(k), \quad -h \leqslant k \leqslant 0 \quad (18b)$$

$$p_i(k) = Q_i x_i(k) + A_i^{*\prime}p_i(k+1) + \alpha_i(k) \quad (19a)$$

$$p_i(K) = 0 \quad (19b)$$

The above equations constitute a linear time-varying discrete-time TPBV problem whose solution through a discrete Riccati formulation will not be necessary by the costate prediction solution since the costate vector will be one of the coordinating variables.

The remaining necessary conditions are

$$\frac{\partial H(\cdot)}{\partial \alpha} = 0 \quad (20)$$

$$\frac{\partial H(\cdot)}{\partial \beta} = 0 \quad (21)$$

$$\frac{\partial H(\cdot)}{\partial x^*} = 0 \quad (22)$$

$$\frac{\partial H(\cdot)}{\partial u^*} = 0 \quad (23)$$

From (22) and (23) we obtain the following expressions for $\alpha(k)$ and $\beta(k)$:

$$\alpha(k) = \left[E'_x(*)+F'_x(*)+G'_x(*)\right]p(k+1)+\psi'_x(*) \quad (24)$$

$$\beta(k) = \left[E'_u(*)+F'_u(*)+F'_u(*)+G'_u(*)\right]p(k+1)+\psi'_u(*) \quad (25)$$

where

$$E_x(*) = \frac{\partial}{\partial x^*}\left[A(x^*(k),u^*(k),x^*(k-h),\ u^*(k-d),(k+1),k)x(k)\right] \quad (26)$$

$$F_x(*) = \frac{\partial}{\partial x^*}\left[B(x^*(k),u^*(k),x^*(k-h),\ u^*(k-d),(k+1),k)u(k)\right] \quad (27)$$

$$G_x(*) = \frac{\partial}{\partial x^*}\left[c(x^*(k),u^*(k),x^*(k-h),\ u^*(k-d),(k+1),k)\right] \quad (28)$$

$$\psi_x(*) = \frac{\partial}{\partial x^*}(x^*(k),u^*(k)) \quad (29)$$

Similarly,

$$\mathbf{F}_u(*) = \frac{\partial}{\partial \mathbf{u}^*}\left[\mathbf{A}(\mathbf{x}^*(k),\mathbf{u}^*(k),\mathbf{x}^*(k-h),\ \mathbf{u}^*(k-d),(k+1),k)\mathbf{x}(k)\right] \tag{30}$$

$$\mathbf{F}_u(*) = \frac{\partial}{\partial \mathbf{u}^*}\left[\mathbf{B}(\mathbf{x}^*(k),\mathbf{u}^*(k),\mathbf{x}^*(k-h),\ \mathbf{u}^*(k-d),(k+1),k)\mathbf{u}(k)\right] \tag{31}$$

$$\mathbf{G}_u(*) = \frac{\partial}{\partial \mathbf{u}^*}\left[\mathbf{c}(\mathbf{x}^*(k),\mathbf{u}^*(k),\mathbf{x}^*(k-h),\ \mathbf{u}^*(k-d),(k+1),k)\right] \tag{32}$$

$$\psi_u(*) = \frac{\partial}{\partial \mathbf{u}^*}(\mathbf{x}^*(k),\mathbf{u}^*(k)) \tag{33}$$

Conditions (20) and (21) lead to

$$\mathbf{x}^*(k) = \mathbf{x}(k) \tag{34}$$

$$\mathbf{u}^*(k) = \mathbf{u}(k) \tag{35}$$

Through the above formulation the problem can be reduced as follows. Substitute predicted sequences $\mathbf{x}^*(k)$, $\mathbf{u}^*(k)$, $\beta(k)$, $\mathbf{p}(k)$ into (17) and (18) to obtain the sequences $\mathbf{u}_i(k)$ and $\mathbf{x}_i(k)$, respectively. These sequences are collated and used to calculate $\alpha(k)$. These computations constitute the first-level solutions. At the second level, update the coordination vector

$$\mathbf{z}(k) = [\mathbf{p}(k),\ \mathbf{x}^*(k),\ \mathbf{u}^*(k),\ \beta] \tag{36}$$

by using equations (19), (25), (34) and (35).

The following algorithm summarizes the costate prediction control of a discrete-time retarded system.

1. Algorithm.

1. Guess the vector sequences $\mathbf{p}(k)$, $\mathbf{x}^*(k)$, $\mathbf{u}^*(k)$, $\beta(k)$, and set the iteration index q=1.

2. At the first level substitute the sequences $\mathbf{p}^q(k)$, $\mathbf{x}^{*q}(k)$, $\mathbf{u}^{*q}(k)$, and $\beta^q(k)$, $0 \leqslant k \leqslant -1$, into the right-hand sides of (17) and (18) to obtain the sequences $\mathbf{u}_i(k)$ and $\mathbf{x}_i(k)$ for $0 \leqslant k \leqslant K-1$ and $i=1,2..., N$. Obtain $\alpha(k)$ using (24).

3. At the second level update (36) by substituting $\mathbf{x}_i^q(k)$ and $\mathbf{u}_i^q(k)$ directly into (19), (34)-(35), and (25). Therefore

$$\mathbf{p}_i^{q+1}(k) = \mathbf{Q}_i\mathbf{x}_i(k) + \alpha'^*_i\mathbf{p}_i^q(k+1) + \alpha_i^q(k) \tag{37}$$

$$\mathbf{x}^{*q+1}(k) = \mathbf{x}^q(k) \tag{38}$$

$$\mathbf{u}^{*q+1}(k) = \mathbf{u}^q(k) \tag{39}$$

$$\beta^{q+1}(k) = \left(\mathbf{E}'_{\mathbf{u}}() + F'_{\mathbf{u}}() + \mathbf{G}'_{-u}()\right)\Big|_{u^q}^{\mathbf{x}^{*q}} + \psi'_{\mathbf{u}}()\Big|_{u^q}^{\mathbf{x}^{*q}} \tag{40}$$

4. If $\|\mathbf{z}^{q+1}(k) - \mathbf{z}^q(k)\| \leqslant \epsilon$ *for* $0 \leqslant k \leqslant 1$, where ϵ is some prespecified tolerance, stop and $\mathbf{u}^q(k)$ is the optimal control. Otherwise go to Step 2.

A schematic representation of this Algorithm is shown in Figure 1.

2. Example. Consider the nonlinear time-delay discrete-time system:

$$\begin{aligned}
x_1(k+1) &= 0.9x_1(k) + 0.2x_2(k) - 0.3x_1(k-1) + 0.1u_1(k) \\
x_2(k+1) &= -0.5x_1(k) + 0.8x_1(k)x_2(k) - 0.3x_2(k) + 0.1u_2(k) \\
\mathbf{x}(\tau) &= \begin{bmatrix} -0.05 \\ -0.50 \end{bmatrix}, \quad -1 \leqslant \tau \leqslant 0
\end{aligned} \tag{41}$$

The problem is to minimize the cost function

$$J = ½ \sum_{k=0}^{19} [0.1x_1^2(k) + 0.1x_2^2(k) + 0.2u_1^2(k) + 0.1u_2^2(k)] \tag{42}$$

while satisfying (41).

Begin by decomposing the system into two first-order subsystems and follow the procedure outlined in Algorithm 1. Since the cost function is quadratic for this example there will be no matrix $\Psi^*()$. In addition, since the only nonlinearity and delay occurs in $\mathbf{x}^*$, the elements of the coordinator vector will be $\mathbf{p}(k)$ and $\mathbf{x}^*(k)$.

Using the predicted state variables, $\mathbf{x}^*$, the first subsystem's state equation becomes,

$$x_1(k+1) = 0.9x_1(k) + 0.2x_2^*(k) - 0.3x_1^{**}(k-1) + 0.1u_1(k) \tag{43}$$

and

$$A = 0.9;\ B_1 = 0.1;\ Q_1 = 0.1;\ R_1 = 0.2 \tag{44a}$$

$$c_1 = 0.2x_2^*(k) - 0.3x_1^*(k-1) \tag{44b}$$

$$x_1(\tau) = -0.05,\ -1 \leqslant \tau \leqslant 0 \tag{44c}$$

The second system is described by

$$x_2(k+1) = -0.3x_2(k) + 0.8x_1^*(k)x_2(k) + 0.1u_2(k) - 0.5x_1^*(k) \tag{45a}$$

$$A_2 = -0.3 + 0.8x_1^*(k); B_2 = 0.1; Q_2 = 0.1; R_2 = 0.1 \tag{45b}$$

$$c_2 = -0.5x_1^*(k) \tag{45c}$$

$$x_2(\tau) = -0.5,\ -1 \leqslant \tau \leqslant 0\,. \tag{45d}$$

The Hamiltonian is

$$H() = \sum_{i=1}^{2} \Bigg\{ \tfrac{1}{2}\mathbf{x}'_i(k)\mathbf{Q}_i\mathbf{x}_i(k) + \tfrac{1}{2}\mathbf{u}'_i(k)\mathbf{R}_i\mathbf{u}_i(k) + p'_i(k+1)\Big\{\mathbf{A}_i^*\mathbf{x}_i(k) + \mathbf{B}_i^*\mathbf{u}_i(k) + \mathbf{c}_i^*(\mathbf{x}_i^*(k) + \alpha'_i(\mathbf{x}_i^* - \mathbf{x}_i)\Big\} \tag{46}$$

Note that since there is no need to predict the control $\mathbf{u}^*(k)$ and β, $\beta(u^*-\mathbf{u})$ is not included.

The first-level problems will be to calculate the sequences $u_1(k)$, $u_2(k)$, $x_1(k)$, $x_2(k)$, $k = 0, 1, \ldots, 19$, using the values of $\mathbf{x}^*(k)$ and $\mathbf{p}(k)$ from the second level. The $\mathbf{u}(k)$ and $\mathbf{x}(k)$ sequences are used to calculate $\alpha(k)$ as in (24). The values of $\mathbf{x}(k)$, $\mathbf{u}(k)$, and $\alpha(k)$ are used at the second level to update $\mathbf{p}(k)$ and $\mathbf{x}^*(k)$, see (37)-(40). To check the interaction error, the Euclidean norm of an error vector was calculated for each iteration. The algorithm converged to an error of 8.75×10^{-5} after 16 iterations, yielding the results shown in Figures 2 through 4.

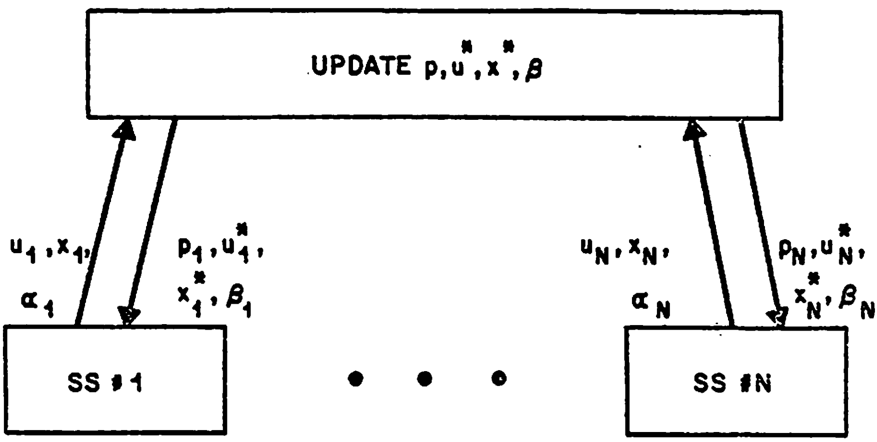

Figure 1. A Two-Level Costate Prediction Algorithm

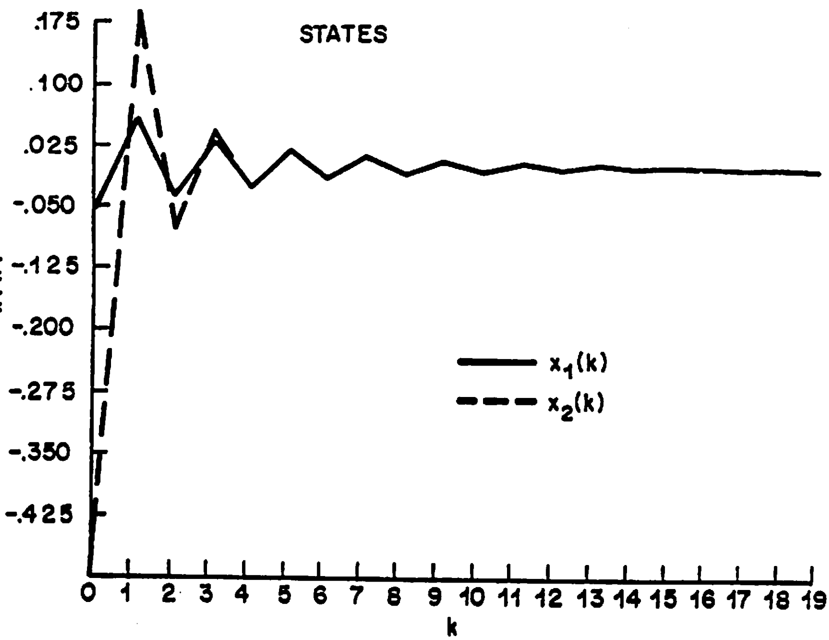

Figure 2. State Variables vs. Responses for Example 2

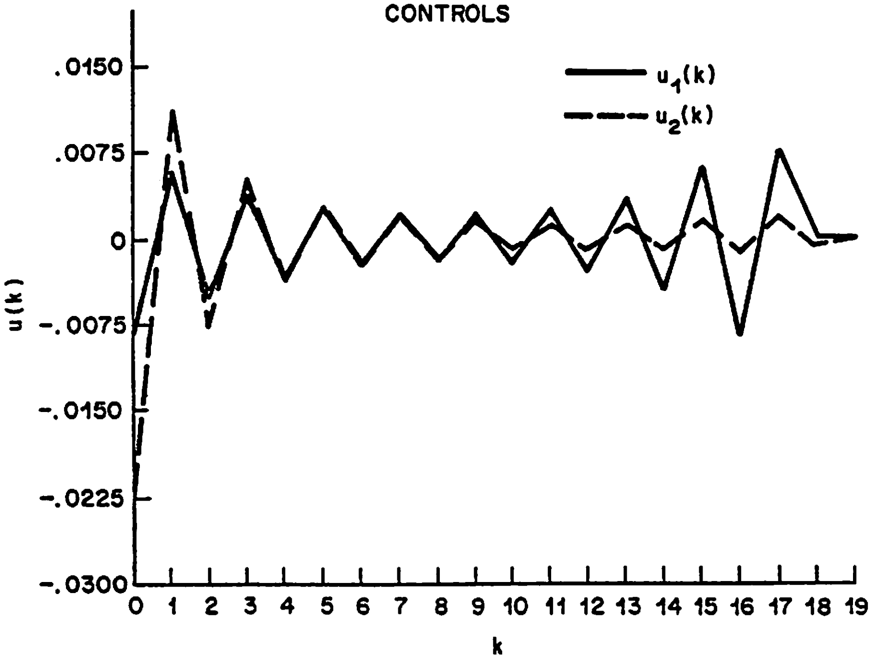

Figure 3. Control Variables vs. k for Example 2

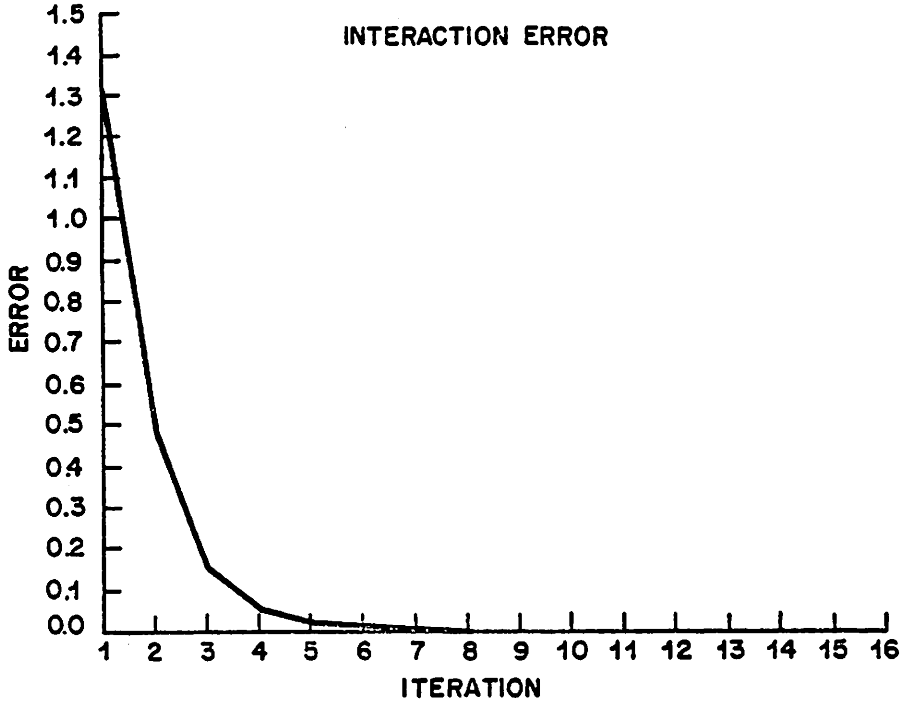

Figure 4. Interaction Error vs. k for Example 2

3. Example As another example consider the system

$$
\begin{aligned}
x_1(k+1) &= 0.5x_1(k) + 0.1x_2(k) - 0.25x_1(k)x_2(k-1) + 0.1u_1^2(k) + 0.5u_2(k) \\
x_2(k+1) &= 0.5x_2(k) - 0.1x_2^2(k) - 0.2x_1(k-1)x_2(k) + 0.1u_1(k-1)u_2(k) \\
&\quad + 0.5x_1(k) - 0.6u_1(k) \qquad (47) \\
x(\tau) &= \begin{bmatrix} -0.5 \\ 0 \end{bmatrix}, \quad -1 \leqslant \tau \leqslant 0
\end{aligned}
$$

with cost function

$$J = \tfrac{1}{2} \sum_{k=0}^{14} \left[0.1x_1^2 + 0.05x_2^2 + 0.05u_1^2 + 0.1u_2^2\right] \qquad (48)$$

This system was also simulated on the Hewlett-Packard 9845A computer by decomposing it into two first-order subsystems. For the first subsystem, the system state equation becomes

$$x_1(k+1) = 0.5x_1(k) - 0.25x_2^*(k-1)x_1(k) + 0.1u_1^*(k)u_1(k) + 0.5u_2^*(k) + 0.1x_2^*(k) \qquad (49)$$

$$A_1 = 0.5 - 0.25x_2^*(k-1);\ B_1 = 0.1u_1^*(k);\ Q_1 = 0.1;$$

$$R_1 = 0.05;\ c_1 = 0.1x_2^*(k) + 0.5u_2^*(k)$$

The second subsystem is described by

$$
\begin{aligned}
x_2(k+1) &= 0.5x_2(k) - 0.1x_2^*(k)x_2(k) + 0.5x_1^*(k) \\
&\quad - 0.2x_1^*(k-1)x_2(k) + 0.1u_1^*(k-1)u_2(k) - 0.6u_1^*(k) \qquad (50)
\end{aligned}
$$

$$
\begin{aligned}
A_2 &= 0.5 - 0.1x_2^*(k) - 0.2x_1^*(k-1) \\
B_2 &= 0.1u_1^*(k-1) \\
c_2 &= -0.6u_1^*(k) + 0.5x_1^*(k) \\
Q_2 &= 0.05;\ R_2 = 0.1
\end{aligned}
$$

Note that in this example there are nonlinearities and delay in both state and control. Therefore, the coordination vector will consist of $\mathbf{p}(k)$, $\mathbf{x}^*(k)$, $\mathbf{u}^*(k)$ and $\beta(k)$. The initial functions are

$$\mathbf{x}(k) = \begin{bmatrix} -0.5 \\ 0 \end{bmatrix}, \quad -1 \leqslant k \leqslant 0 \tag{51a}$$

$$\mathbf{u}(k) = \begin{bmatrix} 0.001 \\ -0.05 \end{bmatrix}, \quad -1 \leqslant k \leqslant 0 \tag{51b}$$

The algorithm converged to an error of 5.67×10^{-5} after 15 iterations. The results are plotted in Figures 5 through 7.

4. Example As a third example, consider a fourth-order system:

$$x_1(k+1) = 0.6x_1(k) + 0.1u_1(k) + 0.2x_2(k) - 0.3x_1(k-1) \tag{52a}$$

$$x_2(k+1) = -0.5x_1(k) - x_2(k-1)x_1(k) + 0.4x_2(k) + 0.1u_2(k) \tag{52b}$$

$$x_3(k+1) = -0.3x_3(k-1) + 0.9x_3(k) + 0.2x_4(k) + 0.1u_3(k) \tag{52c}$$

$$x_4(k+1) = -0.3x_4(k) - 0.5x_3(k) + x_3(k)x_4(k-1) + 0.1u_4(k) \tag{52d}$$

$$\mathbf{x}(\tau) = \begin{bmatrix} 0.3 \\ -0.8 \\ 0.5 \\ 0.6 \end{bmatrix}, \quad -1 \leqslant \tau \leqslant 0 \tag{52e}$$

with cost function

$$J = \tfrac{1}{2}\mathbf{x}'(k) \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix} \mathbf{x}(k) +$$

$$\sum_{k=0}^{14} \left\{ \tfrac{1}{2}\,\mathbf{x}'(k) \begin{bmatrix} 0.05 & 0 & 0 & 0 \\ 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0.05 \end{bmatrix} \mathbf{x}(k) + \right.$$

$$\left. \tfrac{1}{2}\,\mathbf{u}'(k) \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0.05 \end{bmatrix} \mathbf{u}(k) \right\} \tag{53}$$

It is convenient to decompose the system into four first-order subsystems. Following the procedure outlined before the first subsystem is

$$
\begin{aligned}
&x_1(k+1) = 0.6x_1(k) + 0.1u_1(k) + 0.2x_2^*(k) - 0.3x_1^*(k-1) \\
&A_1 = 0.6;\ B_1 = 0.1;\ Q_1 = 0.05;\ S_1 = 0.1;\ R_1 = 0.1; \\
&c_1 = 0.2x_2^*(k) - 0.3x^*(k-1)
\end{aligned} \tag{54}
$$

The second subsystem becomes

$$
\begin{aligned}
&x_2(k+1) = -\ 0.5x_1^*(k) - x_2^*(k-1)x_1^*(k) + 0.4x_2(k) + 0.1u_2(k) \\
&A_2 = 0.4;\ B = 0.1;\ c_2 = -0.5x_1^*(k) - x_2^*(k-1)x_1^*(k); \\
&Q_2 = 0.05;\ R_2 = S_2 = 0.1
\end{aligned} \tag{55}
$$

The third subsystem is

$$
\begin{aligned}
&x_3(k+1) = -\ 0.3x_3^*(k-1) + 0.9x_3(k) + 0.2x_4^*(k) + 0.1u_3(k) \\
&A_3 = 0.9;\ B_3 = 0.1;\ c_3 = 0.2x_4^*(k) - 0.3x_3^*(k-1);\ Q_3 = 0.05; \\
&R_3 = 0.05;\ S_3 = 0.1
\end{aligned} \tag{56}
$$

Finally, the fourth subsystem is described by

$$
\begin{aligned}
&x_4(k+1) = -0.3x_4(k) - 0.5x_3^*(k) + x_3^*(k) + x_3^*(k)x_4^*(k-1) + 0.1u_4(k) \\
&A_4 = -0.3;\ B_4 = 0.1;\ c_4 = x_3^*(k)x_4^*(k-1) - 0.5x_3^*(k);\ Q_4 = 0.05; \\
&R_4 = 0.05;\ S_4 = 0.1
\end{aligned} \tag{57}
$$

As in Example 1, the first level problem is to compute the local state and control sequences, $x_i(k)$ and $u_i(k)$, and Lagrange multiplier $\alpha_i(k)$. The second level uses these values to update the coordination vector, containing $p(k)$ and $x^*(k)$, for the next iteration. The algorithm converged to an error of 0.0465 after 40 iterations. The results are plotted in Figures 8 through 10. In Figure 10, the error is only plotted for the first 20 iterations. To plot more points required more memory than was available.

Comparable results were obtained for similar examples for the non-delayed case by Hassan and Singh [8.36], Mahmoud et. al [8.35] using more powerful computers. Overall, the examples were quite successful, all showing acceptable convergence and system behavior.

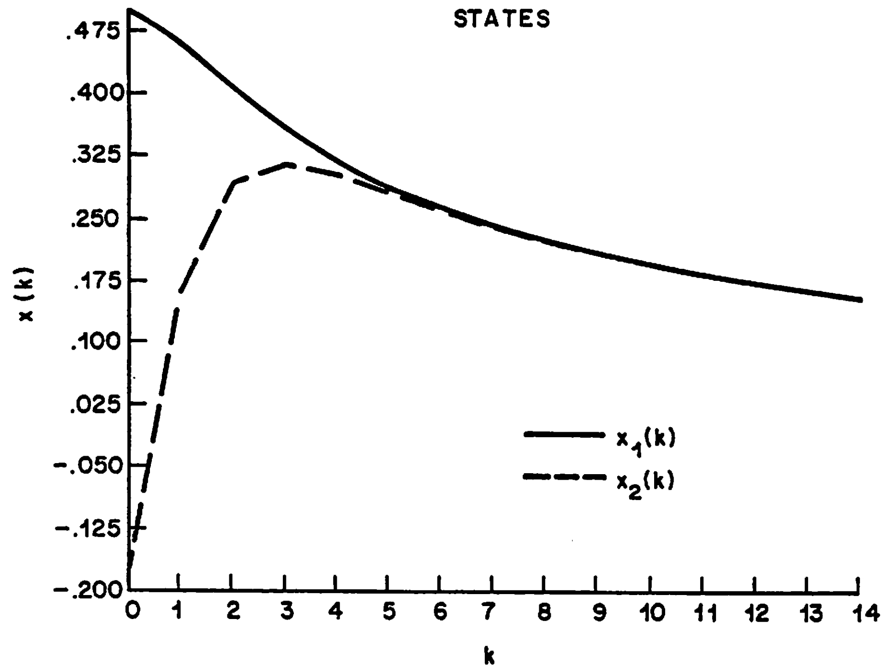

Figure 5. State Variables vs. k for Example 2

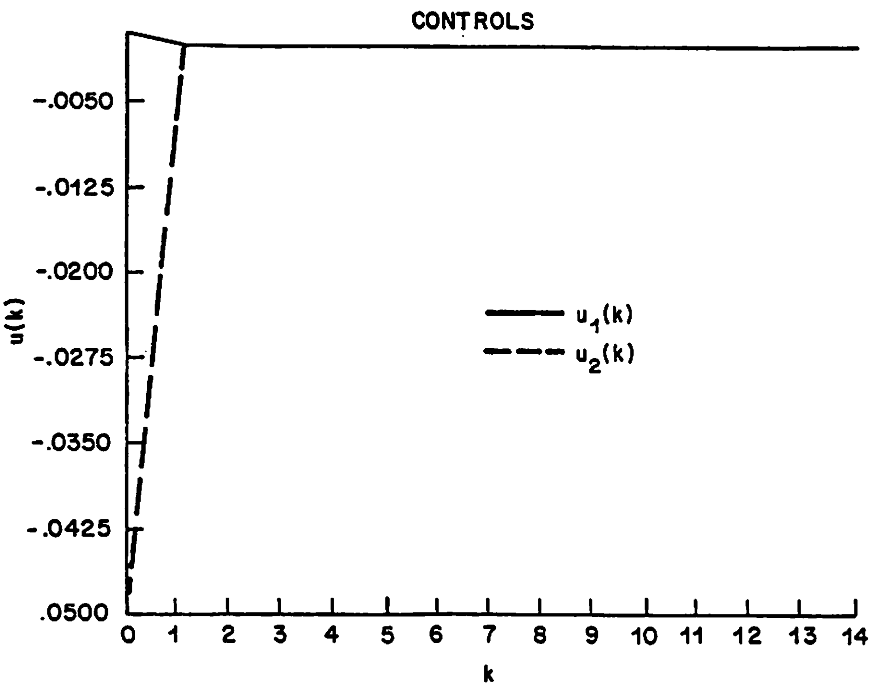

Figure 6. Control Variables vs. k for Example 2

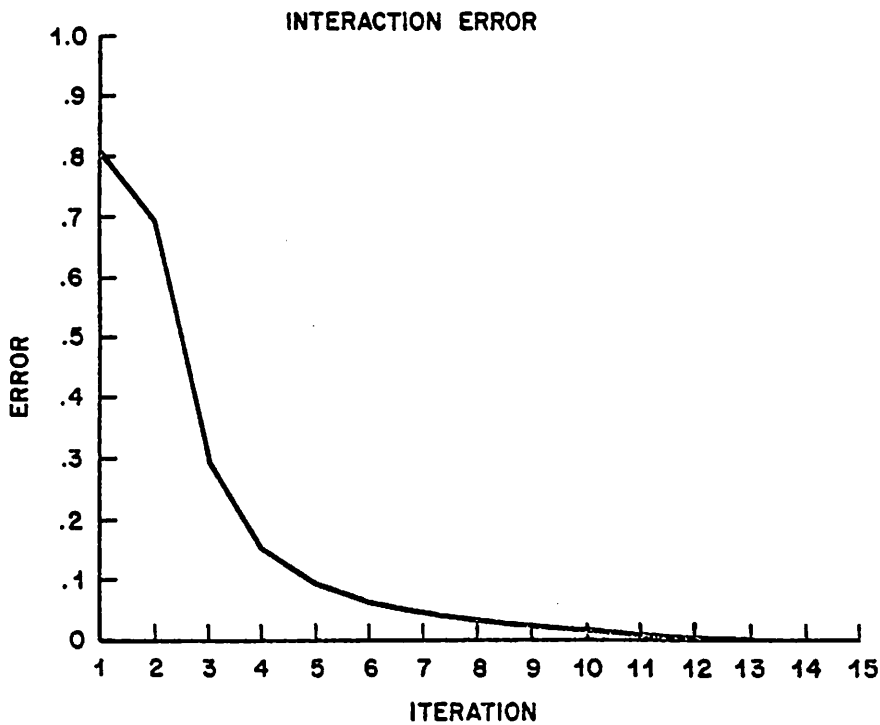

Figure 7. Interaction Error vs. Iteration for Example 2

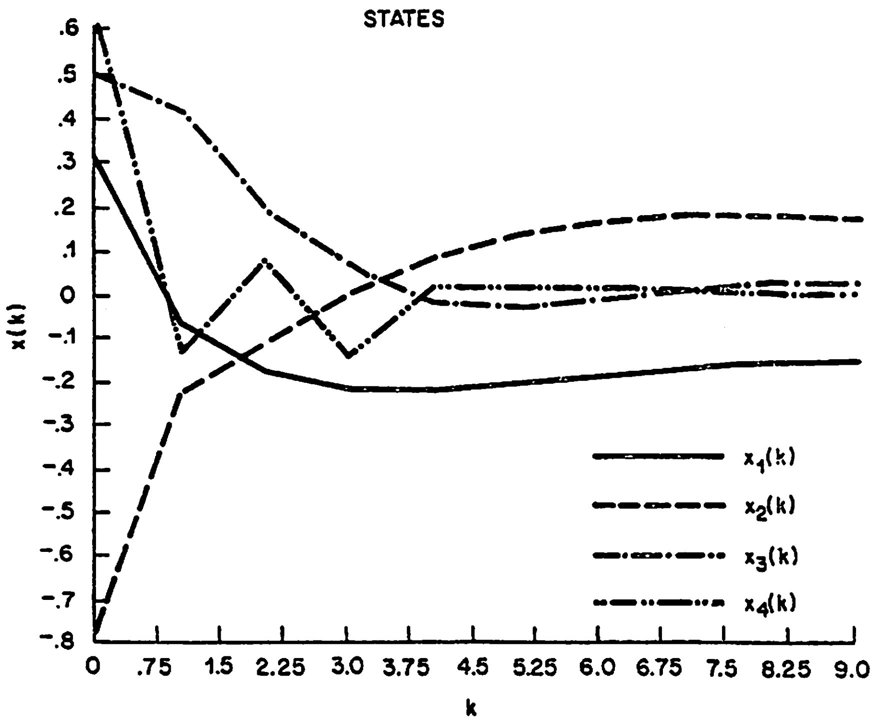

Figure 8. State Sequences for Example 3

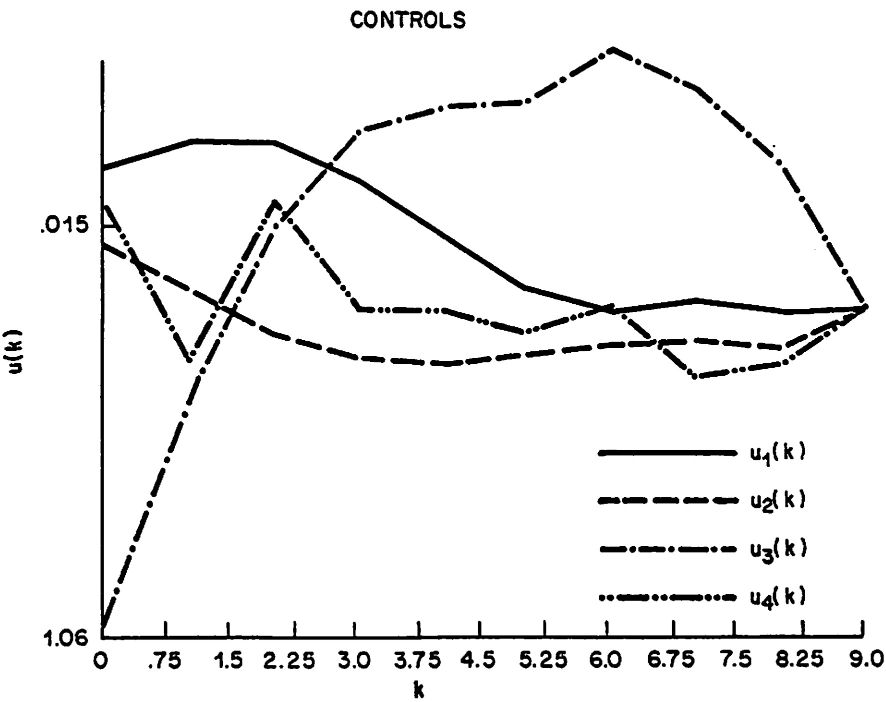

Figure 9. Control Sequences for Example 3

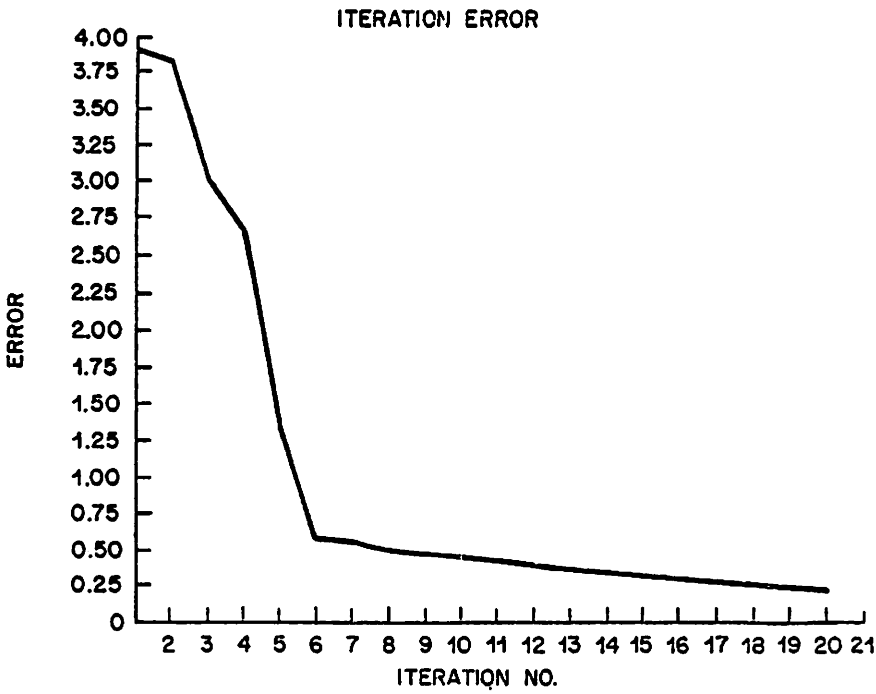

Figure 10. Interaction Error for Example 3

PROBLEMS

8.1 For a linear coupled TD system

$$\dot{x}_1 = -x_1 - 2x_2 - 0.2x_2(t-0.1) + u_1$$
$$\dot{x}_2 = -x_1 - x_2 + 0.1x_1(t-0.1) + u_2$$

and initial functions $x_1(t) = x_2(t) = 1,\ -0.1 \leq t \leq 0$ use Algorithm 8.2.6 to find a near-optimum control for the cost functional,

$$J = 1/2 \int_0^2 \left[x_1^2 + x_2^2 + 1/3u_1^2 + 1/3\, u_2^2\right] dt$$

8.2 Consider a nonlinear coupled TD system

$$\dot{x}_1 = -x_1 + 2x_1x_2 - x_1(t-0.2) + u_1$$
$$\dot{x}_2 = -x_2 - 2x_2x_1(t-0.2) + u_2 - u_2x_2(t-0.2)$$

with initial functions $x_1(t) = 1/2, x_2(t) = 1,\ -0.2 \leq t \leq 0.$ The cost function, to be minimized, is

$$J = 1/2 \int_0^1 \left[x_1^2 + x_2^2 + u_1^2 + u_2^2\right] dt$$

Use Algorithm 8.2.5 and your favorite computer language to find a near-optimum control.

8.3 Consider a 2-subsystem serial TD system defined by,

Subsystem 1:

$$\dot{\mathbf{x}}_1 = \begin{bmatrix} 0 & .1 \\ -1 & -2 \end{bmatrix} \mathbf{x}_1 + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u_1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ y_1 = [1\ \ 1]\, x_1$$

$$Q_1 = R_1 = F_1 = [1/2],\ t_o = 0, t_f = 1,$$
$$\mathbf{x}_1(0) = [-1\ \ 0]'$$

Subsystem 2:

$$\dot{\mathbf{x}}_2 = \begin{bmatrix} 0 & 1 \\ -.1 & -.2 \end{bmatrix} \mathbf{x}_2 + \begin{bmatrix} 1 & .2 \\ .1 & .5 \end{bmatrix} \mathbf{x}_1(t-0.1) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{u}_2 + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

$$Q_2 = F_2 = [1], \; \mathbf{y}_2 = [1 \quad 2]\mathbf{x}_2$$

$$\mathbf{x}_2(0) = [1 \quad -1]'$$

Use Algorithm 8.3.1 to find an optimal hierarchical control.

8.4 Consider a large-scale TD system described by

$$\dot{\mathbf{x}} = \left[\begin{array}{ccccccc} 0.1 & -1 & .2 & .01 & 0 & 0.1 & 0 \\ 0 & 1 & -.1 & 0.2 & 0 & 0 & 0 \\ 1 & -1 & .5 & .2 & 0.05 & 0.1 & 0.2 \\ \hline 0 & 0.1 & 0.1 & 0.1 & -1 & 0.0 & 0.1 \\ 0 & 0.2 & 0 & 0 & -2 & 0 & 0.1 \\ 0 & 0.1 & -0.1 & 0 & 0.01 & 0.5 & -2 \\ 0 & 0.1 & 0.2 & 0 & 0.01 & 0.2 & -1 \end{array}\right] \mathbf{x} +$$

$$\left[\begin{array}{cccccc} 0.2 & 1 & 0 & 0.1 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0.01 \\ \hline 0.1 & 0 & 0 & -0.1 & 0 & 0.1 \\ 0 & 0.1 & 0 & 0 & -.1 & 0.2 \\ 0 & 0 & -0.1 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & -0.1 \end{array}\right] \mathbf{x}(t-0.05) +$$

$$\left[\begin{array}{cc|ccc} 0.1 & 0 & 0 & 1 & 0 \\ 0 & -0.1 & 0 & 0 & 0 \\ -1.0 & 0 & -.1 & 0 & 0.2 \\ \hline 1 & 0.1 & 0.1 & 0 & 0 \\ 0 & 0.2 & 0 & -0.2 & 0 \\ 0 & 0 & 0 & 0 & -0.2 \\ 0 & -0.1 & -0.1 & -1 & -1 \end{array}\right] \mathbf{u}(t) +$$

$$\left[\begin{array}{ccccc} 1 & 0 & 0.05 & 0.5 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & .2 & 1 & 1 \\ \hline 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0.1 & 0.2 \\ 1 & 0.2 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & -0.1 \end{array}\right] \mathbf{u}(t-0.05)$$

with an objective function

$$J = 1/2\mathbf{x}'(1)\mathbf{x}(1) + 1/2 \int_0^1 (\mathbf{x}'\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})dt$$

and initial function $\mathbf{x}'(t) = [1\ 0\ 0\ 0\ 0\ 0\ 0]'$ use the method of Section 8.4 and appropriate computer program to find a near-optimum control.

8.5 Consider a nonlinear TD system

$$\begin{aligned} \dot{x}_1 &= -2x_1x_2 + x_2(t-0.1) + u_1 \\ \dot{x}_2 &= -x_2 - 2x_1x_2(t-0.1) + u_2 \end{aligned}$$

with initial function $x_1(t) = x_2(t) = 1,\ -0.1 \leqslant t \leqslant 0$ and cost function

$$J = 1/2 \int_0^1 \left[u_1^2 + u_2^2\right]dt$$

find a near-optimum control using the interaction prediction method of Section 8.5. Use an appropriate set of nominal trajectories.

8.6 Find a nonlinear discrete-time TD system

$$\begin{aligned} x_1(k+1) &= -2x_1^2(k) + 2x_2(k-1) + 0.1u_1(k) \\ x_2(k+1) &= -0.5x_2(k) + 0.25x_1(k)x_2(k-1) + u_2(k) \end{aligned}$$

with initial functions $x_1(k) = x_2(k) = 0$ and an objective function, to be maximized is

$$J = 1/2 \sum_{k=0}^{4} \left[x_1^2(k) + x_2^2(k) + u_1^2(k) + u_2^2(k)\right]$$

Find a near-optimum control using the costate prediction scheme of Section 8.6.

8.7 Repeat problem 8.5 for a linear discrete-time TD system,

$$\begin{aligned} \mathbf{x}(k+1) = &\begin{bmatrix} 1 & 0.1 \\ -1 & -1 \end{bmatrix}\mathbf{x}(k) + \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}\mathbf{x}(k-1) \\ &+ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{u}(k),\ \mathbf{x}(k) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \end{aligned}$$

$-1 \leqslant k \leqslant 0$. Use the same objective function as before.

8.8 Prove convergence of Algorithm 8.5.1.

8.9 Discretize the system of Problem 8.1 and represent the objective function as

$$J = 1/2 \sum_{k=0}^{k} \left[x_1^2(k) + x_2^2(k) + u_1^2(k) + u_2^2(k) \right]$$

Use Algorithm 8.6.1 to find a near-optimum control for it.

REFERENCES

[8.1] G. L. Kharatishivili, "The maximum principle in the theory of optimal processes with time lags," *Dokl. Akad. Nauk SSSR*, Vol. 136 , pp. 39-42, 1961

[8.2] N. N. Krasovskii, "Optimal processes in systems with time lag." *Proc. 2nd IFAC Congress,* Basel, Switzerland, pp. 327-332, 1963

[8.3] D. H. Eller, J. K. Aggarwal and H. T. Banks, "Optimal control of linear time-delay systems," *IEEE Trans. Aut. Cont.* AC-14, pp. 678-687, 1969

[8.4] J. K. Aggarwal, "Computational of optimal control for time-delay Systems," *IEEE Trans. Aut. Cont.* AC-15, pp. 683-685, 1970

[8.5] K. Inoue, H. Akashi, K. Ogino, and Y. Sawaragi "Sensitivity approaches to optimization of linear systems with time delay," *IFAC J. Automatica* Vol. 7, pp. 671-679, 1971

[8.6] P. Sannuti, "Near-optimum design of time-lag systems by singular perturbation method," *Proc. JACC, Atlanta, GA* Paper No. 20-A, 1970

[8.7] M. Jamshidi, and M. Malek-Zavarei, "Suboptimal design of linear control systems with time delay," *Proc. IEE,* Vol. 119, pp. 1743-1746, 1972

[8.8] H. C. Chan, and W. R. Perkins, "Optimization of time delay systems using parameter imbedding," *IFAC J. Automatica*, Vol. 9, pp. 257-261, 1973

[8.9] M. Jamshidi and M. Razzaghi, "Optimization of linear systems with input time-delay," *Kybernetika II*, 1975

[8.10] M. Jamshidi, Suboptimal control of coupled time-delay systems. *Int. J. Contr.* 17:995-1108, 1973

[8.11] P. V. Kokotovic, and G. Singh , "Optimization of coupled nonlinear systems," *Int. J. Contr.* Vol. 14 p. 51, 1971

[8.12] M. Malek-Zavarei, "Near-Optimum design of nonstationary linear systems with state and control delays," *J. Optimization Theory and Applications*, Vol. 30, pp. 73-88, 1980

[8.13] R. Bate *The optimal control of systems with transport lag.* in C. T. Leondes, ed., Advances in Control Systems, Vol. 7, Academic Press, 1969

[8.14] G. L. Slater and W. R. Wells, "On the reduction of optimal time-delay systems to ordinary ones," *IEEE Trans. Aut. Cont.* AC-17:154-155, 1972

[8.15] M. Jamshidi, "Three-stage near-optimum design of nonlinear-control processes, *Proc. IEE* 121:886-892, 1974

[8.16] M. Jamshidi, "A three-stage design of nonlinear control systems with time-delay" *Int. J. Contr.*, Vol. 21, pp. 753-762, 1975

[8.17] S. A. Gracovetsky and M. Vidyasagar Suboptimal control of neutral systems. *Int. J. Cont.*, Vol. 18, pp. 121-128, 1973

[8.18] M. Malek-Zavarei, "Suboptimal control of systems with multiple delays," *J. Optimization Theory and Applications*, Vol. 30, pp. 621-633, 1980

[8.19] B. D. O. Anderson, "Time delays in large-scale systems," *Proc. IEEE Conf. Dec. Control*, Ft. Lauderdale, FL, 1979

[8.20] M. G. Singh, *Dynamical Hierarchical Control.* Rev. Ed. North-Holland, Amsterdam, 1980

[8.21] M. Ikeda and D. D. Siljak, "Decentralized stabilization of large-scale systems with time delay," *J. Large Scale Systems* 1:273-279, 1981

[8.22] M. Jamshidi and M. Malek-Zavarei, "A hierarchical optimization method of large-scale time-delay systems," *Proc. IEEE Conf. Dec. Control*, San Diego, CA, pp. 1303-1304

[8.23] M. Jamshidi and M. Malek-Zavarei, "A hierarchical control for large-scale time-delay systems, "*J. Large-Scale Systems*, Vol. 4, pp. 149-163, 1983

[8.24] M. Jamshidi and C. M. Wang "A Computational Algorithm for Large-Scale Nonlinear Time-Delay Systems," *Proc. Int. Large-Scale Systems Symposium*, Virginia Beach, VA, pp. 258-266, October, 1982

[8.25] M. Jamshidi and C. M. Wang "Optimal Control of large-scale nonlinear systems with time-delay, " *IEEE Trans. Syst. Man Cyber.*, Vol. SMC-14, pp. 2-9, 1984

[8.26] M. Jamshidi and J. Merryman, "On the hierarchical optimization of retarded systems via costate prediction", *Proc. 1982 Amer. Control Conf.*, Arlington, VA, pp. 899-904, June 1982

[8.27] J. M. Breadeau and M. Jamshidi, "On the hierarchical control of discrete time systems with time delay" *J. Large-Scale Systems*, Vol. 7, pp. 33-46, 1984

[8.28] Jamshidi, M. A feedback near-optimum control for nonlinear systems. *Inform. Contr.* 32:75-84, 1976

[8.29] M. D. Mesarovic, D. Macko, Y. Takahara, "Two coordination principles and their applications in large-scale systems control", *Proc IFAC Congr., Warsaw*, Poland, 1969

[8.30] J. D. Pearson, "Dynamic decomposition techniques in Optimization methods for large scale systems", Wismer, D. A., Ed., McGraw-Hill Book Co, 1971

[8.31] M. Jamshidi, "*Large-Scale Systems, Modeling and Control*", Elsevier North-Holland Publishing Co., 1983

[8.32] A. M. Geoffrion, "Elements of large-scale mathematical programming: PartI, Concepts, PartII, Synthesis of Algorithms and bibliography". *Management Sci.* 16. 1970

[8.33] A. M. Geoffrion, "*Large-scale linear and nonlinear programming in Optimization Methods for Large-Scale Systems*", D. A. Wismer, Ed., McGraw-Hill Book Co., 1971

[8.34] Y. Takahara, "A Multi-level structure for a class of dynamical optimization problems", *M. S. Thesis*, Case Western Reserve University, Cleveland, Ohio, 1965

[8.35] M. Mahmoud, W. Vogt and M. Mickle, "Multilevel Control and Optimization using Generalized Gradients Technique, "*Int J. Control*, Vol. 25, pp. 525-543, 1977

[8.36] Hassan, M. and Singh, M. G. "A Two-Level Costate Prediction Algorithm for Non-Linear Systems," *IFAC J. Automatica*, 13: pp. 629-634, 1977

[8.37] A. Newhouse, "Chebyschev curve fit.", Comm ACM. 5: 281., 1962.

[8.38] M. Jamshidi and R. Banning, "LSSPAK/PC - A CAD Package for Large-Scale Control Systems", *Tech. Report No. LCAD 85-01*, CAD Laboratory for Systems/Robotics, EECE Dept., UNM, Albuquerque, NM, January, 1985.

[8.39] A. E. Taylor and D. C. Lay, *Introduction to Functional analysis*, 2nd ed., John Wiley and Sons, Inc., 1980

[8.40] R. J. Peters, K. C. Chu and M. Jamshidi, Optimal Operation of a Water Resources System by Stochastic Programming," *Math. Programming*, Study 9, 1978

[8.41] D. A. Chyung, "Discrete Optimal Systems with Time-Delay," *IEEE Trans. Auto. Control*, p. 117, 1967

Part IV APPLICATIONS

CHAPTER 9

SOME APPLICATIONS OF TIME-DELAY SYSTEMS

9.1 INTRODUCTION

In this chapter we will discuss some industrial applications of TD systems. The first application deals with the tension and thickness control of a steel strip passing through a multi-stand cold rolling mill. This application has already been discussed in Sections 1.3 and 8.5 without presenting a formal solution for it. An existing 3-stand system will be discussed. The second application area, discussed in Section 9.3, is urban traffic control. The hierarchical control and a time-delay algorithm are used to treat this class of problems. The application of a hierarchical control algorithm for discrete-time systems with distributed delays to traffic control problems is used here. The third application is a multi-reservoir water-resources system which was first introduced in Section 1.3. Here an existing 3-reservoir water-resources system will be presented and a solution for it will be given. The chapter concludes by looking into a hydraulic control system in Section 9.5. The celebrated Smith Predictor and an adaptive algorithm for it are presented in Section 9.6.

9.2 COLD ROLLING MILLS

In this Section a near-optimum design and computer control algorithm is presented for a large-scale cold rolling mill. Such mills are mathematically represented by systems of a large number (thirty to seventy) of nonlinear coupled scalar differential-difference equations. Due to the complexities of these models, relatively little has been done in the way of applying optimal control theory to rolling mills [9.1-9.5]. Among these attempts the paper by Smith [9.5] seems to be a good starting point. However, the model used is not complete since it does not include the dynamics and nonlinearities due to work roll drives, winding and unwinding reels.

The model used in this section takes into account the dynamics of winding, unwinding reels; pay-off and delivery bridles; screw-down mechanism and work roll drives. To simplify the design, the initially nonlinear model is transformed into a parametric family of linear models. This family is generated by 'slow' variations of winding reel radius and moment of inertia. In the first stage of the design the 'slow', 'basic', and 'fast' states of the system are identified and the *three-time-scale* concept is used to eliminate the slow and fast states. Thus the dimensionality of the model is reduced from $(2k+1)(n_s+n_b+n_f)$ to n_b , where k is the number of first-order elements used to approximate an interstand time delay. In the second stage, the reduced model is approached as a set of weakly coupled subsystems convenient for the 'ϵ-coupling' procedure [9.7] (see Section 8.2). Then the state regulator design is applied to the whole family of models depending on winding real radius r. The concept of 'parameter embedding' [9.8, 9.9] is used to avoid repeated solutions of the accompanied matrix Riccati equation. A near-optimum feedback gain matrix is obtained as a function of r which is easily programmable on a process control computer. While the developed control program is general, the numerical evaluation of the program and the simulation tests are performed using the dfrom an Allegheny Ludlum's three-stand high-speed mill [9.10]. The numerical data of this mill may be obtained in Reference [9.11].

9.2.1 The Rolling Process

A schematic of an N-stand cold mill is given in Figure 1 which shows the decoiler (pay-off reel and bridles), coiler (delivery bridles and reel) as well as the three 'mill housing', *i.e.*, work rolls, back-up rolls and screw-down mechanisms. The strip is fed from the pay-off reel through the bridles to the first stand where its thickness is reduced by plastic deformation due to a combination of interstand tension and roll separating force. The strip then passes through the remaining stands, delivery bridles and is wound on the winding reel. The work rolls drive motors are assumed to be both armature and field-controlled, while the screw-down motors are only armature-controlled.

There are many rolling theories [9.12-9.14]. However, due to its computationally convenient expression of rolling force, torque and strip thickness in terms of the mill variables, the theory of Bland and Ford [9.15] is used:

$$\begin{bmatrix} \hat{F} \\ \hat{\tau} \\ \hat{h}_n \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & k_3 & k_4 & k_5 \\ l_1 & l_2 & l_3 & l_4 & l_5 \\ m_1 & m_2 & m_3 & m_4 & m_5 \end{bmatrix} \begin{bmatrix} \hat{h}_i \\ \hat{h}_o \\ \tilde{t}_i \\ \tilde{t}_o \\ \zeta \end{bmatrix} \tag{1}$$

where the variables are defined as follows:

i : 'entry' or inlet plane

$\hat{}$: $\hat{F} = \dfrac{F - F}{F^*}$

o : 'exit', or outlet plane

$\sim$: $\tilde{f} = \dfrac{t - t^*}{\text{area}}$

* denotes the operating value

n : neutral plane

t : tension per unit area, *p.s.i.*

r : roll torque per unit width, *lb.-ft./in.*

F : roll force per unit width, *lb./in.*

ζ : coefficient of friction

h : thickness of the strip, inches

Figure 2 shows the mill variables associated with each stand. Common assumptions in all two-dimensional rolling theories is that the density and width of the rolled material remain constant throughout the process. As a consequence of these assumptions, the volume of the material per unit time passing through each vertical segment is constant:

$$v_i h_i = \cdots = v_n h_n = \cdots = v_o h_o \tag{2}$$

where v is the strip velocity. The linearized form of (2) reduces to

$$\hat{v}_i = \hat{v}_n + \hat{h}_n - \hat{h}_i, \tag{3}$$

$$\hat{v}_o = \hat{v}_n + \hat{h}_n - \hat{h}_o. \tag{4}$$

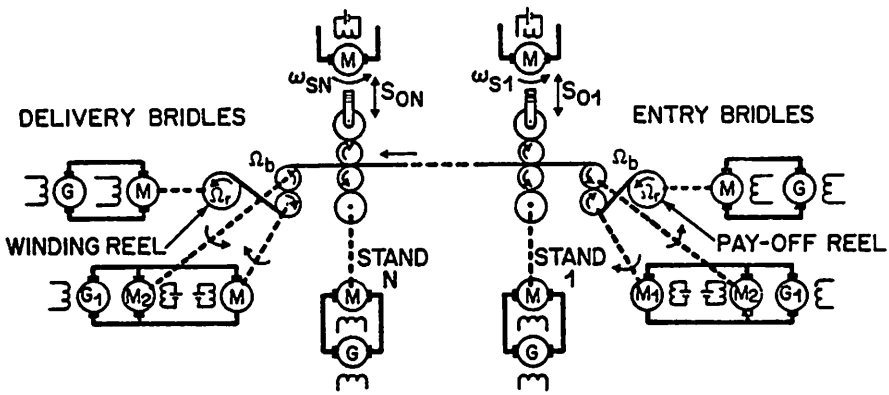

Figure 1. An N-stand Cold-Rolling Mill

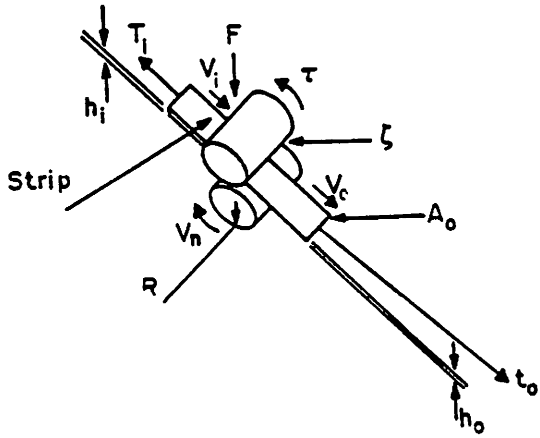

Figure 2. Mill Stand Parameters

The output tension and thickness at the *jth* stand is assumed to be related to the input quantities of the $(j+1)st$ stand by

$$t_{i,j+1}(t) = t_{oj}(t) \tag{5}$$

$$h_{i,j+1}(t) = h_{o,j}(t-\lambda). \tag{6}$$

In (5) and (6) j represents the number of stand, $j=1,2,\ldots,N$; N is the maximum number of stands in the mill and λ is the time delay (strip transit time) in seconds.

The interstand strip is usually described by Hook's elasticity law in its time-derivative form [9.1, 9.5]:

$$\dot{t}_{oj} = \frac{E}{T}\,(\hat{v}_{ij+1} - \hat{v}_j), \tag{7}$$

where E is the modulus of elasticity of the material and λ is an average value of the time delay assumed to be constant. The elastic stretch of the mill housing, relating output thickness h_o to the screw-down pitches (positive-upward position) and the rolling force F is

$$\hat{h}^{*}_{oj} = \left\{\frac{1}{h^{*}_{oj}}\right\} s_{oJ} + \left\{\frac{F^{*}_{j}}{K_M h^{*}_{oj}}\right\} \hat{F}_j. \tag{8}$$

where K_M is the modulus of elasticity of the mill housing, s_{oJ} represents screw-down motors setting in inches, K_M represents modulus of elasticity (spring constant) of "mill housing" in *lb./in.* and an asterisk * stands for operating value of a quantity.

The work rolls drive motors must provide torque sufficient to balance friction, acceleration, inlet tension and rolling torques. The torque caused by outlet tension helps the rolls rotate in the prescribed direction:

$$\tau_{mj} = J\dot{\omega}_j + B\omega_j + \frac{1}{n} RT_{ij} - \frac{1}{n} Rt_{oj} + \tau_j \tag{9}$$

where J, B and R represent moment of inertia $(lb-ft\, sec^2/rad)$, friction loss $(lb-ft-sec/rad)$ and the roll's radius (ft). Similar equations can be written for torque relation at the shaft of the bridles' rolls drive and screw-down motors. However, the torque relation for the winding reel has a term due to variations of the reel's moment of inertia, *i.e.*,

$$\sum_i \tau_i = \tau_r = b\omega + J\dot{\omega} + \dot{J}(\Omega-\Omega_b) + \frac{exp(-2\zeta a)}{n} \mathrm{rt_o}, \tag{10}$$

where

$$\dot{J} = \frac{dJ(r)}{dr}\dot{r} = \frac{d(r)}{dr}\frac{nh}{2\pi}\omega \tag{11}$$

Ω is angular velocity $(rad/sec.)$, n is the gear ratio a is the bridles' rolls wrap angle and r is the slowly varying winding reel's radius.

9.2.2 A State-Space Model

The above relations are combined in a state equation model for the entire process. Table 1 shows a possible selection of the state and control variables. Note that the output tension of the third stand and that of the coiler are the same. The variables $\hat{F}$, $\hat{t}_i$, and $\hat{h}_n$ do not appear as state variables since it is possible to express them in terms of $\hat{h}_o$, $\hat{t}_o$ and $\hat{h}_o(t-\lambda)$ from (1), (8), (5) and (6). Thus the following parametric family of state space models with multiple delays is obtained:

$$\dot{\mathbf{X}} = \mathbf{A}(r)\mathbf{X} + \mathbf{B}(r)\mathbf{U} + \mathbf{A}_d\mathbf{X}(t-\lambda) + \mathbf{A}_{2d}\mathbf{X}(t-2\lambda) \tag{12}$$

TABLE 1. A choice of state and control vectors

Vector		Coiler and decoiler ($j=d,c$)	Stand j ($j=1,2,3$)	Total
State	Physical	$\hat{\imath}_{f0j}\hat{\imath}_{f1j}\hat{\imath}_{fj}\,\hat{\Omega}_j\,\hat{\Omega}_j\hat{\imath}_{t1}\hat{\imath}_{a1j}\hat{\imath}_{a2j}$	$\hat{\imath}_{aj}\ \hat{s}_j\ \hat{v}_{sj}\ \hat{\imath}_{f1j}\ \hat{\imath}_{2j2}\hat{\Omega}_j\ \bar{t}_{0j}\ \hat{\imath}_{asj}\hat{\imath}_{a1j}\hat{\imath}_{a2j}$	$n=44$
	Mathematical	$x_1\ x_2 x_3\ x_4\ x_5\ x_6\ x_7\ x_8$	$x_9\ x_{10}\ x_{11}\ x_{12}\ x_{13}\ x_{14}\ x_{15}\ x_{16}\ x_{17}\ x_{18}$	
Control	Physical	$\hat{e}_{0j}\quad \hat{e}_{1j}\quad \hat{e}_j$	$\hat{e}_{f1j}\quad \hat{e}_{f2j}\quad \hat{e}_{asj}\quad \hat{e}_{a1j}\quad \hat{e}_{a2j}$	$m=21$
	Mathematical	$u_1\quad u_2\quad u_3$	$u_4\quad u_5\quad u_6\quad u_7\quad u_8$	

with an initial function

$$\mathbf{X}(t) = \eta(t),\ t_0 - 2\lambda \leqslant t \leqslant t_o. \tag{13}$$

where matrices $\mathbf{A}$, $\mathbf{A}_d$, $\mathbf{A}_{2d}$ are $n\times n$, and $\mathbf{B}$ is $n\times m$. The expressions for $\mathbf{A}$, $\mathbf{A}_d$, $\mathbf{A}_{2d}$ and $\mathbf{B}$ as well as the detailed derivation of (12) can be found in Reference [9.11]. For each fixed value of r, (12) represents a linear time-delay system. Thus, nonlinearities due to r are interpreted as slow parameter variations.

9.2.3 Optimal Control Design

The optimal control design approach taken here follows the singularly-perturbed scheme of Section 7.4 which has been independently treated by Sannuti [9.16] and Inoue et. al. [9.17]. Let the time delay λ be divided into k equal subintervals and introduce k new state vectors. Thus (12) and (13) may be approximated by the following system of vector differential equations:

$$\dot{\mathbf{X}} = \mathbf{A}(r)\mathbf{X} + \mathbf{B}(r)\mathbf{U} + \mathbf{A}_d + \mathbf{Y}_k + \mathbf{A}_{2d}\mathbf{W}_k,$$

$$\frac{\lambda}{k}\,\dot{\mathbf{Y}}_1 = \mathbf{X} - \mathbf{Y}_1,$$

$$\vdots$$

$$\frac{\lambda}{k}\,\dot{\mathbf{Y}}_i = \mathbf{Y}_{i-1} - \mathbf{Y}_i, \tag{14}$$

$$\vdots$$

$$\lambda_k\,\dot{\mathbf{W}}_i = \mathbf{Y}_i - \mathbf{W}_i,\ i=1,\ldots,k,$$

with initial conditions

$$\mathbf{X}(t_0) = \eta(t_0),\ \mathbf{Y}_i(t_o) = \eta\left\{t_0 - \frac{i\lambda}{k}\right\},\ \mathbf{W}_i(t_0) = \eta\left\{t_0 - \frac{2i\lambda}{k}\right\}. \tag{15}$$

System (14) presents $M = (2k+1)n$ ordinary differential equations. For example, if $k=10$ and $n=44$, then $M=924$. In an attempt to solve (14) by classical optimal control theory one would be confronted with a problem of solving an $M\times M$ symmetric matrix Riccati equation equivalent to a system $M(M+1)/_2=427350$ scalar equations.

Considering Table 1 and the mill data [9.11] it is observed that the variables associated with the screw-down mechanism and some flux equations form a 'slow' state and those associated with motors armature circuits and time delays form a 'fast' state. The remaining variables such as angular velocities or tensions which do not fall under either of the above two categories are considered as the 'basic' state. Thus we divide **X** into slow, basic, and fast states, $\mathbf{X}' = [\mathbf{s} \vdots \mathbf{x} \vdots \mathbf{z}]'$. Then model (14) may be rewritten in the following partitioned form:

$$\dot{s} = u(e_0 s + c_o z + B_o u), \tag{16a}$$

$$\dot{x} = e_1 s + A_0(r)x + c_1(r)z + C_2 Z_k + B_1(r)u, \tag{16b}$$

$$\mu\dot{z} = e_2 s + A_1 z + c_3 z + B_2 u, \tag{16c}$$

$$\mu\dot{z} = A_2 X + C_4 Z, \tag{16d}$$

where $Z = [Y : W] = [Y_1, \ldots, Y_k : W_1, \ldots, W_k]'$, $Z_k = [Y_k, W_k]^1$ and $\mu = \lambda/k$. The coefficient matrices A_i, B_i, C_i, C_4 and B_i are the partitioned form of A, A_d and A_{2d},

$$A(r) = \begin{bmatrix} \mu c_0 \\ \cdots\cdots \\ e_l \\ \cdots\cdots \\ e_2/\mu \end{bmatrix} \begin{matrix} 0 \\ \cdots\cdots \\ A_0(r) \\ \cdots\cdots \\ A_1/\mu \end{matrix} \begin{bmatrix} C_0 \\ \cdots\cdots \\ C_1(r) \\ \cdots\cdots \\ C_3/\mu \end{bmatrix} = \left[\begin{array}{c|c} A_1(r) & \epsilon A_{12}(r) \\ \cdots\cdots & \cdots\cdots \\ \epsilon A_{21}(r) & A_2(r) \end{array}\right]$$

$$B(r) = \begin{bmatrix} \mu B_0 \\ \cdots\cdots \\ B_1(r) \\ \cdots\cdots \\ B_2/\mu \end{bmatrix} \tag{17}$$

$$
\mathbf{A}_2 = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \\ \\ \\ \\ \cdot \\ \cdot \\ \cdot \\ \\ \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad
\mathbf{C}_4 = \left[\begin{array}{cccccc|cccc}
-\mathbf{I} & & & & & \mathbf{0} & & & & \\
\mathbf{I} & -\mathbf{I} & & & & \vdots & & & & \\
\mathbf{0} & \mathbf{I} & -\mathbf{I} & & & & & & \mathbf{0} & \\
\vdots & & & \mathbf{I} & -\mathbf{I} & \mathbf{0} & & & & \\
\mathbf{0} & & & & \mathbf{I} & -\mathbf{I} & & & & \\
\hline
\mathbf{I} & & & & & & -\mathbf{I} & & & \\
 & \mathbf{I} & \mathbf{0} & & & & & -\mathbf{I} & \mathbf{0} & \\
 & & \cdot & & & & & & \cdot & \\
 & & & \cdot & & & & & \cdot & \\
 & & \mathbf{0} & & \cdot & & & \mathbf{0} & & \cdot \\
 & & & & & \mathbf{I} & & & & -\mathbf{I}
\end{array}\right]
\tag{18}
$$

where $\mathbf{I} = \mathbf{I}_n$ is the identity matrix. The initial condition corresponding to (16) is

$$
\begin{aligned}
\mathbf{X}'(0) &= [\mathbf{s}(0) \vdots \mathbf{x}(0) \vdots \mathbf{z}(0)]' = [0 \vdots \mathbf{x}_0 \vdots \mathbf{z}_0']', \\
\mathbf{Z}'(0) &= [\eta(-i\mu) \vdots \eta(-2i\mu)], \; i = 1, \ldots, k.
\end{aligned}
\tag{19}
$$

It is emphasized that fast modes of system (16), *i.e.*, (16c) and (16d), can be interpreted as a 'singular perturbation' which has been treated in Sannuti and Kokotovic' [9.18].

By a comparison of the time constants corresponding to each of the state variables, the vectors s, x and z are identified as follows:

$$
\begin{aligned}
\mathbf{s} &= [\hat{\mathrm{s}}_j \hat{v}_{sj} \hat{\iota}_{f1j} \hat{\iota}_{f2j}], \\
\mathbf{x} &= [\hat{\iota}_{f0i} \hat{\iota}_{f1i} \hat{\iota}_{fi} \; \Delta\hat{\ }_i \bar{t}_{ii} \vdots \Delta\hat{\ }_j \bar{t}_{oj}] \\
\mathbf{z} &= [\hat{\iota}_{a1i} \hat{\iota}_{a2i} \hat{\iota}_{ai} \vdots \hat{\iota}_{asj} \iota_{a1j} \hat{\iota}_{a2j}], \; j=1,2,3 \text{ and } i=c,d.
\end{aligned}
\tag{20}
$$

Note that (20) represents $n_s = 12$ slow, $n_b = 17$ basic and $n_f = 15$ fast state vectors, respectively. This 'time scale decomposition' [9.6] represents the first stage of the design.

Considering (19) and setting $\mu \to 0$, the high dimensional model (16) is reduced to

$$\dot{s} = 0, \tag{21a}$$

$$\dot{x} = e_1 s + A_0(r)x + C_1(r)z + C_2 Z_k + B_1(r)u, \tag{21b}$$

$$0 = e_2 s + A_1 x + C_3 z + B_2 u, \tag{21c}$$

$$0 = A_2 X + C_4 Z. \tag{21d}$$

Eliminating z and Z from algebraic equations (21c), (21d) in terms of x and u, respectively, *i.e.*

$$z = -C_3^{-1}(A_1 x + B_2 u),\ Z = -C_4^{-1} A_2 X = [X, \ldots, X : X, \ldots, X], \tag{22}$$

$$Z_k = [X \vdots X] = [s, x, z \vdots s, x, z],$$

(21b) becomes

$$\dot{x} = A(r)x + B(r)u, \tag{23}$$

where $A(r)$ and $B(r)$ are numerically defined by

$$A(r) = A_o(r) + C_2 - C_1(r)C_3{}^{-1}A_1 \tag{24a}$$

$$
\mathbf{A}(r) = \left[\begin{array}{cccc|ccccc}
-0.318 & & & & & & & & \\
 & -0.318 & & & & & & & \\
 & -0.33 & & & & & & & \\
 & & -0.11 & -0.532 & & & & & \\
a_{51} & a_{53} & a_{55} & a_{56} & & & & \epsilon \mathbf{A}_{12} & \\
 & 23.4a_{65} & -0.805 & -0.805 & & & & & \\
 & & -7.65 & -19.0 & & & & & \\
 & & 0.173 & -0.1195 & & & & & \\
\hline
 & & & & -19.0 & & & & \\
 & & & & -0.111 & & & -0.0216 & \\
 & & & & -4.3 & -18.8 & & & 6.53 \\
 & & & & & -0.318 & & & \\
 & \epsilon \mathbf{A}_{21} & & & & & & -0.318 & \\
 & & & & & & & -0.33 & \\
 & & & & & & 2.21 & -0.11 & -0.205 \\
 & & & & & a_{16,12} & a_{16,14} & a_{16,16} & a_{16,17} \\
 & & & & & -0.973 & & & -0.19a_{17,18}
\end{array}\right] \tag{24b}
$$

where $a_{52}(r) = N_1(r)\, D_1(r)$, $a_{53}(r) = 0.1a_{51}(r)$, $a_{55}(r) = -0.7a_{51}(r)$, $a_{65}(r) = N_2(r)/D_1(r)$, $a_{65}(r) = 6.6r$, $a_{16,12}(r) = N_1(r)/D_2(r)$, $a_{16,14}(r) = 0.1\ a_{16,12}(r)$, $a_{16,17}(r) = N_2(r)/D_2(r)$, and $a_{17,16}(r) = 7.96r$ with $N_1(r)N_1(r) = 1.98(r+0.152r^3)$, $D_1(r) = 1+0.0185r^4$,

$$
\mathbf{B}(r) = \mathbf{B}_1(r) + \mathbf{C}_1(r)\mathbf{B}_2 : \tag{25a}
$$

$$
\mathbf{B}(r) = \left[\begin{array}{c|c} \mathbf{B}_1(r) & 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & \mathbf{B}_2(r) \end{array}\right]
$$

$$
= \left[\begin{array}{ccc|ccccccc}
0.318 & & & & & & & & & \\
-0.318 & & & & & & & & & \\
 & b_{33}(r) & & & & & & & & \\
 & & & & & & 0 & & & \\
 & & -1.22 & -1.22 & & & & & & \\
 & & & & -0.826 & -0.826 & & & & \\
\hline
0 & & & & & -0.424 & -0.424 & & & \\
 & & & & & & 0.318 & & & \\
 & & & & & & & 0.318 & & \\
 & & & & & & & & & b_{14,12}(r) \\
 & & & & & & & & 0 & \\
 & & & & & & & & 0 & \\
 & & & & & & & & 0 &
\end{array}\right] \qquad (25b)
$$

where $b_{33}(r) = 0.105/r(1.5.25r^2))$, $b_{14,12}(r) = 0.105/(r(1-8.78r^2))$. Note that $n_1 = 8$ $n_2 = 9$, $m_1 = 5$, $m_2 = 7$, $\epsilon = 0.50$,

$$
\mathbf{A}_{12} = \begin{bmatrix} & 0 & \\ & & \\ 0{\cdot}0448 & \cdots & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{A}_{21} = \begin{bmatrix} & -7.9 \\ & 0.082 \\ & 0 \\ 0 & \cdot \\ & \cdot \\ & \cdot \\ & 0 \end{bmatrix} \qquad (25c)
$$

In this way the 924-dimensional model (16) is reduced to $n_b = 17$-dimensional model (23) which represents a parametric family of non-delayed linear differential equations.

Recall that the state variables in the model (23) are deviations of tensions, velocities and some of the motor field currents from their nominal values (Table 1). Our aim is to design a regulator which will keep these variables as close to zero as possible, without using too large values of control voltages. A convenient way to express this objective mathematically is to introduce a quadratic cost functional

$$J(\mathbf{u}) = \int_{t_o}^{t_f} (\mathbf{x}'Q\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})\, dt, \tag{26}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are positive definite weighting matrices. Then our design problem can be formulated as follows.

For each value of $r \epsilon [r_o, r_f]$ determine a feedback control $\mathbf{u}$ which minimizes the cost functional (24), while $\mathbf{x}$, $\mathbf{u}$ and r are related by (23). If an attempt is made to solve this problem by a standard Riccati equation procedure, the amount of computation is prohibitively large. Inspection of the system matrices $\mathbf{A}(r)$ and $\mathbf{B}(r)$ in (24)-(25) motivates us to apply the 'ϵ-coupling' procedure to reduce part of the computational difficulties. The other part of these difficulties will be reduced by the 'parameter embedding' method [9.20].

9.2.4 Computational Procedure

The system matrices $\mathbf{A}(r)$ and $\mathbf{B}(r)$ are expressed in terms of a coupling parameter ϵ while $\mathbf{Q}$ and $\mathbf{R}$ are assumed to be block-diagonal:

$$\mathbf{A}(r,e) = \left[\begin{array}{c|c} \mathbf{A}_1(r) & \epsilon\mathbf{A}_{12} \\ \cdots\cdots & \cdots\cdots \\ \epsilon\mathbf{A}_{21} & \mathbf{A}_2(r) \end{array}\right], \quad \mathbf{B}(r) = \left[\begin{array}{c|c} \mathbf{B}_1(r) & 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & \mathbf{B}_2(r) \end{array}\right],$$

$$\mathbf{Q} = \left[\begin{array}{c|c} \mathbf{Q}_1 & 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & \mathbf{Q}_2 \end{array}\right], \quad \mathbf{R} = \left[\begin{array}{c|c} \mathbf{R}_1 & 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & \mathbf{R}_2 \end{array}\right], \tag{27}$$

where $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_{12}$, $\mathbf{A}_{12}$, $\mathbf{B}_1$, and $\mathbf{B}_2$ are $n_1 \times n_1$, $n_2 \times n_2$, $n_1 \times n_2$, $n_2 \times n_1$, $n_1 \times m_1$ and $n_2 \times m_2$

matrices, respectively. The submatrices A_1 to B_2 are defined in (24)-(25) and $\epsilon=0.5$ is a fixed value throughout. The first subsystem consists of pay-off reel, bridles and stand one while the stands two, three delivery bridles and widing reel constitute the second subsystem. Then for every fixed ϵ the optimum control $\mathbf{u}$ for the above problem is [9.21]

$$\mathbf{u} = - \mathbf{R}^{-1}\mathbf{B}'\mathbf{K}(r)\mathbf{x} = - \mathbf{F}(r)\mathbf{x}, \tag{28}$$

where $\mathbf{K}(r)$ is the steady-state solution of the matrix Riccati equation

$$\dot{\mathbf{K}} = -\mathbf{K}\mathbf{A} - \mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{S}\mathbf{K} - \mathbf{Q}, \;\; \mathbf{K}(t_1,r,\epsilon) = 0, \tag{29}$$

where $\mathbf{S} = \mathbf{B}\mathbf{R}^{-1}\mathbf{B}'$.

Using the 'ϵ-coupling' technique we avoid solving the system of $n(n+1)/2$ nonlinear algebraic equations [9.19]:

$$\mathbf{K}\mathbf{A} + \mathbf{A}'\mathbf{K} - \mathbf{K}\mathbf{S}\mathbf{K} + \mathbf{Q} = 0 \tag{30}$$

by approximating its solution by a truncated MacLaurin series in ϵ,

$$\mathbf{M} = \mathbf{K}(t,r,0) + \epsilon\mathbf{K}^1(t,r,0) + \cdots + \frac{\epsilon^m}{m!}\mathbf{K}^m(t,r,0). \tag{31}$$

where

$$\mathbf{K}^i(t,r,0) \triangleq \frac{\delta^i \mathbf{K}(t,r,\epsilon)}{\delta\epsilon^i}\Bigg|_{\epsilon=0}$$

If we partition $\mathbf{K}$,

$$\mathbf{K} = \left[\begin{array}{c|c} \mathbf{K}_1(t,r) & \epsilon\mathbf{K}_{12}(t,r) \\ \cdots\cdots & \cdots\cdots \\ \epsilon\mathbf{K}'_{12}(t,r) & \mathbf{K}_2(t,r) \end{array}\right] \tag{32}$$

and substitute it into (31), three truncated series in ϵ will be obtained for $\mathbf{K}_1$, $\mathbf{K}_{12}$ and $\mathbf{K}_2$. It has been shown that the even-order partials of $\mathbf{K}$ are diagonal and odd-order partials are anti-diagonal, *i.e.*, [9.19].

$$\mathbf{K}^{2j} = \left[\begin{array}{c|c} \mathbf{K}_1^{2j} & 0 \\ \hline 0 & -\mathbf{K}_2^{2j} \end{array}\right] \tag{33}$$

and

$$\mathbf{K}^{2j+1} = \left[\begin{array}{c|c} 0 & -\mathbf{K}_{12}^{2j+1} \\ \hline (\mathbf{K}_{12}^{2j+1} & 0 \end{array}\right], \; j=0,1,\ldots, \tag{34}$$

Substituting **M** of (31) into (28) for **K**, noting (32)-(34) and equating the coefficients of ϵ^2, $m=0,1,\ldots,$ we obtain that for $m=0$, (33) is the solution of

$$\mathbf{K}_i^0(r)\mathbf{A}_i(r) + \mathbf{A}'_i(r)\mathbf{K}_i^0(r) - \mathbf{K}_i^0(r)\mathbf{S}_i(r)\mathbf{K}_i^0(r) + \mathbf{Q}_i = 0, \; i = 1,2. \tag{35}$$

For $m=1$, (32) is the solution of

$$\mathbf{K}_{12}^1(r)\mathbf{G}_2(r) + \mathbf{G}'_1(r)\mathbf{K}_{12}^1(r) = -\mathbf{F}_{12}^0(r), \tag{36a}$$

where

$$\mathbf{G}_i(r) = \mathbf{A}_i(r) - \mathbf{S}_i\mathbf{K}_i, \; i = 1,2$$

$$\mathbf{F}_{12}^0 = \mathbf{A}'_{12}\mathbf{K}_2^0(r) + \mathbf{K}_1^0(r)\mathbf{A}_{12}. \tag{36b}$$

Note that $\mathbf{K}_{21}^1 = (\mathbf{K}_{12}^1)'$. For $m=2$, (33) is the solution of

$$\mathbf{K}_i^2(r)\mathbf{G}_i(r) + \mathbf{G}'_i(r)\mathbf{K}_i^2(r) = -\mathbf{F}_i^1(r), \; i = 1,2, \tag{37a}$$

where

$$\mathbf{F}_1^1 = 2(\mathbf{A}_{21}-\mathbf{S}_2\mathbf{K}_{21}^1)'\mathbf{K}_{21}^1 + 2\mathbf{K}_{12}^1\mathbf{A}_{21}, \tag{37b}$$

$$\mathbf{F}_2^1 = 2(\mathbf{A}_{12} - \mathbf{S}_1\mathbf{K}_{12}^1)'\mathbf{K}_{12}^1 + 2\mathbf{K}_{21}^1\mathbf{A}_{12}. \tag{37c}$$

For $m=3$ and 4 matrix equations similar to (36a) and (37a) are obtained except that the forcing terms would now be $\mathbf{F}_{12}^1(r)$ and $\mathbf{F}_i^2(r)$, respectively. This procedure can be repeated to as many terms as necessary for a prescribed accuracy. In this paper $m=2$; then the resulting approximate Riccati matrix becomes

$$\mathbf{M}(r) = \left[\begin{array}{c|c} \mathbf{M}_1(r) & \mathbf{M}_{12}(r) \\ \hline \mathbf{M}_{12}(r) & \mathbf{M}_2(r) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{K}_1^0(r) + \frac{\epsilon^2}{2!}\mathbf{K}_1^2(r) & \epsilon\mathbf{K}_{12}^1(r) \\ \hline \epsilon(\mathbf{K}_{12}^1(r))' & K_2^0(r) + \frac{\epsilon^2}{2!}\mathbf{K}_2^2(r) \end{array}\right] \tag{38}$$

This completes the three-term MacLaurin series approximation of $\mathbf{K}(r,\epsilon)$.

The Riccati matrix $\mathbf{K}$ is a function of reel radius r. Thus each term of the series in (38) must be computed as a function of r. The 'radius-embedding' procedure will simplify this problem [9.20].

The solutions $\mathbf{K}_1^0$ and $\mathbf{K}_2^0$ of (35) as continuous functions of r are obtained as follows. Differentiate (35) *w.r.t.* r to get the embedding equation

$$\frac{d\mathbf{K}_i^0}{dr}\mathbf{G}_i + \mathbf{G}'_i\frac{d\mathbf{K}_i^0}{dr} = -\mathbf{K}_i^0\alpha_i - \alpha'_i\mathbf{K}_i^0 + \mathbf{K}_i^0\zeta_i\mathbf{K}_i^0, \quad i=1,2. \tag{39}$$

where

$$\mathbf{G}_i \triangleq \mathbf{A}_i - \mathbf{S}_i\mathbf{K}_i, \; \alpha_i \triangleq \frac{d\mathbf{A}_i}{dr}, \; \zeta_i \triangleq \frac{d\mathbf{S}_i}{dr} \quad \text{with} \quad \mathbf{K}_i^0(r_o) = \mathbf{K}_{io}.$$

This differential equation must be integrated *w.r.t.* r starting with a known matrix $\mathbf{K}_{oi}$ at $r=r_o$. At each step of a routine for integration of this equation we solve a linear Lyapunov-type matrix equation for $d\mathbf{K}_i^0/dr$. Note that the linear algebraic equations (36) and (37) for $\mathbf{K}_{12}^1(r)$, $\mathbf{K}_1^2(r)$ and $\mathbf{K}_2^2(r)$ are also of Lyapunov type. Figure 3 presents a flow-chart of the computational procedure.

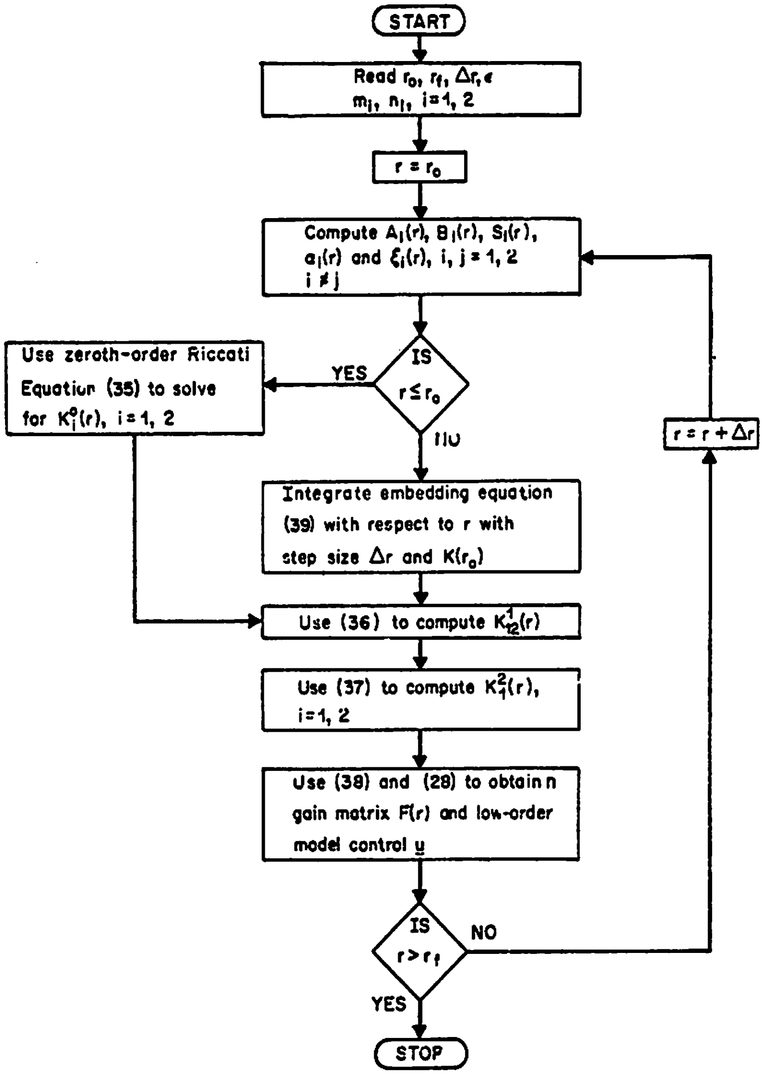

Figure 3. Flow-Chart for the Feedback Controller

9.2.5 Numerical Solution of the Feedback Gain Matrix

The following choice of $\mathbf{Q}_i$ and $\mathbf{R}_i$, $i = 1,2$ matrices was made:

$$\mathbf{Q}_1 = diag(0.5\ 0.5\ 0.5\ 1.0\ 1.0\ 5.0\ 1.0\ 5.0),\quad \mathbf{R}_i = \mathbf{I}_5. \tag{40}$$

$$\mathbf{Q}_2 = diag(1.0\ 5.0\ 1.0\ 0.5\ 0.5\ 0.5\ 1.0\ 1.0\ 5.0),\ \mathbf{R}_2 = \mathbf{I}_7, \tag{41}$$

where 0.5, 1.0 and 5.0 are the penalties assumed for field current, angular velocity and tension errors, respectively. The solution of (33) at r=0.84 is

$$\mathbf{K}_1^0(r_o) = \begin{bmatrix} 1.11 & -0.03 & 0.05 & 0.03 & 0.12 & -0.376 & -0.09 & 0.526 \\ & 18.87 & -0.003 & 11.04 & 0.07 & -1.91 & -0.51 & 1.066 \\ & & 0.688 & 16.78 & 0.01 & -0.04 & -0.01 & 0.054 \\ & & & 28.01 & 39.2 & 1.66 & 3.63 & 1.21 \\ & & & & 10.22 & 0.50 & 0.93 & 0.24 \\ & & & & & 0.85 & -3.06 & 0.07 \\ & & & & & & -8.63 & -0.02 \\ & & & & & & & 12.71 \end{bmatrix} \tag{42}$$

$$\mathbf{K}_2^0(r_0) = \begin{bmatrix} -8.78 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ & 23.32 & -2.39 & -0.16 & -0.01 & -0.02 & 0.90 & -0.35 & 0.75 \\ & & -10.2 & 0.14 & -0.50 & 0.01 & 3.48 & -1.31 & 3.3 \\ & & & 1.14 & -1.11 & 0.05 & -0.72 & 0.15 & -0.32 \\ & & & & 20.70 & -0.11 & 12.9 & -1.0 & 0.92 \\ & & & & & 0.73 & -0.07 & 0.02 & -0.032 \\ & & & & & & 20.01 & -0.347 & -1.64 \\ & & & & & & & 1.05 & 0.58 \\ & & & & & & & & 0.12 \end{bmatrix} \tag{43}$$

The matrices (42)-(43) were used as initial conditions to integrate the embedding equation (39) for $0.85 \leqslant r \leqslant 3.0$ and $\Delta r = 0.022$. Then for each increment Δr, $\mathbf{K}_1^0(r)$

and $\mathbf{K}_2^0(r)$ were obtained. Using (40) and $\epsilon = 0.5$ the matrices in (38) are

$$\mathbf{M}_1(r_0) = \begin{bmatrix} -0.08 & -1.4 & -0.08 & -1.31 & -0.17 & -0.43 & -0.08 & -13.7 \\ & 16.95 & -0.15 & 9.07 & -0.33 & -1.98 & -0.44 & -19.0 \\ & & 0.7 & -0.14 & -0.02 & -0.04 & -0.08 & -1.46 \\ & & & 26.0 & 3.52 & 1.44 & 3.3 & -21.0 \\ & & & & 0.94 & 0.45 & 0.83 & -43.8 \\ & & & & & 0.94 & -2.76 & -0.95 \\ & & & & & & -7.8 & -0.03 \\ & & & & & & & -233.0 \end{bmatrix} \tag{44}$$

$$\mathbf{M}_2(r_0) = \begin{bmatrix} -19.4 & 0.03 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ & 32.93 & -11.5 & -0.2 & -0.25 & -0.02 & 4.12 & -1.55 & 3.74 \\ & & -50.0 & 0.65 & -2.33 & 0.06 & 17.0 & -6.44 & 16.6 \\ & & & 1.13 & -1.1 & 0.04 & -0.9 & 0.22 & -0.5 \\ & & & & 20.6 & -0.11 & 13.5 & -1.23 & 1.52 \\ & & & & & 0.73 & 0\text{--}009 & 0.02 & -0.05 \\ & & & & & & 15.5 & -1.76 & -6.0 \\ & & & & & & & 0.4 & 2.25 \\ & & & & & & & & -4.22 \end{bmatrix} \tag{45}$$

$$\mathbf{M}_{12}(r_0) = \begin{bmatrix} 0.02 & 0.057 & 0.0 & -0.25 & -0.03 & -0.05 & -0.07 & 0.0 \\ 0.15 & 0.58 & -0.01 & -0.2 & -0.02 & -0.09 & -0.06 & 0.02 \\ 0.0 & 0.75 & 0.0 & -0.02 & 0.0 & 0.0 & -0.01 & 0.0 \\ -0.36 & 0.76 & -0.01 & -0.19 & 0.05 & 0.02 & -0.07 & 0.0 \\ -0.1 & 0.16 & 0.0 & -0.04 & 0.02 & 0.03 & -0.01 & 0.0 \\ 0.13 & 0.04 & 0.0 & -0.01 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ -10.44 & 8.4 & -0.03 & -2.1 & 1.05 & 1.32 & -0.73 & -0.08 \end{bmatrix} \tag{46}$$

Equations (44)-(46) are then used to obtain the feedback gain matrix $\mathbf{F}(r)$:

$$\begin{bmatrix} \mathbf{u}_1 \\ \cdots\cdots \\ \mathbf{u}_2 \end{bmatrix} = -\left[\begin{array}{c|c} \mathbf{R}_1^{-1}\mathbf{B}_1(r)\mathbf{M}_1(4) & \mathbf{R}_1^{-2}\mathbf{B}'_1(r)\mathbf{M}_12(r) \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \mathbf{R}_2^{-1}\mathbf{B}'_2(r)M_{21}(r) & \mathbf{R}_2^{-1}\mathbf{B}_2(r)\mathbf{M}_2(r) \end{array}\right] \begin{bmatrix} \mathbf{x}_1 \\ \cdots\cdots \\ \mathbf{x}_2 \end{bmatrix}$$

$$= -\left[\begin{array}{c|c} \mathbf{F}_1(r) & \mathbf{F}_{12}(r) \\ \cdots\cdots & \cdots\cdots \\ \mathbf{F}_{21}(r) & \mathbf{F}_2(r) \end{array}\right] \begin{bmatrix} \mathbf{x}_1 \\ \cdots\cdots \\ \mathbf{x}_2 \end{bmatrix}, \qquad (47)$$

where $\mathbf{u}_1$, $\mathbf{u}_2$ are five- and seven-dimensional controls while $\mathbf{x}_1$, $\mathbf{x}_2$ are eight-and nine-dimensional state vectors. The gain matrix $\mathbf{F}(r)$ in (47) has to be obtained as a function of reel radius r. The initial form of $\mathbf{F}(r)$ is shown below:

$$F(0.84) = \left[\begin{array}{cccccccc|ccccccccc}
-0.02 & -0.45 & & -0.42 & & -0.14 & & -4.38 & 0.0 & 0.22 & & -0.08 & & & & & \\
-0.45 & 5.4 & & 2.9 & -0.1 & -0.63 & -0.14 & -6.05 & 0.0 & 0.25 & & -0.06 & & & & & \\
& & 0.14 & & & & & -0.29 & 0.0 & 0.0 & & & 0 & & & & \\
0.09 & 0.54 & & -4.0 & -1.01 & 3.38 & 9.48 & -0.04 & 0.0 & 0.0 & & & & & & & \\
0.09 & 0.54 & & -4.0 & -1.01 & 3.38 & 9.48 & -0.04 & 0.0 & 0.0 & & & & & & & \\
\hline
& -0.12 & & 0.29 & 0.1 & -0.1 & & 8.62 & 16.0 & & & & & & & & \\
& -0.12 & & 0.29 & 0.1 & -0.1 & & 8.62 & 16.0 & & & & & & & & \\
& & & & 0 & & & 0.0 & & 4.9 & 21.6 & -0.27 & 0.99 & 0.0 & -7.22 & 2.73 & -7.03 \\
& & & & & & & & & & -7.03 & -0.15 & 0.48 & & 0.0 & & \\
-0.08 & -0.06 & & & & & & -0.66 & & & 0.2 & 0.36 & -0.35 & 0.0 & -0.29 & 0.07 & -0.15 \\
0.0 & & & & & & & 0.33 & & -0.08 & -0.74 & -0.35 & 6.55 & 0.0 & 4.28 & -0.39 & 0.48 \\
0.0 & & & & & & & 0.0 & & 0.0 & & & & & & & 0.0
\end{array}\right] \tag{48}$$

Note that many elements are zero up to two significant decimals and it was found that some remain constant. The final form of $\mathbf{F}(r)$ is

$$\mathbf{F}(r) =$$

$$\left[\begin{array}{ccccccc|ccccccccc}
-0.12 & f_1 & -0.34 & & -0.14 & -0.11 & f_2 & 0.0 & 0.25 & & & & & & & \\
f_1 & f_3 & f_4 & -0.1 & f_5 & -0.11 & -6.42 & 0.0 & 0.25 & & & & & & & \\
 & & 0.10 & & & & f_6 & 0.0 & 0.0 & & & 0 & & & & \\
f_7 & 0.41 & -3.42 & f_8 & f_9 & 10.0 & f_{10} & 0.0 & 0.0 & & & & & & & \\
f_7 & 0.41 & -3.42 & f_8 & f_9 & 10.0 & f_{10} & 0.0 & 0.0 & & & & & & & \\
\hline
 & & 0.15 & 0.1 & & & 8.62 & 16.0 & & & & & & & & \\
 & & 0.15 & 0.1 & & & 8.62 & 16.0 & & & & & & & & \\
 & & & & & & & & 5.68 & 24.4 & f_{11} & 0.90 & -0.09 & -6.9 & f_{12} & \\
 & & & & & & & & 5.68 & 24.4 & f_{11} & 0.90 & -0.09 & -6.9 & f_{12} & \\
 & & & 0 & & & f_{13} & & & f_{14} & f_{15} & 0.0 & 0.0 & f_{16} & f_{17} & f_{18} \\
 & & & & & f_{19} & & & & -0.68 & -0.37 & f_{20} & 0.0 & f_{21} & -0.29 & 0.32 \\
0.0 & 0.0 & & & & & 0.0 & & & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0
\end{array}\right] \tag{49}$$

where $f_i \triangleq f_i(r)$, $i = 1, \ldots, 21$ are obtained by quadratic approximation of the corresponding elements of $\mathbf{F}(r)$ at a sufficient number of values of r. Table 2 shows coefficients of these approximations. Figure 4 shows six of the r-varying elements of matrix $\mathbf{F}(r)$ whose dependence on r is appreciable. This controller can be implemented by a special purpose process computer for a direct digital control of the three-stand mill. It is hoped that the results of this section could serve as a first step of an industrial design project where the second step is testing at the plant.

TABLE 2. Gain matrix $\mathbf{F}(r)$ radius-varying functions

Index i	$f_i(r)=cr^2+br+a$ c	b	a
1	0.0300	-0.2772	-0.2368
2	0.0761	-1.3872	-3.2614
3	0.6815	-3.9518	8.0923
4	0.6082	-3.2900	5.0952
5	-0.2021	0.9876	-1.2674
6	-0.0976	-0.1867	-0.2675
7	-0.0524	0.5730	-0.3798
8	0.1740	-1.6105	0.2618
9	-0.0091	-0.3970	3.5874
10	-0.0427	-0.0032	-0.0458
11	0.3355	-1.1138	0.0612
12	0.0008	0.6142	-0.1578
13	0.2825	-1.6840	0.9191
14	-0.1114	3.2700	0.2078
15	-0.2032	1.1963	-0.6670
16	0.0885	-0.1733	0.4491
17	-0.0116	-0.1222	-0.1760
18	0.0520	0.0107	0.0297
19	0.0112	-0.1636	-0.0262
20	0.9100	-4.7564	9.6920
21	1.0215	-5.1246	7.6907

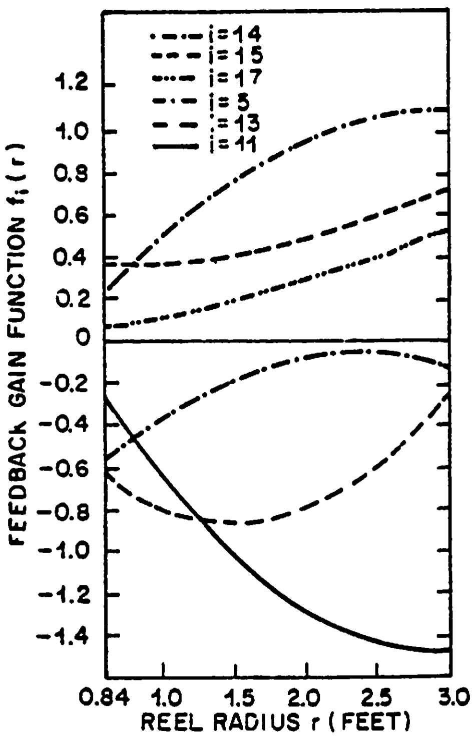

Figure 4. Six Elements of Gain Matrix $\mathbf{F}(r)$ as a Function of r

9.3 TRAFFIC CONTROL

In this section, the hierarchical control methods are applied to urban traffic systems. After the introduction of the time-delay algorithm of Tamura [9.27, 9.28] the dynamic model of such systems is presented. The above algorithm is then applied to a three-intersection traffic system in West London.

9.3.1 A Hierarchical TD Algorithm

In this section one of the more powerful algorithms for multilevel optimization of large-scale linear discrete-time systems with delays in both state and control will be presented. The problem was first introduced by Tamura [9.27], and has been further considered by many others [9.28-9.30].

Consider the following large-scale discrete-time system with delays:

$$\mathbf{x}(k+1) = \mathbf{A}_o\mathbf{x}(k) + \mathbf{A}_1\mathbf{x}(k-1) + \mathbf{A}_2\mathbf{x}(k-2) + \cdots + \mathbf{A}_s\mathbf{x}(k-s) + \mathbf{B}_o\mathbf{u}(k) + \mathbf{B}_1\mathbf{u}(k-1) + \cdots + \mathbf{B}_s\mathbf{u}(k-s) \tag{1}$$

where $\mathbf{A}_i$ and B_i, $i = 0,1,2,\cdots s$, are $n \times n$ and $n \times r$ matrices, respectively, and $\mathbf{x}$ and $\mathbf{u}$ are n and r−dimensional state and control vectors. System (1) has $2s$ delayed terms; hence it requires $2s$ discrete-time initial functions, assumed to be zero without loss of generality:

$$\mathbf{x}(k) = 0,\ \mathbf{u}(k) = 0,\ -s \leqslant k < 0,\ \mathbf{x}(0) = \mathbf{x}_0 \tag{2}$$

The physical interpretation of the initial functions in (2) for system (1) is that the system is operating at its steady-state and receives a disturbance at $k=0$ and drives it to a known value $\mathbf{x}_0$. The system cost function is assumed to be quadratic:

$$J = \frac{1}{2}\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \sum_{k=1}^{K}\frac{1}{2}\{\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}(k)\mathbf{R}(k)\mathbf{u}(k)\} \tag{3}$$

where $\mathbf{Q}(k)$ and $\mathbf{R}(k)$ are both assumed to be positive-definite. The optimal control problem is to find a sequence of control vectors $\mathbf{u}(0), \mathbf{u}(1), \ldots, \mathbf{u}(k-1)$ such that (3) is minimized while (1)-(2) and a set of inequality constraints

$$\mathbf{x}_{min} \leqslant \mathbf{x}(k) \leqslant \mathbf{x}_{max} \tag{4}$$

$$\mathbf{u}_{min} \leqslant \mathbf{u}(k) \leqslant \mathbf{u}_{max} \tag{5}$$

are satisfied. It goes without saying that a solution to the problem (1)-(5) in usual "centralized" methods by the application of the maximum principle results in a TPBV problem which involves both delay and advance terms, making the attainment of an optimum solution very difficult indeed, if not impossible. Here the objective is the application of hierarchical control via the interaction balance principle of Section 8.5.

Following the formulation of discrete-time maximum principle [9.31], let us define the Hamiltonian:

$$H(\mathbf{x}(k),\mathbf{u}(k),\lambda(k),k) = \tfrac{1}{2}\{\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}(k)\mathbf{u}(k)\} + \sum_{i=0}^{s} \lambda'(k+i)(\mathbf{A}_i\mathbf{x}(k-i) + B_u(k-i)) \tag{6}$$

where $k=0,1,\ldots,K-1$, $\lambda(k)$ is a vector of Lagrange multipliers at k, and $\lambda(K)$, $\lambda(K+1),\cdots$ are defined as zero vectors. In a manner similar to previous discussions regarding large-scale TD systems in Chapter 8, for a given vector $\lambda = \lambda^*$, the Lagrangian can be defined as

$$L(\mathbf{x},\mathbf{u},\lambda^*,k) = \tfrac{1}{2}\,\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) - \lambda'(k-1)\mathbf{x}(k) + \sum_{k=0}^{K-1}\{H(\mathbf{x}(k),\mathbf{u}(k),\lambda^*(k),k) - \lambda^*(k-1)\mathbf{x}(k)\}$$

Thus the optimization problem is to minimize (6) subject to (4)-(5). As before, this problem can be altered to that of maximizing the minimum of L(.) with respect to λ. The power behind the "time-delay algorithm" of Tamura [9.27] is the decomposition of this problem into $(K+1)$ independent minimization problems for a given λ^*, as in the three-level coordination formulation discussed in the last section, which reduces a "functional" optimization problem to a "parametric" one.

1. Problem $k=0$. By virtue of (7), definition (6), and constraints (4)-(5), the optimization problem for $k=0$ is

$$\underset{\mathbf{u}(0)}{Min}\, H(\mathbf{x}(0),\mathbf{u}(0),\lambda(0)) = Min\ \tfrac{1}{2}\{x'(0)\mathbf{Q}(0)x(0) + u'(0)R(0)u(0)\} + \sum_{i=1}^{s} \lambda^{*\prime}(i)(A_i x(0)+B_i u(0)) \tag{8}$$

subject to

$$\mathbf{x}(0) = \mathbf{x}_o,\ \mathbf{u}_{min} \leqslant \mathbf{u}(0) \leqslant \mathbf{u}_{max} \tag{9}$$

Now if $\mathbf{R}(0)$ is assumed to be a diagonal matrix, the necessary conditions for (8)-(9) lead to a set of m independent relations, each of which has an explicit solution given by $\partial H(\cdot)/\partial \mathbf{u}(0)=0$, *i.e.*,

$$\mathbf{u}^*(0) = sat_u\left\{-R^{-1}(0)\sum_{i=0}^{s} B'_i\lambda^*(i)\right\} \tag{10}$$

where the "saturation" function $sat_u(\cdot)$ is

$$sat_u(\sigma_j) = \begin{cases} \mathbf{u}_{max,j} & \text{if } \sigma_j > \mathbf{u}_{max,j} \\ \sigma_j & \text{if } \mathbf{u}_{min,j} \leqslant \sigma_j \leqslant \mathbf{u}_{max,j} \\ \mathbf{u}_{min,j} & \text{if } \sigma_j < \mathbf{u}_{min,j} \end{cases} \tag{11}$$

and the index j represents the jth element of control $\mathbf{u}_j$, $j=1,2,\ldots,m$.

2. Problem $k=1,2,\ldots,K-1$. The intermediate problem is defined by

$$\underset{\mathbf{x}(k),\mathbf{u}(k)}{Min}\ H\,(\mathbf{x}(k),\mathbf{u}(k),\lambda^*(k),k)-\lambda^{*\prime}(k-1)\mathbf{x}(k) \tag{12}$$

subject to

$$\mathbf{x}_{min} \leqslant \mathbf{x}(k) \leqslant \mathbf{x}_{max} \tag{13}$$

and

$$\mathbf{u}_{min} \leqslant u(k) \leqslant \mathbf{u}_{max} \tag{14}$$

Once again, assuming that $R(k)$ and $Q(k)$ are diagonal, the partial derivatives $\partial H(\cdot)/\partial x_i(k)$ and $\partial H(\cdot)/\partial u_j(k)$ for $i=1,2,\ldots,n$ and $j=1,2,\ldots,m$ lead to a set of $n+m$ independent one-parameter equations whose general solution is

$$\mathbf{x}^*(k) = sat_x\left\{-Q^{-1}(k)\left[-\lambda^*(k-1) + \sum_{i=0}^{s} A'_i\lambda^*(k+i)\right]\right\}$$

$$\mathbf{u}^*(k) = sat_u\left\{-R'(\mathrm{k})\left[\sum_{i=0}^{s} B'_i\lambda^*(k+i)\right]\right\} \tag{15}$$

where the ℓth element of the saturation function $sat_x(v)$ is

$$sat_x(v_\ell) = \begin{cases} \mathbf{x}_{max,\ell} & \text{if } v_\ell > \mathbf{x}_{max,\ell} \\ v_\ell & \text{if } \mathbf{x}_{min,\ell} \leqslant v_\ell \leqslant \mathbf{x}_{max,\ell} \\ \mathbf{x}_{min,\ell} & \text{if } v_\ell < \mathbf{x}_{min,\ell} \end{cases} \tag{16}$$

3. Problem $k = K$. This problem is

$$\underset{x(k)}{Min}\left\{\tfrac{1}{2}\,\mathbf{x}'(K)Q(K)\mathbf{x}(K) - \lambda^{*\prime}(K-1)\mathbf{x}(K)\right\} \tag{17}$$

subject to

$$\mathbf{x}_{min} \leqslant \mathbf{x}(K) \leqslant \mathbf{x}_{max} \tag{18}$$

whose solution is similarly given by

$$\mathbf{x}(K) = Sat_x\{Q^{-1}(K)\lambda^*(K-1)\} \tag{19}$$

The above so-called "time-delay algorithm" can be summarized as follows.

4. Time-Delay Algorithm

Step 1: At level one, solve $K+1$ analytic problems defined by (10)-(15) and (19) for a fixed set of Lagrange multiplier vector $\lambda(k) = \lambda^*(k)$, $k=0,1,\ldots,K-1$.

Step 2: At level two, the value of $\lambda^*(k)$ is improved through a gradient-type iteration

$$\lambda^{*r+1}(k) = \lambda^{*r}(k) + \delta^r \mathbf{d}^r(k) \tag{20}$$

where $\mathbf{d}^r(k)$ is a function of the error $\mathbf{e}^r(k)$, *i.e.*,

$$\mathbf{d}^r(k) = \mathbf{f}(\mathbf{e}^r(k))$$

$$= \mathbf{f}\left\{\sum_{i=0}^{s}[\mathbf{A}_i\mathbf{x}(k-i) + \mathbf{B}_i\mathbf{x}(k-i)] - \mathbf{x}(k+1)\right\} \tag{21}$$

which follows from our previous discussions, *i.e.*, Equation (1).

The following example illustrates the time-delay algorithm.

5. Example. Consider a simple second-order system

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \end{bmatrix}$$

$$+ \begin{bmatrix} 0.5 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} + \begin{bmatrix} 0 & 0.25 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1(k-1) \\ u_2(k-1) \end{bmatrix} \tag{22}$$

with cost function

$$J = \tfrac{1}{2}\, \mathbf{x}'(5)\mathbf{Q}(5)\mathbf{x}(5) + \tfrac{1}{2} \sum_{k=0}^{4} \{\mathbf{x}'(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}'(k)\mathbf{R}(k)\mathbf{u}(k)\} \tag{23}$$

and constraints

$$\begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \leqslant \mathbf{x}(k) \leqslant \begin{Bmatrix} 2 \\ 2 \end{Bmatrix}, \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} \leqslant \mathbf{u}(k) \leqslant \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}, \mathbf{x}(0) = \begin{Bmatrix} 0.5 \\ 0.5 \end{Bmatrix} \tag{24}$$

and $\mathbf{Q}(5) = diag(1,2)$, $\mathbf{Q}(k) = \mathbf{I}_2$, and $\mathbf{R}(k) = diag(1,0.5)$.

The problem was solved for an error tolerance of 0.001, a step size of 0.1 for a conjugate gradient iteration, and $\lambda(0) = (0.1\ ,\ 0.1)'$. The algorithm converged in 49 iterations, as shown in Figure 1. Several other initial values for $\mathbf{x}(0)$ and $\lambda(0)$ were tried and the convergence was achieved in a similar fashion.

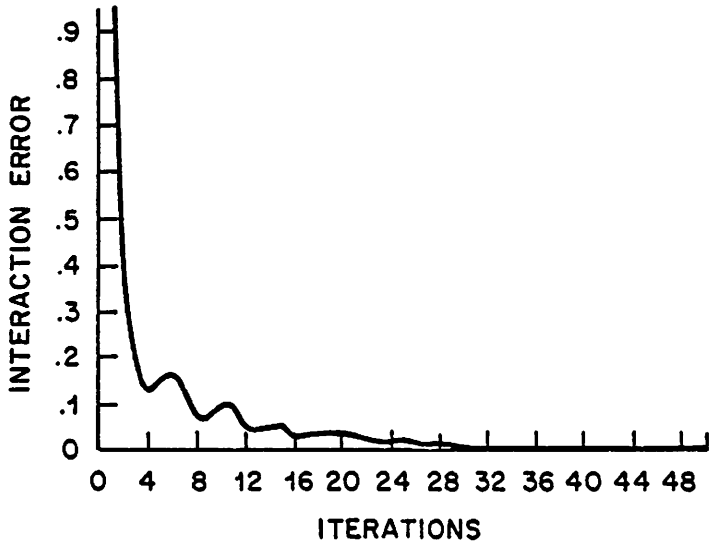

Figure 1. Interaction Error Versus Iterations for the TD Algorithm in Example 9.3.5

9.3.2 Model Of An Oversaturated Intersection

The urban traffic systems consist of intersections and streets. Consider a two one-way oversaturated intersection shown in Figure 2. Let the horizontal and vertical directions be represented by 1 and 2, respectively. Let $a_i(t)$ and $d_i(t)$ be the arrival and averaged partial rates of vehicles at the intersection, respectively, for direction $i=1,2$. Furthermore, let $s_i(t)$ denote the saturation flow rate of vehicles in direction i, *i.e.*, the maximum number of vehicles which can pass through the intersection per cycle in direction i when this direction has all available green lights [9.30].

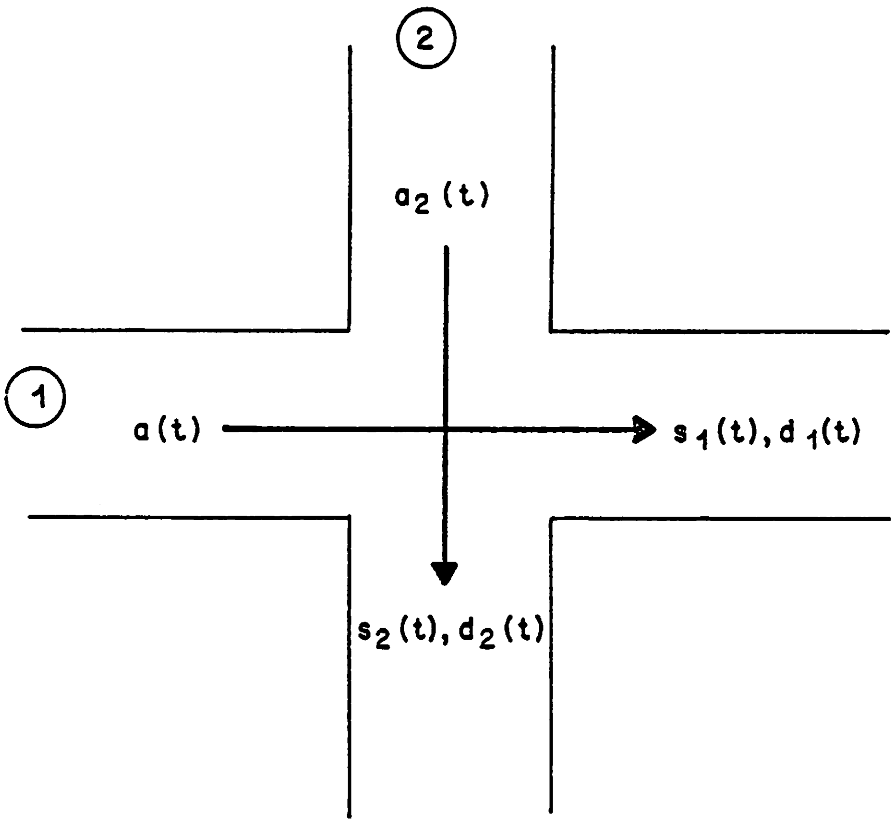

Figure 2. A Two One-Way Over Saturated Intersection

If G_i, $-$ 1,2 are the durations of green in the two directions and L the amber time loss, then the cycle time during $C = G_1 + G_2 + L$. Let $u_i(t)$, the fraction of time during which light is green in the *ith* direction, be the control variable. Then it can be readily seen that [9.30]

$$u_i(t) = d_i(t)/s_i(t) = G_i/(C - L) \tag{25}$$

for i=1,2. This formulation is based on the classical experiments of Webster and Cobb [9.32], where cycle time C is assumed to be constant. In this case, for each intersection, only one control $u_1 = u_i$ would suffice with the other control being $u_2 = C - u_i - L = G_e - u_i$, where G_e is the effective green. The control variable in a direction i is bounded by

$$\underline{u}_i \leqslant u_i(t) \leqslant \bar{u}_i \tag{26}$$

where the lower value $\underline{u}_i$ would not allow a waste of green light, while the upper value $\bar{u}_i$ would prevent the drivers from believing that the traffic signals are stuck [9.30]. Both limits in (26) are obviously positive constants. The state variables $x_i(t)$, $i=1,2$ of the intersection are defined as the instantaneous queue lengths in the two directions. The dynamic model's equation of the single intersection is obtained by considering the evolution of the queues, *i.e.*,

$$x_i(k+1) = x_i(k) + a_i(t) - d_i(t) \tag{27}$$

which leads to

$$x_i(k+1) = x_i(k) + a_i(t) - s_i(k)u_i(k) \tag{28}$$

following (25). The state of the intersection is furthermore constrained by physical limitations, *i.e.*,

$$0 \leqslant x_i(k) \leqslant \bar{x}_i \,. \tag{29}$$

For a complex traffic intersection shown in Figure 3, the portions of queues which turn would also be considered as state variables. The usual practice in traffic engineering has been to take the portions of turns to non-turning as fairly constant. Thus, a turning movement can be thought of as a release from the main queue [9.33]. Thus, the state equation for the intersection of Figure 3 can be written by

$$x(k+1) = x(k) + a(k) - d(k) \tag{30}$$

where $x' = (x_{11}, x_{12}, x_{21}, x_{22})$, $a' = (a_{11}, \ldots, a_{22})$, and $d' = (d_{11}, \ldots, d_{22})$. Now if the control u in the horizontal direction is defined as

$$u = d_{11}/s_{11} = d_{12}/s_{12} \tag{31}$$

then the vertical direction's control will be

$$G_e - u = d_{21}/s_{21} = d_{22}/s_{22} \tag{32}$$

where, as defined earlier, G_e is the effective green duration. Based on the control definitions (31)-(32), the intersection's state equation (30) can be rewritten by

$$x(k+1) = x(k) + u(k) + c \tag{33}$$

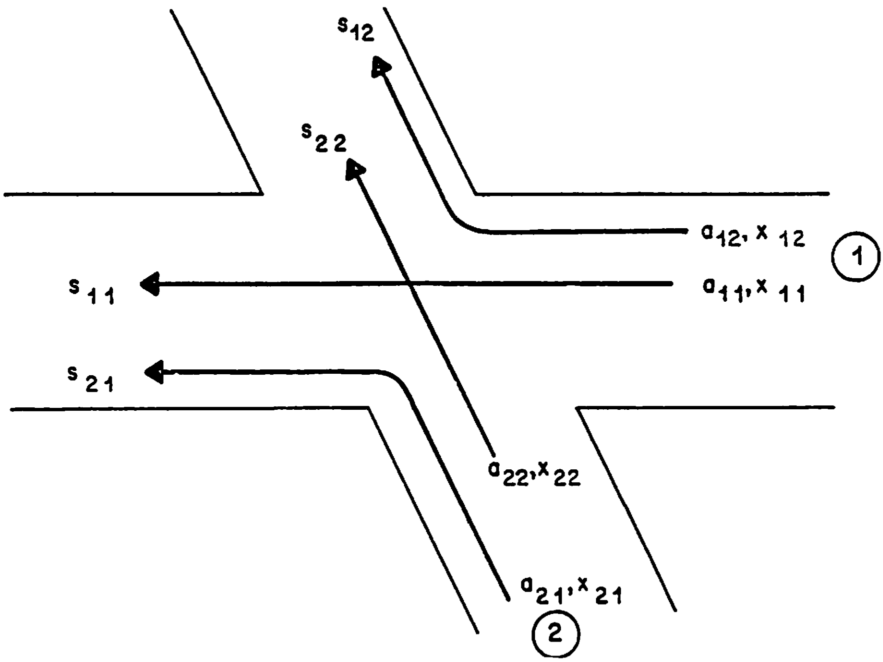

Figure 3. A More Complex Intersection Point

where $b' = (-s_{11}, -s_{12}, s_{21}\ s_{22})$ and $c' = (d_{11}, d_{12}, d_{21} - G_3\, s_{21}, d_{22} - G_e\, s_{22})$.

It must be noted that in traffic engineering practice, two kinds of control are commonly used. These are "split" and "offset" controls [9.33]. The first, which has been used here, is based on splitting the effective green at each intersection as described earlier. The second control represents the time difference between the start of green from one intersection to that of the next. In the present study, since the macro-behavior of the system is of concern, the offset is considered to be constant.

The actual traffic situations, as mentioned earlier, would include a number of intersections connected by streets. Consider a three-intersection area shown in Figure 4.

Under such conditions, it is a common practice to consider the interconnecting streets as pure delay elements [9.30]. Therefore, the arrival rate in the *ith* direction at $(i+1, j+1)$ intersection, shown in Figure 4, would be

$$a_{i,j+1} = s_{ii}\, u_{ii}(k-h) \tag{34}$$

where h is number of unit delays. With this observation, the dynamic model of a network of intersections and streets can be given by [9.30]

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \sum_{j=0}^{h} B_j\, \mathbf{u}(k-j) + \mathbf{z}(k) \tag{35}$$

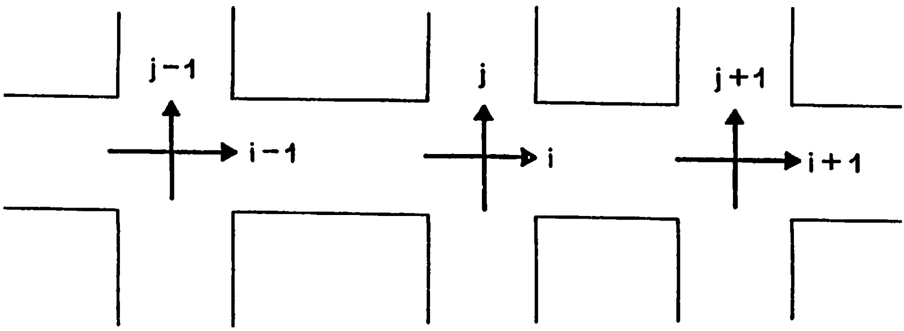

Figure 4. A Three-Intersection Traffic Area

where $\mathbf{x}$ is the state vector (queues), $bu(k)$ is control vector, $\mathbf{B}_j$, $j=0,1,\ldots,h$ are the control weighting matrices, $\mathbf{u}(k-j)$, $j=1,2,\ldots,h$ are delayed controls and $\mathbf{z}(k)$ is the vector of external inputs, *i.e.*, from outside the system. The state and control vectors are further bounded as before, *i.e.*,

$$0 \leqslant \mathbf{x}(k) \leqslant \bar{\mathbf{x}}\,,\ \underline{\mathbf{u}} \leqslant \mathbf{u}(k) \leqslant \bar{\mathbf{u}}\,. \tag{36}$$

The cost function for a multi-junction urban traffic control system can be chosen to be quadratic,

$$J = 1/2 \sum_{k=0}^{K-1} \left\{ \mathbf{x}'(k)\ \mathbf{Q}\mathbf{x}(k) + (\mathbf{u}(k) - \mathbf{u}^*)'\ \mathbf{R}\left[\mathbf{u}(k) - \mathbf{u}^*\right] \right\} \tag{37}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are positive definite weighting matrices. The vector $\mathbf{u}^*$ denotes the nominal control which can be chosen in advance for efficient use of the system. The optimal control of the urban traffic system is to find a sequence of green signals periods $u_i(t)$ which would satisfy dynamic constraint (35), bounds (36) while minimizing cost function (37). In the next section, this problem is solved by the time-delay algorithm of Tamura [9.27, 9.28], which was considered in detail in Section 9.3.1.

9.3.3 A Traffic Control Case Study

In this section a case study representing a modified form of a three intersections system of West London network which has been discussed in various reports by Singh and Tamura [9.30, 9.33] is considered. The system is shown in Figure 5 where 12 possible queues (states) are present with three control variables, one for each intersection. Following the development of the previous section, the state equations can be seen to be

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + a_1(k) - s_1u_1(k) \\
x_2(k+1) &= x_2(k) + a_2(k) - s_2u_2(k) \\
x_3(k+1) &= x_3(k) + a_3(k) - s_3(G_e-u_1(k)) \\
x_4(k+1) &= x_4(k) + a_4(k) - s_4(G_e-u_1(k)) \\
x_5(k+1) &= x_5(k) + a_5(k) - s_5(G_e-u_1(k)) \\
x_6(k+1) &= x_6(k) + b_{26}s_2u_2(k-2) + b_{16}(G_e-s_3u_2(k-2)) \\
&\quad - s_6u_2(k) \qquad\qquad (38) \\
x_7(k+1) &= x_7(k) + b_{27}s_2u_2(k-2) + b_{17}(G_e-s_3u_2(k-2)) \\
&\quad - s_7u_2(k) \\
x_8(k+1) &= x_8(k) + a_8(k) - s_8(G_e-u_2(k)) \\
x_9(k+1) &= x_9(k) + a_9(k) - s_9(G_e-u_2(k)) \\
x_{10}(k+1) &= x_{10}(k) + s_7u_2(k-1) + s_9(G_e-u_2(k-1)) \\
x_{11}(k+1) &= x_{11}(k) + a_{11}(k) - s_{11}u_3(k) \\
x_{12}(k+1) &= x_{12}(k) + a_{12}(k) - s_{12}(G_e-u_3(k))
\end{aligned}
$$

where G_e is the equivalent green and sixth, seventh, and tenth directions have no inflow rate of vehicle since they are internal with respect to the system. The numerical data for this system is taken from the Planning and Transportation Department of the Greater London Council as reported by Singh [9.30], Singh and Tamura [9.33]. Table 1 provides the numerical data for the 3-intersection system. Based on this data and the general discrete-time-delay state equation (35), the corresponding state equation for the system becomes

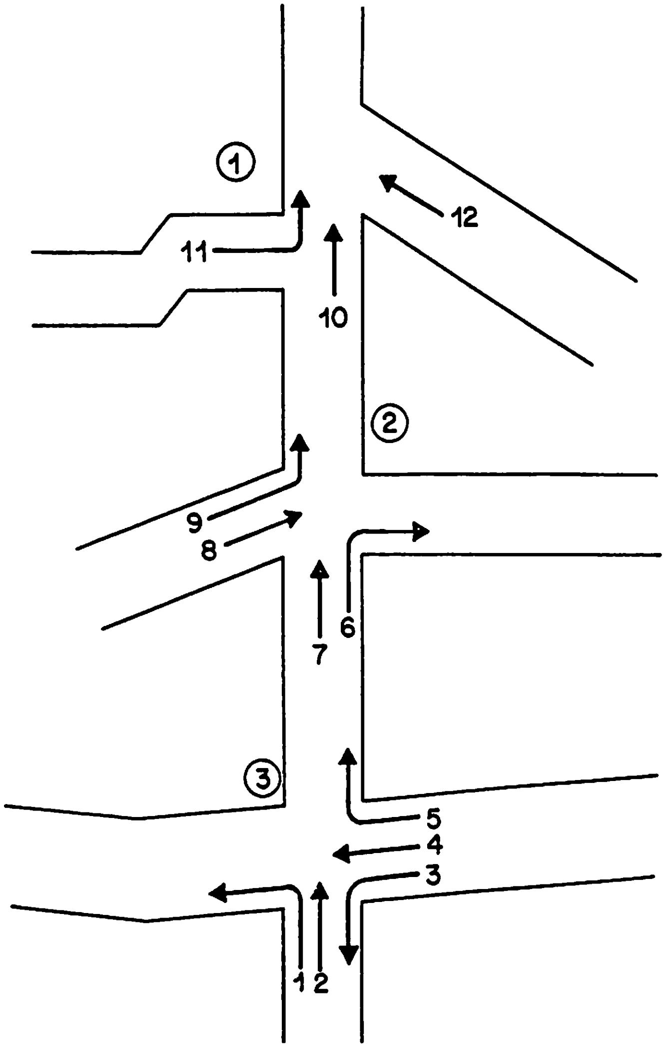

Figure 5. A Three-Intersection West London System

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{B}_o\Delta\mathbf{u}(k) + \mathbf{B}_1\Delta\mathbf{u}(k-1) + \mathbf{B}_2\Delta\mathbf{u}(k-2) + \mathbf{z}(k) \tag{39}$$

where new control variables $\Delta\mathbf{u}_i$ $i=1,2,3$ are defined by

$$\Delta u_i = u_i - u_i^*, u^* = (0.45 \quad 0.3 \quad 0.5) \,. \tag{40}$$

Matrices $\mathbf{B}_j$, $j=0,1,2$ and vector $\mathbf{z}(k)$ are

$$\mathbf{B}_o = \begin{bmatrix} -25 & 0 & 0 \\ -65 & 0 & 0 \\ 4 & 0 & 0 \\ 31 & 0 & 0 \\ -25 & 0 & 0 \\ 0 & -26 & 0 \\ 0 & -64 & 0 \\ 0 & 132 & 0 \\ 0 & 34 & 0 \\ 0 & 0 & -96 \\ 0 & 0 & -25 \\ 0 & 0 & 90 \end{bmatrix}, \quad \mathbf{z}(k) = \begin{bmatrix} -8.20 \\ -21.60 \\ 0.90 \\ 7.70 \\ 8.40 \\ 1.62 \\ 2.56 \\ -64.20 \\ -16.40 \\ -8.40 \\ -10.50 \\ -33.50 \end{bmatrix} \tag{41}$$

with non-zero elements of $\mathbf{B}_1$ and $\mathbf{B}_2$, $B_1(10,2) = 30$, $B_2(6,1) = 18.3$ and $B_2(7,1) = 42.7$. The initial state and the bounds on state and control vectors are given by

$$\mathbf{x}'(0) = (30 \;\; 30 \;\; 70 \;\; 70 \;\; 70 \;\; 40 \;\; 40 \;\; 30 \;\; 30 \;\; 20 \;\; 30 \;\; 30) \tag{42}$$

$$\bar{\mathbf{x}}' = (40 \;\; 40 \;\; 80 \;\; 80 \;\; 80 \;\; 50 \;\; 50 \;\; 40 \;\; 40 \;\; 25 \;\; 40 \;\; 40)$$

$$\underline{\mathbf{u}}' = (-0.25 \;\; -0.1 \;\; -0.3), \; \bar{u}' = (0.25 \;\; 0.4 \;\; 0.2) \; .$$

The cost function was assumed to be in quadratic form (36) with the new control vector Δu, *i.e.*,

$$J = 1/2 \sum_{k=0}^{3} \left\{ \mathbf{x}'(k)\mathbf{Q}\mathbf{x}(k) + \Delta\mathbf{u}'(k)\mathbf{R}\Delta\mathbf{u}(k) \right\} , \tag{43}$$

where $\mathbf{Q} = diag(1, 1, 1, 1, 1, 1.5, 1.5\ 1, 1, 2, , 1)$, $\mathbf{R} = \mathbf{I}_3$. It is noted that in the choice of Q, the internal flows through routes 6, 7 and 10 are weighted higher due to limited storage capacities of the main interconnecting street (see Figure 5). The system was simulated on an IBM 3032 using FORTRAN 77 and the time-delay algorithm of Section 9.3.1. The convergence was reached in 189 iterations and 9.2 minutes. The same results on this system were obtained by Singh and Tamura [9.33] on an IBM 370/165 with 153 iterations and 2.73 minutes of CPU time. Typical optimum queues and greens for the system are shown in Figure 6. The results indicate that the queues reduce considerably in all directions except in directions 3, 4, and 5. One reason for this is the higher values of the incoming flow rates in those directions and narrower streets, as is evident from the saturation values s_i, $i = 3$, 4 and 5 in Table 1. Next section treats the graph theoretical approach and hierarchical control for large-scale urban traffic systems.

TABLE 1. Numerical data for the three-intersection system of Figure 5

i	1	2	3	4	5	6	7	8	9	10	11	12
$s_i(k)$ per cycle	25	65	4	31	34	26	64	132	34	96	25	90
$a_i(k)$ per cycle	3.0	7.6	27	21.7	23.7	0	0	15	4.0	0	2.0	2.4
	$G_e = 0.9$, $b_{16} = b_{26} = 0.3$, $b_{17} = b_{27} = 0.7$.											

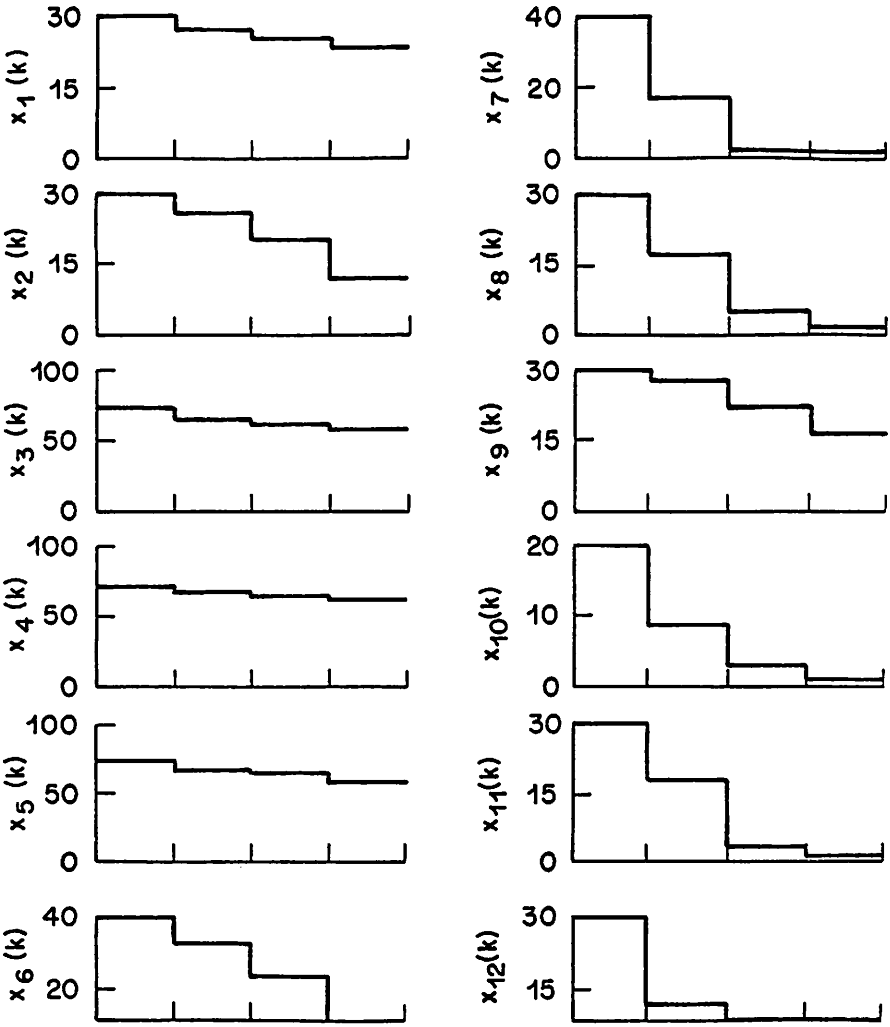

Figure 6a. Typical Optimum Queues for the London System

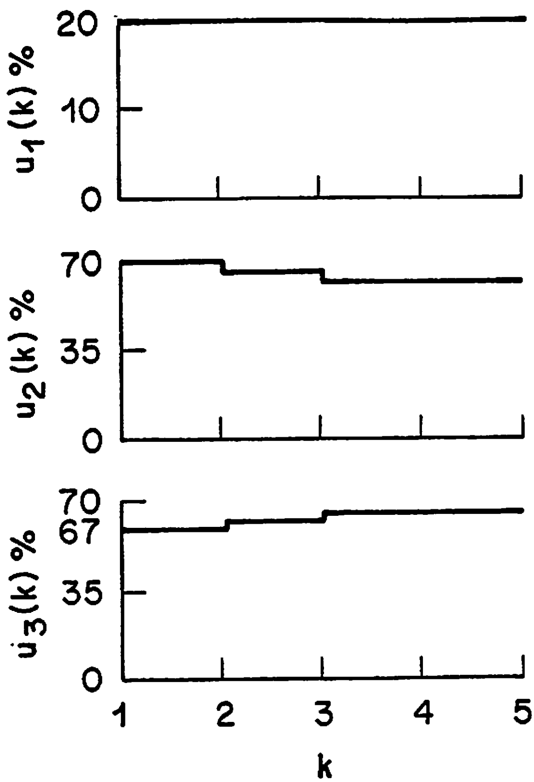

Figure 6b. Typical Optimum Queues for the London System

9.4 WATER RESOURCES SYSTEMS

In this section a discrete-differential dynamic programming scheme is introduced for the optimization of TD discrete-time systems and applied to an existing water resources system.

It is a well known fact that straight application of Dynamic Programming either in its forward or backward algorithm, requires a great deal of computer storage and computations. Computational requirements increase almost exponentially with a system's order--the "curse of dimensionality." The computer requirements of dynamic programming (DP) have been reduced by a number of researchers, including Bellman. One such modification is discrete differential dynamic programming (DDDP). DDDP is an iterative technique in which the DP's recursive formula is used about a nominal trajectory to improve upon the discrete decisions and state [9.22].

9.4.1 Discrete Differential Dynamic Programming

Consider the discrete state equation

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k),\ \mathbf{u}(k),\mathbf{x}(k-h),k),\ k=0,1,\ldots,K-1, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{u}$ are n–and r–dimensional state and control vectors, respectively and h is the delay. The system constraints may be defined as

$$\mathbf{x}(k)\epsilon X(k)\ \subset R^n,\ \ \mathbf{u}(k)\epsilon U(k)\ \subset R^r \tag{2}$$

where $X(k)$ and $U(k)$ are the states and control admissible domains, respectively. The optimization problem is to obtain the sequence of controls (decisions) $\mathbf{u}_i(k)$, $i=1,\ldots,$ which satisfy (1) while maximizing (or minimizing) an objective function,

$$F = \sum_{k=0}^{K-1} R(\mathbf{x}(k),\ \mathbf{u}(k),\ k) \tag{3}$$

where F is the sum of returns R obtained as the result of applying decision $U(k)$ for one stage when the system is at state $X(k)$. Using DP's forward recursive algorithm,

$$F^*(\mathbf{x}(k+1),k+1) = \max_{\mathbf{u}(k)\epsilon U(k)} \{R(\mathbf{x}(k),\ \mathbf{u}(k),h) + F^*(\mathbf{x}(k),\ k)\} \tag{4}$$

where $F^*(\mathbf{x}(k),k)$ is the maximum total returns from stage 0 to stage k. Assuming that the system (1) is invertible [9.22] one can solve for $\mathbf{x}(k)$, *i.e.*,

$$\mathbf{x}(k) = \mathbf{h}(\mathbf{x}(k+1),\ \mathbf{x}(k),\mathbf{u}(k),\ k) \tag{5}$$

Substituting (5) in the recursive algorithm (4) results in

$$F^*(\mathbf{x}(k+1),\ k+1) = \max_{\mathbf{u}(k)\epsilon U(k)} \{R\{\mathbf{h}(\mathbf{x}(k+1),\ \mathbf{x}(k),\mathbf{u}(k),k) + F^*(\mathbf{x}(k),\ k)\}\} \tag{6}$$

which can be solved for every state $\mathbf{x}(k+1)\epsilon X(k+1)$ as a function of $\mathbf{x}(k)$ only. In many applications, such as water resources systems, it is often to have the states (reservoirs storage levels) satisfy certain conditions at the beginning and ended of the operating period, *e.g.*, a year. Hence, the optimization problem (1)-(3) has, in addition, the following boundary conditions:

$$\mathbf{x}(1) = \mathbf{a}_1, \ \mathbf{x}(N) = \mathbf{a}_N \tag{7}$$

In the approximate DDDP algorithm proposed by Heidari et. al., [9.22], a sequence of initial decision or policy, $\bar{\mathbf{u}}(k)$, $k = 0,1, \ldots, N-1$, which results in a nominal sequence of state $\bar{\mathbf{x}}(k)$, is assumed. The corresponding total return (3) due to the nominal state and control vectors becomes

$$\bar{F} = \sum_{t=0}^{N-1} R(\bar{\mathbf{x}}(k), \bar{\mathbf{u}}(k), k) \tag{8}$$

In an attempt to reduce the computer time and memory requirements of the dynamic programming problem (1)-(3) and (7), the state admissible domain $X(k)$ in (12) is reduced by the following method. Consider a set of m–dimensional incremental state vectors

$$\Delta\mathbf{x}^i(k) = \left\{\delta x_1^i(k), \ \delta x_2^i(k), \ldots, \delta x_m^i(k)\right\}' \tag{9}$$

for $k=0,1, \ldots, N$ and $i=1,2, \ldots, M^m$ where M is the total number of quantized increments of the state domain. The *ith* component δx^i can take on any value σ_t^i, $i=1,2, \ldots, M^m$ and $t=1,2, \ldots, m$ within the admissible state domain. When the incremental state vectors (g) are added to the nominal trajectory, a subdomain, called $D(k)$, is obtained,

$$D(k): \ \mathbf{x}(k) + \Delta\mathbf{x}^i(k), \ i=1,2, \ldots, M^m \tag{11}$$

Figure 1 shows two such subdomains corresponding to $m=2, M=3$, and $m=3, M=2$. All lattice points in $D(k)$, called the "corridor" by Heidari et al., [9.22] are shown in Figure 2 for $m=1$, $M=$, and $N=12$.

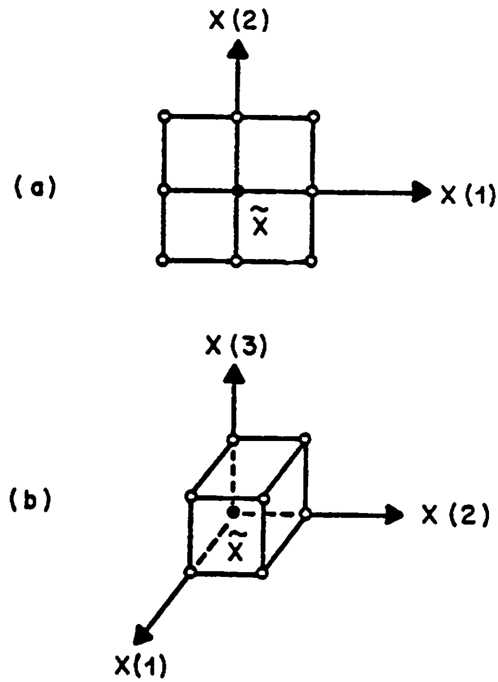

Figure 1. Two Subdomains for DDDP Algorithms (a) $m=2$, $M=3$ and (b) $m=3$, $M=2$

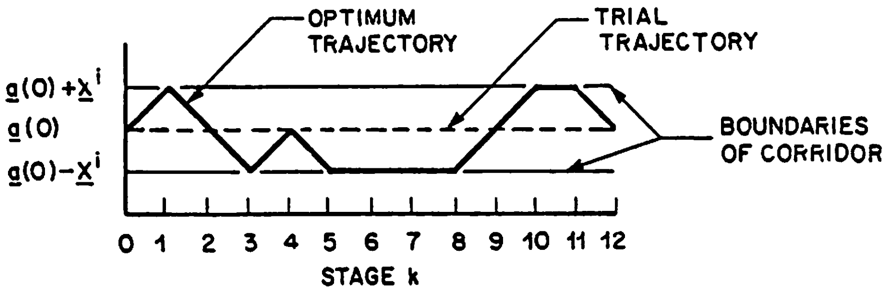

Figure 2. An Illustration of a "corridor" in DDDP Algorithm for $m=1$, $M=3$ and $N=12$

In the application of DDDP, at the ℓth iteration, the state domain (2) is reduced to the corridor $C(t)$. Hence, using the recursive formula given by (6) a new objective function F_ℓ^* is obtained. Now, if $F_\ell^* > \bar{F}_\ell$, the new trajectory is used as nominal for the $(\ell + l)th$ iteration; otherwise, the previous nominal trajectory is kept. The corridoring technique for ℓth iteration of DDDP is described by the following algorithm.

1. Algorithm DDDP [9.23]

Step 1: Use the $(\ell-1)-th$ optimal trajectories as the ℓth nominal trajectory, *i.e.*,

$$\left\{\bar{\mathbf{x}}(k)\right\}_{\ell} = \left\{\mathbf{x}^*(k)\right\}_{\ell-1}, \left\{\bar{\mathbf{u}}(k)\right\}_{\ell} = \left\{\mathbf{u}^*(k)\right\}_{\ell-1} \tag{12}$$

Step 2: Select $(\sigma_\ell^i)^k$, $i=1,2,\ldots,M^m$, $k=1,\ldots,m$ to define corridor C_ℓ and use (6) to maximize F subject $\mathbf{x}(k_o)\epsilon C_\ell(t)$.

Step 3: Trace and obtain the optimum trajectory satisfying boundary conditions, (7), *i.e.*, $(\mathbf{x}^*(k))_\ell$ and $(\mathbf{u}^*(k))_\ell$, obtain F_ℓ^*.

Step 4: If $F_{\ell-l}^* > F_{\ell-l}^*$ go to Step 1.

Step 5: Stop.

A few points regarding the above approximate DDDP algorithm must be discussed. The choice of the corridor widths $\sigma^i(t)$, $i=1,2,\ldots,M^m$ are somewhat arbitrary at this point. However, in this approach the corridor $C(t)$ is formed in such a manner that the nominal trajectory $(\bar{\mathbf{x}}(k))_\ell$ lies in the middle of its domain. The initial choice σ_0^i in this study was on the trial and error basis. The computational cost of analysis could easily be tolerated to find a set of σ_0^i that could rapidly lead to a near-optimum solution. Therefore, the algorithm may be repeated to isolate the most favorable set of the initial corridor widths σ_0^i. If the width is kept constant for every iteration and very little or no more improvements can be achieved after the *kth* iteration, it is recommended that σ_ℓ^i be reduced at the $(\ell + l)$th iteration and continue the optimization process. No theoretical justification for the nature of the solution is available at this point. However, one practical way of testing the local or global optimality of the algorithm is to try two or more different nominal trajectories at the middle and extreme lattice points of the state domain which satisfy the boundary conditions and system constraints. If these trajectories lead to the same solution, one may conclude that a nearly global solution is obtained. Further theoretical or computational investigations for global optimization is subject to further studies. The following existing example illustrates the DDDP algorithm.

9.4.2 Khuzestan Water Resources System

The water resources system considered in this section is presented in Figure 3. As shown, there are three reservoirs receiving water from two main streams, three hydropower plants and three irrigatable areas. This system is to operate for 12 months of the year. The parameters in Figure 3 are defined as follows:

$y_i(k)$ = Inflow into ith-reservoir during period (month or stage k, i=1,2,3)

$u_i(k)$ = Outflow from ith-reservoir during period k, i=1,2,3 and inflow into the irrigatable areas, i=4,5,6

$x_i(k)$= Storage (state) of the ith-reservoir during the period k, i=1,2,3.

As seen from Figure 3. The outflows from reservoirs 1 and 2 are diverted to three irrigatable areas in the basin. Hence, it is assumed that a portion of diverted irrigation water, returns to the main stream with a delay of one month. The value of $\alpha(k)$ depends on the types of soils, crops cultivated, and the chronological conditions. Furthermore, the following assumptions are made for this system:

(i) No appreciable evaporation loss is considered.

(ii) Municipal and industrial use are considered to be relatively insignificant compared to irrigation [9.24].

(iii) Water inflows, irrigation and power demands are as summed to be deterministic and statistically known.

(iv) No water quality considerations are assumed for this system.

(v) No recreational benefits were assigned to the system.

In lieu of the above assumptions, the system's state equation is obtained through the continuity principle,

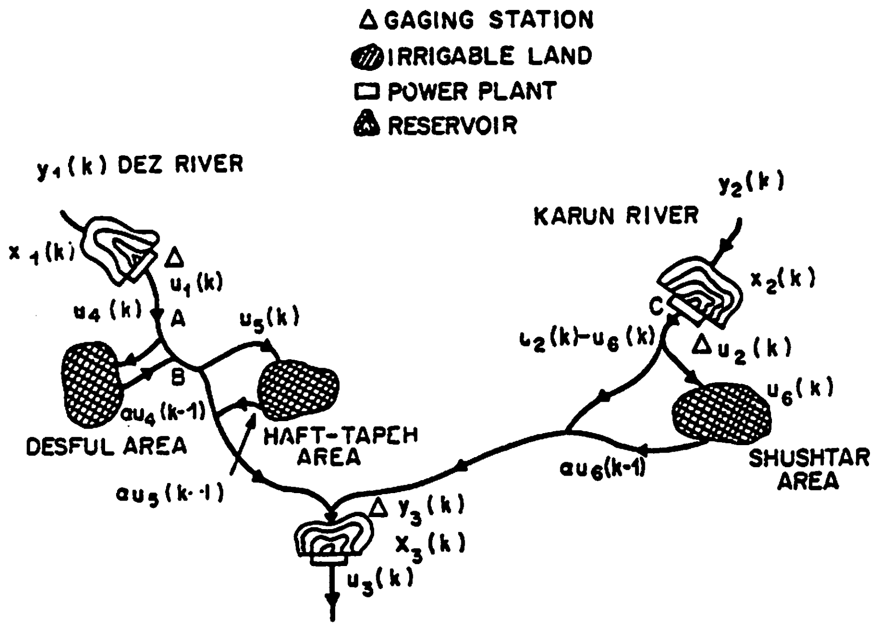

Figure 3. A Schematic of Khuzestan Water Resources System

$$x_1(k+1) = x_1(k) + y_1(k) - u_1(k) \tag{13}$$

$$x_2(k+1) = x_2(k) + y_2(k) - u_2(k) \tag{14}$$

$$\begin{aligned} x_3(k+1) &= x_3(k) + y_3(k) - u_3(k) \\ &= x_3(k) + u_1(k) + u_2(k) - u_3(k) - u_4(k) - u_5(k) - u_6(k) \\ &\quad + \alpha(k)[\, u_4(k-1) + u_5(k-1) + u_6(k-1)] \end{aligned} \tag{15}$$

The irrigation demands $u_i(k)$, $i=4,5,6$ depend on the availability of water at points A, B and C and the corresponding agricultural needs. The irrigation return of the Dezful and Haft-Tapeh areas, as well be seen shortly, are almost identical, hence there will be no conflicting demands for $u_4(k)$ and $u_5(k)$. The demand for water at Dezful area (node A Figure 3) is initially met and then diversions required for Haft-Tapeh at node B will be satisfied. However, if the irrigations demands cannot be satisfied, steps are taken to penalize the returns for the shortage so that the outflows $u_i(k)$, $i=1,2,3$ would be modified. Based on the data gathered, the following constraints are obtained for the state and decision variables

$$\begin{aligned} 0 &\leqslant x_1(k) \geqslant 2500\ MCM\ (\textit{million cubic meters}) \\ 0 &\leqslant x_2(k) \leqslant 1200\ MCM \\ 0 &\leqslant x_3(k) \leqslant 800\ MCM \end{aligned} \tag{16}$$

$$\begin{aligned} 0 &\leqslant u_1(k) \leqslant x_1(k) + y_1(k) \\ 0 &\leqslant u_2(k) \leqslant x_2(k) + y_2(k) \\ 0 &\leqslant u_3(k) \leqslant x_3(k) + y_3(k) \end{aligned} \tag{17}$$

The parameters for power generation are:

$$\begin{aligned} &E_t = 11 \times 10^6\ MW\text{-}hr\ yr \\ &P_1 = 520\ MW,\ P_2 = 1000\ MW,\ P_3 = 400\ MW \end{aligned} \tag{18}$$

where P_1, P_2 and P_3 are installed capacities of power plants 1,2 and 3, respectively, and E_t is the total annual energy demand.

The total return during a time interval starting at state k is

$$R(k) = R_i(k) + R_p(k) \tag{19}$$

where $R(k)$ is the system's total return, $R_i(k)$ and $R_p(k)$ are the gross returns from irrigation and hydropower generation during period k, respectively.

One of the major activities of the water resources system shown in Figure 3 is to supply irrigation water to the three given areas. These areas with wheat, barley and sugar cane as their main crops, have a growing season of about 300 days. In this system, it is assumed that the crop pattern remains the same from year to year, but the irrigatable land can vary depending on the annual target irrigation water and land available. Table 1 presents the monthly distribution of the annual target irrigation requirements. The return $R_I(k)$ is defined to be

$$R_I(k) = b_4(k)u_4(k) + b_5(k)u_5(k) + b_6(k)u_6(k) \tag{20}$$

TABLE 1. MONTHLY DISTRIBUTION OF THE ANNUAL TARGET IRRIGATION DIVERSION REQUIREMENTS

Percentage of Target Annual Diversion For Irrigation			
Month	Haft-Tapeh	Dezful	Shushtar
October	8.98	31.24	1.57
November	10.07	6.67	3.64
December	13.33	3.77	6.09
January	11.98	4.33	5.58
February	10.48	3.26	11.46
March	10.40	4.60	13.88
April	8.98	4.37	28.17
May	8.74	4.14	4.63
June	7.70	6.39	5.02
July	5.61	9.66	10.67
August	3.69	11.88	5.33
September	2.04	9.79	4.02
Total	100%	100%	100%

(Source: Khuzestan Water and Power Authority)

where the unit gross irrigation return coefficients $b_j(k)$, $j=4,5,6$ are assumed to be a function of the yearly irrigation target output [9.25] and are shown in Figure 4. There may also be some irrigation shortages during one or more months of the year. Depending on such shortages, there are some lost crops and hence a penalty function is needed to take them into account. The penalty function for all three areas used in this analysis is shown in Figure 5.

The other important activity of the system of Figure 3 is hydropower generation. This energy would be distributed among domestic, industrial, and irrigation demands. The demand schedule for energy of the Khuzestan system is shown by Table 2. During every month of the year, depending on the excess or shortage of energy, the system may have to export or import energy. Typical energy output functions are calculated from the following equation [9.26]

$$E_i = C_i e u_i H_i = 1.2\times 10^{\ 3} C_1 e u_{1cki} x_i \tag{21}$$

where

E_i = Energy output (MW-hr)/mo from power plant i, i=1,2,3

C_i = Constant of proportionality (MW hr/mm^4)

e = turbine efficiency (%)

u_i = water flow through turbine of ith power plant (CM mo)

H_i = effective head of ith reservoir (m)

x_i = ith reservoir storage (MCM).

In (21), the effective head H_i is assumed to be proportional to the square root of storage, *i.e.*, $H_i = 0.0012,\ x_i^{1/2}$. The return $R_p(k)$ was chosen to be.

$$R_p(k) = b_1(k)E_1(k) + b_2(k)E_2(k) + b_3(k)E_3(k) \tag{22}$$

where $b_j(k)$, j = 1,2,3 are the unit gross power generation return coefficients. These coefficients were set for three different situations: (a) when power generated was equal to demand, b_i = 0.002. i=1,2,3 Rials MW-hr. (b) when power generated was above the demand, it was assumed that the surplus could be sold at a price of b_i = 0.0008,i = 1,2,3 Rials MW-hr and the demand energy would have a price as (a); and (e) when generated power was below the demand, t was assumed that a penalty of

0.001 Rials MW-hr would be imposed on the return from power* This penalty forced the system to generate the required energy without violating the system's constraints. Table 3 presents the remaining data of the system in Figure 1. The total annual return of the system would be

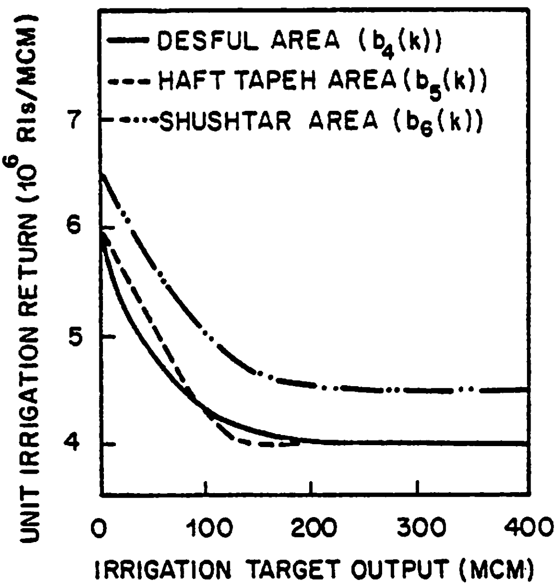

Figure 4. Annual Unit Irrigation Returns (Million Rials) as a Function of the Target Output (MCM)

* These figures are estimates based on Khuzestan Water and Power Authority's data.

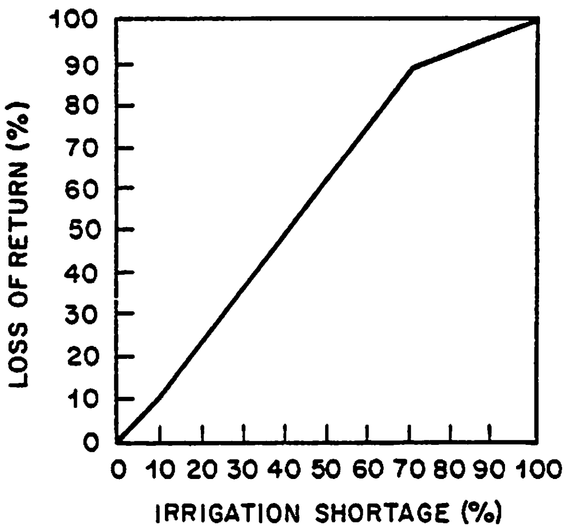

Figure 5. Penalty Function for Shortage of Water in the Three Irrigation Areas

TABLE 2. MONTHLY DISTRIBUTION OF ANNUAL TOTAL ENERGY OUTPUT

Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Total
Percent Monthly Requirement	8.2	8.1	8.3	8.2	7.5	7.3	7.7	8.3	8.9	9.1	9.3	9.1	100.0

(Source: Khuzestan Water and Power Authority)

TABLE 3. NUMERICAL DATA FOR KHUZESTAN WATER RESOURCES

Darn	Head (m)	Turbine Efficiency	Water Flow (m^3/mo)	Max Energy Output (Mw hr/mo)	Rated Power Capacity Mw
Dez	50	.85	1240×10^6	3000×10^2	520
Karun	45	.85	1500×10^6	3500×10^2	1000
Karun No. 2	40	.85	800×10^6	2100×10^2	4000

(a) Dams Specifications

River	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Dez y_1	51	224	365	279	306	760	915	634	403	238	151	104
Karun y_2	50	200	250	300	450	500	650	400	350	200	120	100

(b) Monthly River Flow (m^3/sec)

(Source: Khuzestan Water and Power Authority)

$$F = \sum_{k=1}^{N-1} R(k) = \sum_{k=1}^{N-1} [R_p(k) - R_i(k)] \tag{23}$$

where N represents the total number of stages in the time interval under consideration.

The optimal scheduling problem can be stated as follows: obtain a sequence of decision variables $u_1(k)$,$u_2(k)$ and $u_3(k)$ which would maximize the total return function (23) subject to the constraints (13) through (18) over a twelve month period. As discussed before, the method considered for this optimization problem is an approximate approach based on dynamic programming, *i.e.*, DDDP.

9.4.3 Computational Results

The DDDP algorithm discussed before was coded on the digital computer using FORTRAN language. A main program and seven subroutines were written. The computer simulations were done on an IBM 370/135 and an IBM 360/75 computers. In this study, in addition to the constraints considered by equations (16)-(18) and Tables 1-3, the desired initial and final state vectors were chosen to be

$$s(1) = s(13) = \begin{bmatrix} 1500 \\ 700 \\ 400 \end{bmatrix} MCM. \tag{24}$$

The initial state trajectories were assumed to be constant for all twelve months of the year. The irrigation return percentage was chosen to be $x(k) = 0.30$ for all k. Using these data and parameters, three near-optimal trajectories were obtained for the three reservoirs after ten iterations. Typical results are shown in Figure 6. Two other initial nominal trajectories were tried and the same optimal release policies were obtained. The trajectories in this Figure, except for reservoir 1, are constant with time. This is because the input into reservoir 2 is sufficient to meet the irrigation demands of Shushtar area which are rather stable. Therefore, the system does not use the stored water, and keeps the storage at maximum simply because the effective head of the power plants is assumed proportional to the square root of storage. Trajectory of reservoir I shows some fluctuation. This is in response to the higher percentage of target annual diversion for irrigation during months of October to the end of March for Haft-Tapeh irrigation area. Had the other demands (irrigation and power) been chosen more realistically, (*i.e.*, if monthly demand fluctuations had been better projected in Tables 1 and 2) the storage of the reservoirs would have been used more efficiently.

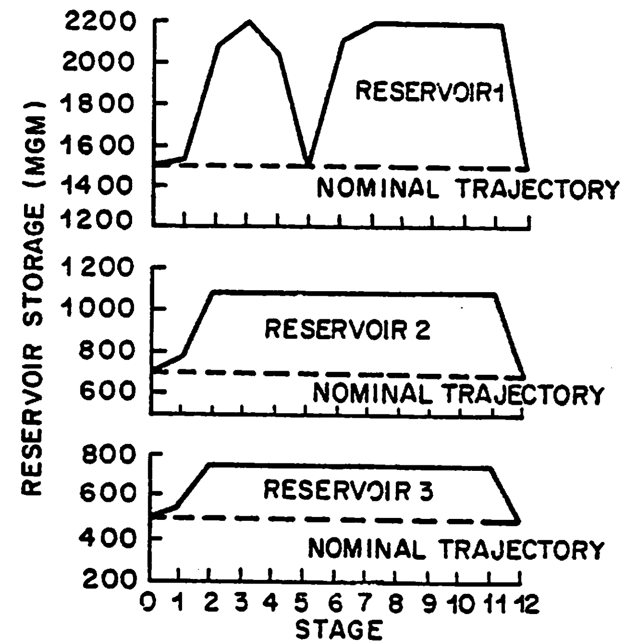

Figure 6. Typical Near-Optimum State Trajectories for the Three Reservoirs

Figure 7 is a plot of total return versus the number of iterations for this system. It shows that after 12 iterations the return has increased from the nominal return of 7182×10^6 Rials to 8150×10^6 Rials. Since in iterations 13, 14, etc., no major improvement in total return was observed, it was concluded that return of iteration 12 represents the near-optimal solution of this study.

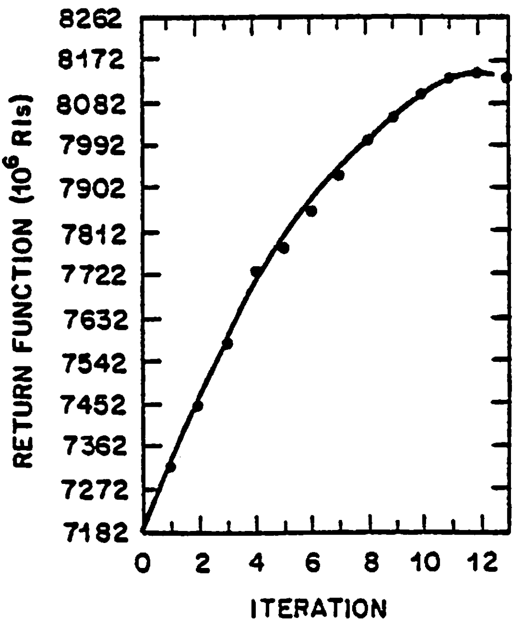

Figure 7. Convergence of Near-Optimum Returns (Million Rials) vs Number of Iterations for the Khuzestan System

9.5 A HYDRAULIC LEVEL SYSTEM

In this section a hydraulic level control system is described and its stability is studied. The problem is in part discussed by Dorf [9.34] also.

Let us consider the hydraulic system of Figure 1.3.4. The inherent delay in this system is the transport time for the liquid to travel a distance d with a velocity v, *i.e.* the delay is

$$h = d/v \tag{1}$$

This delay is, as can be seen from Figure 1.3.4, between the liquid output and the valve adjustment. A block diagram of this system is given in Figure 1. As seen, the open-loop transfer function is given by

$$G(s)H(s) = G_a(s)G_t(s)G_f(s)e^{-sh} \tag{2}$$

Where the actuator, tank and float transfer functions are, respectively

$$\begin{aligned} G_a(s) &= \frac{10}{s+1} \\ G_t(s) &= \frac{K}{s + 1/30} \\ G_f(s) &= \frac{9}{s^2 + 3s + 9} \end{aligned} \tag{3}$$

and K is the dc gain whose nominal value is 1.05. The nominal value of the delay is assumed to be $h=1$.

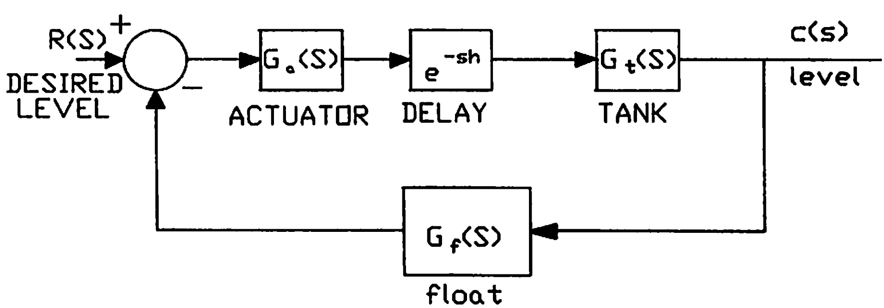

Figure 1. Block Diagram of a Hydraulic Level Control System

Here, the stability of the system is studied through frequency response. The Bode plots of this system without time delay (assuming that it is negligible) is shown in Figure 2. As seen, the non-delay system has a gain crossover frequency at $\omega_2 = 0.8$ and the phase margin is 40°. This indicates that the system with time delay is stable. Figure 3 shows the Bode diagram for the hydraulic system with delay. As shown, the system has now a phase margin of -3° which indicates that the system is unstable. One possible way of stabilizing the system is to reduce the gain of the system to have a phase margin for a reasonable stability margin. For example, a reduction of 5 dbs, leading to a gain of $K' = 1.05/1.78 = 0.59$.

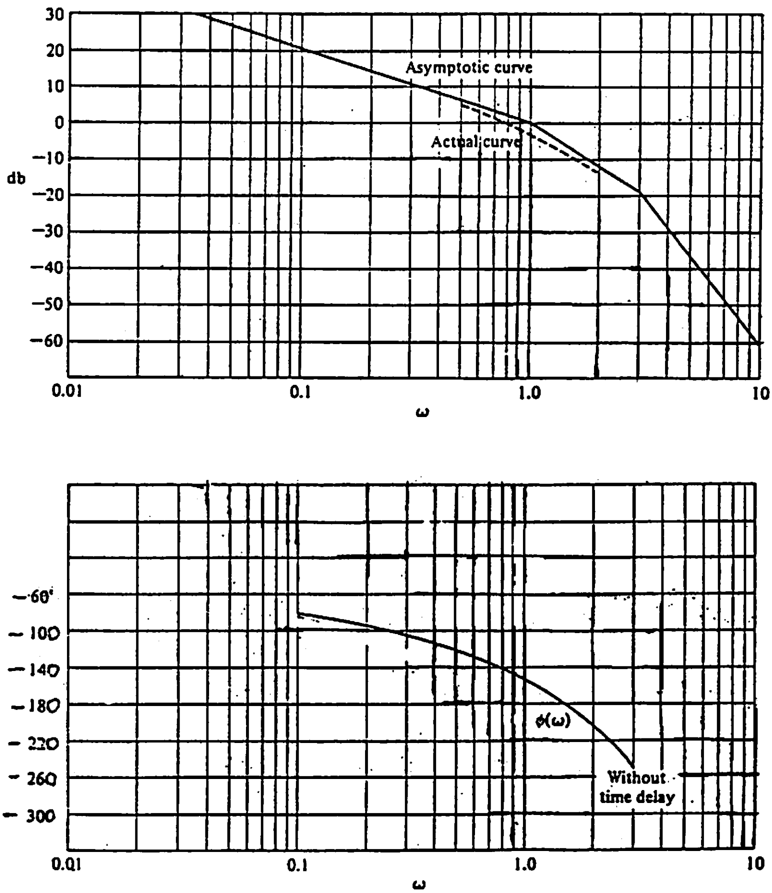

Figure 2. Bode Diagram of the Hydraulic System Without Delay

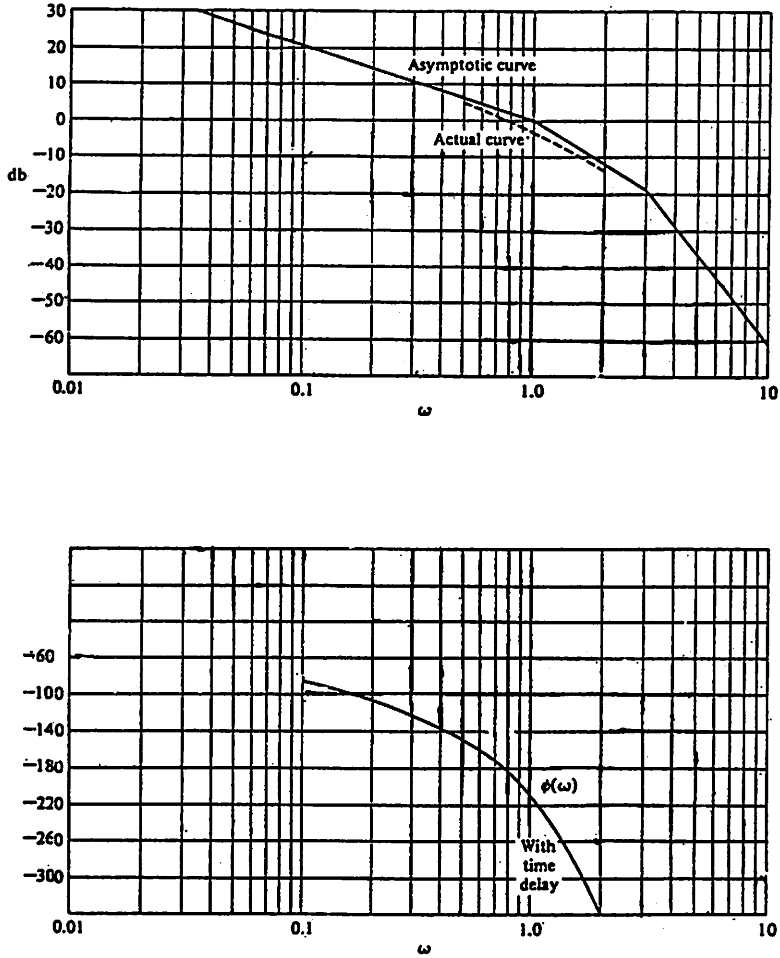

Figure 3. Bode Diagram of the Hydraulic System With Delay

In this simple problem, it is deduced that the addition of a time delay introduces an additional phase lag, resulting in a less stable system. Thus, when time delays can not be avoided, a reduction in loop gain would help stabilize the system. The price to pay here would be an increase in the steady-state error.

9.6 CONTROL OF TIME DELAY SYSTEMS VIA SMITH PREDICTOR

The basic conclusion of the previous section in dealing with a hydraulic time-delay system was that the introduction of delays in a tuned system would require a reduction in gain to maintain stability. Smith Predictor [9.35, 9.36] control scheme can be used to avoid this dilemna provided that the model parameters match those of the plant exactly. To accommodate this matching process, an adaptive control scheme can be added to the Smith Predictor. In this section a brief introduction to the structure and stability of the Smith Predictor is presented.

9.6.1 The Smith Predictor

Consider a conventional time-delay control system shown in Figure 1. Here the notation of Marshall [9.35] has been used to represent shorthand purpose, *e.g.* $R(s)$ is represented by R,G_c is the controller $G_c(s)$; D represents disturbance $D(s)$; G_p is the plant transfer function $G_p(s)$; H_p represents plant delay $H_p(s)$, and C_p is the output of the plant. If the fictitious variable B in Figure 1 were measurable, then it could be fed back to the cascade controller, thereby moving the time delay outside of the control loop. Under such condition, the system output would be delayed value of the delay-free portion of the system. This would improve the response of the system. However, this separation of delay from the system is hardly possible since the delay is often distributed and not lumped and there is no viable reason to place the delay after the plant rather than before it.

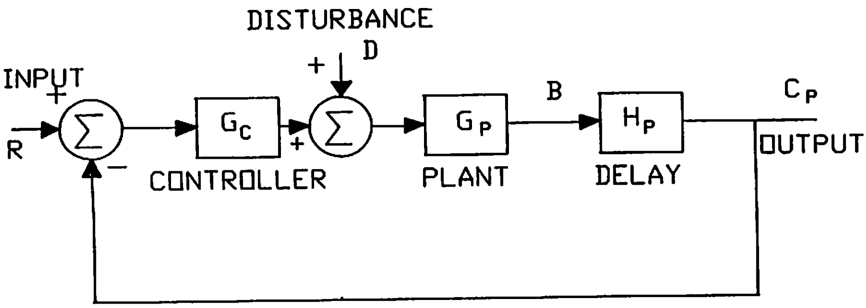

Figure 1. A Typical Time-Delay System

In order to improve the time-delay system's design, one can model the plant as indicated by Figure 2. Here, G_m represents the plant's dynamic model, H_m is the time-delay's model and E is the error between the plant output and the model output. Although the fictitious variable B is not available, B_m can, however, be used for feedback purpose as shown in Figure 2. This scheme would only control the model well, but not the system output C_p. This would indicate that both inaccurate models and load disturbances can not be accommodated. In order to compensate for these errors, a second feedback loop is introduced as shown in Figure 3. This is the Smith Predictor control scheme for TD systems. In this way, the effect of the time delay in the feedback loop is minimized allowing one to use conventional controllers such as PD, PI or PID for G_c.

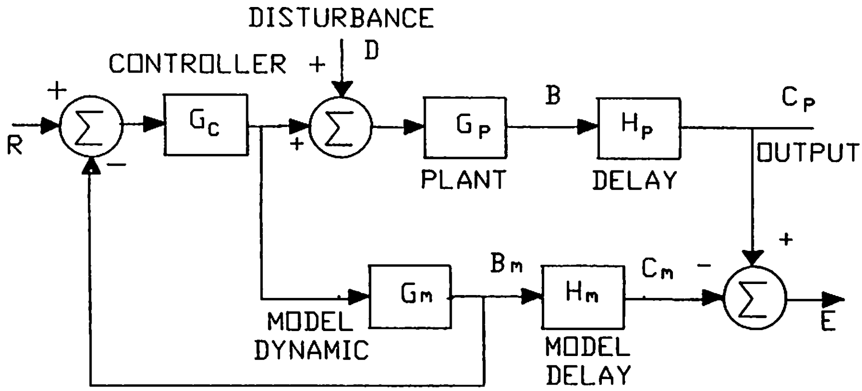

Figure 2. A Preliminary Form of the Smith Predictor

An alternative representation of the Smith Predictor is shown in FIgure 4. Assuming that disturbance $D=0$, the closed-loop transfer function of the system is given by

$$\frac{C_p(s)}{R(s)} = \frac{G_c G_p H_p}{1 + G_c G_m - G_c G_m H_m + G_c G_p H_p} \tag{1}$$

If $G_p = G_m$ and $H_p = H_m$, the above equation reduces to

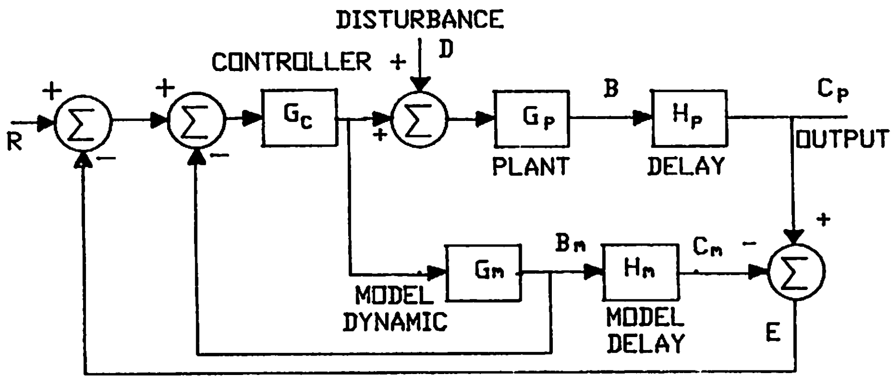

Figure 3. A Complete Smith Predictor TD Control System

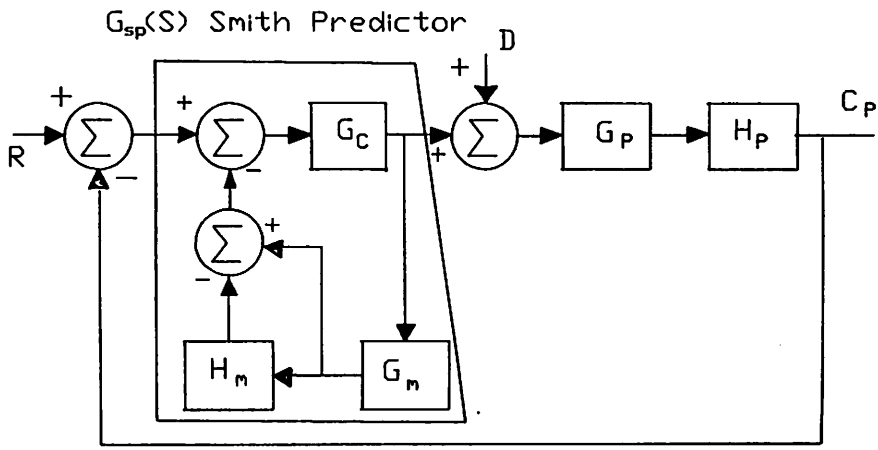

Figure 4. An Alternative Schematic for the Smith Predictor

$$\frac{C_p(s)}{R(s)} = \frac{G_c G_p H_p}{1 + G_c G_m} \tag{2}$$

As it can be seen from the above relation, the effect of the delay has been eliminated from the denominator of the transfer function.

9.6.2 An Adaptive Smith Predictor TD Control System

It is very common to have plant deteriorations due to age, wear, fatigue, temperature, disease, etc. Such variations result in changes in the plant behavior calling for a subsequent change in the controller. In this case, let us assume that the delay itself in the TD system changes. It is desired to device a Smith Predictor which adapts itself to such changes. In this section, a model reference adaptive control scheme discussed by Landau [9.37] for TD systems is presented.

A special form of the model reference adaptive control consists of a time-varying plant with adjustable parameters and is parallel to an accurately defined model [9.36]. In this manner, the system's parameters are changed to make the controlled system behave like the parallel model. This close description to a Smith Predictor makes it applicable for adaptive time-delay control system.

In the present consideration, a performance criterion and a scheme to minimize it need to be chosen. In accordance to the works of Marshall [9.35] and Landau [9.37], a reasonable performance criterion is integral square error (ISE) of the system and model outputs, *i.e.*

$$J = 1/2 \int e^2(t)dt \tag{3}$$

where $e(t) = C_p(t) - C_m(t)$. The minimization schemes are often a gradient-based method such as conjugate gradient and Davidon-Fletcher-Powell, to name two. Let the magnitude of the delay be represented by h_p with its initial value, h_{po}. Assume that the delay has been changed by a small amount δh_p. In the model reference adaptive control the system's gain is changed to minimize the cost criterion J. For the case of Smith Predictor, the model is modified such that it would track the time delay, thereby minimizing the cost. If we utilize a gradient method, the upgrading relation is

$$\delta h_m = -k \frac{\partial J}{\partial h_m} \tag{4}$$

Where k is a gain or step size. In order to avoid mathematical complexities, it is assumed that plant delay h_p changes continuously, but the model delay h_m would change only at discrete intervals. The procedure has, thus, two phases - one to identify the new delay and the other is to control the plant. The change in h_p is given by

$$\delta h_p = -k\frac{\partial J}{\partial h_p} \tag{5}$$

The rate of change of the delay is

$$hp = \frac{dhp}{dt} = \frac{d}{dt}(hpo + \delta hp) = \frac{d}{dt}\delta hp$$

h_p

Using (5), one has

$$hp = -k\frac{d}{dt}\frac{\partial J}{\partial hp}$$

Assuming that the adaptation is slow, one can interchange the order of differentiation,

$$\begin{aligned} hp &= -k\frac{\partial}{\partial h_p}\frac{dJ}{dt} = -k\frac{\partial}{\partial h_p}\frac{d}{dt}(1/2\int e^2(t)dt) \\ &= -\frac{k}{2}\frac{\partial^2 e}{\partial h_p^2} \end{aligned}$$

Using the chain rule, one obtains

$$hp = -ke\frac{\partial e}{\partial h_p}$$

and finally,

$$\delta hp = \int h_p^o dt$$

Now, the desired change in the model time delay can be completed as

$$\delta h_m = \delta h_p = \int h_p\, dt = -k\int e\frac{\partial e}{\partial h_p}dt \tag{6}$$

The problem, thus, reduces to one of finding $\partial e/\partial h_p$. Within a first approximation, C_m is not a function of h_p. This possible dependence, however, should be verified by simulation as suggested by Marshall [9.35]. Thus,

$$\frac{\partial e}{\partial h_p} = \frac{\partial C_p}{\partial h_p} \tag{7}$$

Using (7) in (6) one gets

$$\delta h_m = -k\int e\frac{\partial C_p}{\partial hp}dt \tag{8}$$

This last relation is somewhat desirable, because the latter part of the integrand $\partial C_p/\partial h_p$ represents a sensitivity function and can be easily computed. Using (1), we have

$$P(s) = \frac{C_p(s)}{R(s)} = \frac{G_c G_p H_p}{1 + G_c G_m - G_c G_m H_m + G_c G_p H_p} \tag{9}$$

calculating the partial derivative $\partial C_p/\partial hp$ and omitting complex frequency s argument results in

$$\frac{\partial C_p}{\partial h_p} = \frac{(-sG_c G_p H_p)(1 + G_c G_m - G_c G_m H_m)R}{(1 + G_c G_m + G_c G_p H_p - G_c G_m H_m)^2}$$

Now, using (9) twice,

$$\frac{\partial C_p}{\partial hp} = -sP(1 - P)R$$

A schematic to produce $\partial C_p/\partial h_p$ is shown in Figure 5. Here, it is assumed that the Laplace transform of $\partial C_p/dt$ is $\partial C_p(s)/\partial H_p(s)$. The pure differentiation shown in this figure can be digitally approximated over a given frequency range. The blocks labeled G_p and H_p in the lower Smith Predictor are assumed to contazin the plant parameter values, but they are unavailable. Thus, one needs to use G_m and H_m to update them as often as possible.

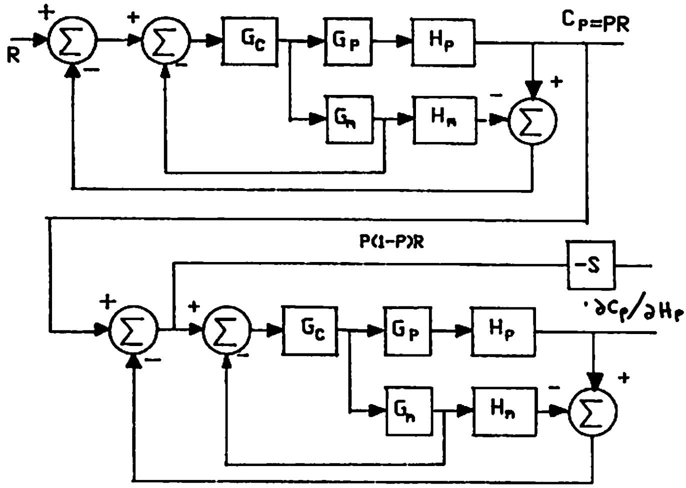

Figure 5. A Method of Generating the Sensitivity Function $\partial C_p/\partial h_p$ Based on [9.35]

The remaining value to be determined is the gain k in (8). From (6) and (8) one has

$$\delta hp = -k \int_{t_1}^{t_2} e \frac{\partial C_p}{\partial hp}\, dt \tag{10}$$

The error, e is a function of the plant time delay. For a small change in the time delay, $\delta\, hp$, using the definition of differential and assuming that $e_o = 0$, *i.e.* $\delta e = e$, one gets

$$e = \delta hp \frac{\partial e}{\partial hp} \tag{11}$$

Substituting (11) and (7) into (10) yields

$$\delta hp = -k \int_{t_1}^{t_2} \delta hp \frac{\partial C_p}{\partial h_p} \frac{\partial C_p}{\partial hp}\, dt$$

Now, removing δhp from inside the integral and dividing both sides by $k\ \delta hp$ results

$$1/k = -\int_{t_1}^{t_2} |\partial C_p/\partial h_p|^2 dt \tag{12}$$

Since $\partial C_p/\partial h_p$ is usually available, the constant k can be precomputed. This would now complete the algorithm for changing the model time delay:

$$\delta\theta_m = -k \int_{t_1}^{t_2} e \frac{\partial C_p}{\partial h_p} dt$$

This algorithm implementation is shown in Figure 6.

Marshall [9.35] has simulated the system of Figure 8 on a digital computer using

$$G_p(s) = \frac{1}{a(s+a)}, \; H_p(s) = e^{-s\tau_p}$$

where τ_p = 2.0 seconds and α = 0.5 and $G_c(s)$ = 0.5. The adaptive control of the delay was achieved for a range of k. The system, thus simulated, was stable and had satisfactory dynamics when the initial model delay was off by as much as 80%. This adaptation of the delay is shown in Figure 7. As seen, the nominal delay of 2 seconds was achieved for offsets in model delay of up to ± 80%. Marshall [9.35] has tested the adaptive control algorithm for a wide range of computational facilities such as hybrid and analog computers with both reduced and full sensitivity models.

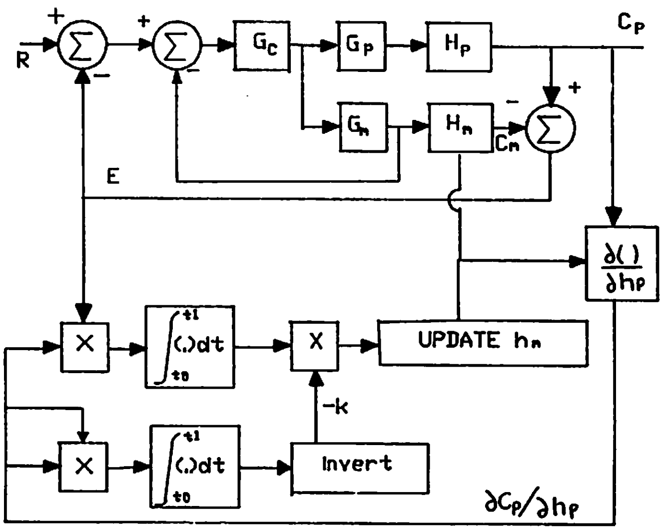

Figure 6. An Adaptive Controller for a TD System with a Smith Predictor [9.35]

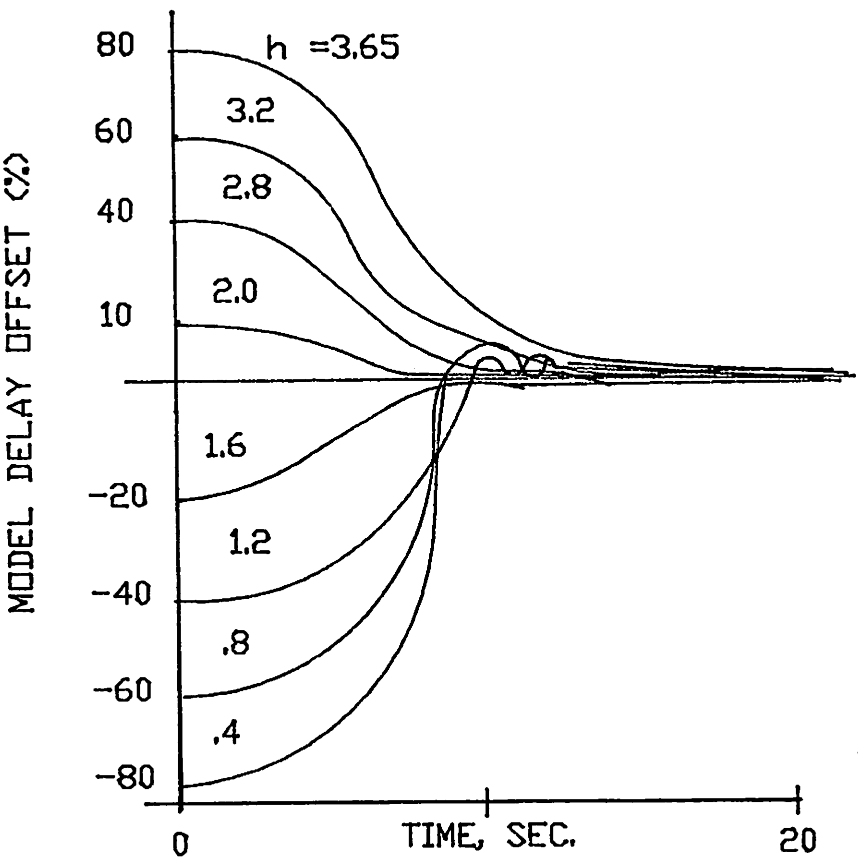

Figure 7. Model Delay Offset in Percent Versus Time for Adaptive Controller and a Smith Predictor

REFERENCES

[9.1] R. A. Phillips, "Analysis of Tadem Cold Reduction Mill with Automatic Gauge Control," *AIEE Trans.*, pp. 355-363, January, 1957

[9.2] A. J. Durelli, E. A. Phillips and C. H. Tsao, *Introduction to the Theoretical and Experimental Analysis of Stress and Strain,* McGraw Hill Co., 1958

[9.3] J. H. Courcoulas and J. M. Ham, "Incremental Equations for Tandem Rolling Mills," *AIEE Trans.*, pp. 363-371, January, 1957

[9.4] F. G. Johnson, J. H. Stoltz and R. G. Chapman, Jr., "Electrical Features of Combinational-Temper-Reduction Mills," *Iron and Steel Engrs., Journal,* pp. 125-134, October, 1965

[9.5] H. W. Smith, "Dynamic Control of a Two-Stand Cold Mill, *JFAC J. Automatica,* Vol. 5., pp.183-190, 1969

[9.6] C. Hadlock, M. Jamshidi and P. V. Kokotovic', "Near-Optimum Design of Three Time-Scale Systems," *Proc. 4th Princeton Conference on Info. Science and Systems,* Princeton University, Princeton, N.J., March, 1970

[9.7] P. V. Kokotovic', W. R. Perkins, J. B. Cruz, Jr and G. D'Ans, "ϵ Coupling Method for Near-Optimum Design of Large-Scale Linear System," *Proc-IEE,* Vol. 116, No. 5, pp. 889-892, May, 1969

[9.8] M. Jamshidi, G. C. D'Ans and P. V. Kokotovic', Application of a Parameter - Imbedded Riccati Equation, *IEEE Trans. Aut. Cont.,* Vol. AC-15, No. 6, pp. 682-683, December, 1970

[9.9] M. Jamshidi, "A Near-Optimum Controller for Cold Rolling Mill," *Int. J. Contr.,* Vol. 16, pp. 1137-1154, 1972

[9.10] E. A. Randich, "Allegheny Ludlum's High-Speed Three-Stand Tandem Cold Mill, *Iron and Steel Engr.,* October, 1968, pp. 101-121

[9.11] M. Jamshidi, "An Integrated Near-Optimum Design of Cold Rolling Mills," *CSL Report R-499,* Univ. Illinois, Urbana, Il., 1971

[9.12] E. Orowan, "The Calculation of Roll Pressure in Hot and Cold Flat Rolling," *Proc. Inst. Mech. Engr.*, Vol. 150, pp. 140-167, 1943

[9.13] D. P. Bland and H. Ford, "The Calculation of Roll Force and Torque in Cold Strip Rolling with Tensions," *Proc. Inst. Mech. Engrs.*, Vol. 155, pp. 144-163, 1948

[9.14] A. H. El-Waziri, "An Up-to-Date Examination of Rolling Theory," *Iron and Steel Engrs.*, pp. 73-80, October, 1963

[9.15] M. Jamshidi and P. V. Kokotovic', "Dynamic Model of a Continuous Cold Rolling Mill," *CSL Report R-459*, Univ. of Illinois, Urbana, February 1970

[9.16] P. Sannuti, "Near-Optimum Design of Time-Lag Systems by Singular Perturbation Method," *Proc. JACC*, Atlanta, GA., pp. 489-496, June, 1970

[9.17] K. Inoue, H. Akashi, K. Ogino, and Y. Swargi, "Sensitivity ApproaChes to Optimization of Linear Systems with Time Delay," *IFAC J. Automatica*, Vol. 7, pp. 671-679, 1971

[9.18] P. Sannuti and P. V. Kokotovic', "Near-Optimum Design of Linear Systems by a Singular Perturbation Method," *IEEE Trans. Aut. Contr.*, Vol. AC-14, pp. 15-21, 1969

[9.19] G. C. D'Ans, "An Embedding Approach for Solving Riccati Equation," *CSL Report* R-428, September, 1969

[9.20] P. V. Kokotovic' and G. C. D'Ans, "Parameter Imbedding Design of Linear Optimal Regulators," *Proc. 3rd Princeton Conference on Info. Science and Systems*, Princeton University, Princeton, NJ pp. 378-379, March, 1969

[9.21] R. E. Kalman, "Contributions to the Theory of Optimal Control," *Bal. Soc. Math.*, Mexico, 2nd Ser., Vol. 5, pp. 102-118, 1960

[9.22] M. Heidari, V. T. Chow, P. V. Kokotovic', D. D. Meredith, "The Discrete Differential Dynamic Programming Approach to Water Resources Systems Optimization," *Water Resources Research* Vol. 7, pp. 273-282, 1971

[9.23] M. Jamshidi and M. Heidari, "Application of Dynamic Programming to Control Khuzestan Water Resources Systems," *IFAC J. Automatica*, Vol. 13, pp. 287-293, 1976

[9.24] B. D. Gardner, Y. Madhi, S. Partovi, M. Hendizadeh, and M. Saghatoleslami, "Price of Irrigation Water in Iran," *Water Resources Research*, Vol. 10, pp.1080-1084, 1974

[9.25] R. E. Larson and W. G. Keckler, "Applications of Dynamic Programming to Control of Water Resources Systems," *IFAC J. Automatica*, Vol. 5, pp. 15-26, 1969

[9.26] A. Maas, M.M. Hufschmidt, R. Dorfman, H. E. Thomas, S. A. Marglin and G. M. Fair, "Design of Water Resources Systems," *Harvard University Press*, 1962

[9.27] H. Tamura, "Decentralized Optimization for Distributed Lag Models of Discrete Systems," *IFAC J. Automatica*, Vol. 11,pp. 593-602, 1975

[9.28] M. Jamshidi, *Large-Scale System - Modeling and Control*, Elsevier, North-Holland, 1983

[9.29] F. Fallaside and P. Perry, "Hierarchical Optimization of a Water Supply Network," *Proc. IEE*, Vol. 122, pp. 202-2-8, 1975

[9.30] M. G. Singh, *Dynamic Hierarchical Control*, Rev. Ed., North-Holland, 1980

[9.31] P. Dorato and A. H. Levis, "Optimal-Linear Regulators: The Discrete-Time Case," *IEEE Trans. Aut. Cont.*, Vol. AC-16, pp.613-620, 1966

[9.32] F. V. Webster and B. M. Cobb, "Traffic Signals." Road Research Technical Paper 56, Her Majesty's Stationary Office, London, 1966

[9.33] M. G. Singh and H.Tamura, "Modeling and Hierarchical Optimization for Oversaturated Urban Road Traffic Networks," *Int. J. Contr.*, Vol. 20, pp. 913-934, 1974

[9.34] R. C. Dorf, "*Modern Control Systems*," Addison-Wesley Publishing Co., 1980

[9.35] J. E. Marshall, "*Control of Time-Delay Systems,*" Peter Peregrinus, Ltd., Stevenage, England 1979

[9.36] A. T. Bahill, *A Simple Adaptive Smith-Predictor for Controlling Time-Delay Systems,*" *IEEE Control Systems Magazine,*" Vol. 3, No. 2, May, 1983, pp.16-22.

[9.37] I. D. Landau, "*Adaptive Control: The Model Reference Approach,*" *Marcel Dekker, Inc., 1979.*

APPENDIX A

REVIEW OF LINEAR ALGEBRA

A.1 INTRODUCTION

This appendix is concerned with the review of fundamental topics in linear algebra which are of direct or indirect use in the study of time-delay systems. In order to keep the size of this appendix within reason, some background knowledge on this subject is assumed on the part of the reader. Thus we dispense with some elementary definitions. Also the proofs of some theorems may be omitted or may be left as exercises to the reader.

We will start with some basic definitions and concepts. Then we will review linear transformations and functions of matrices.

A.2 FIELDS AND VECTOR SPACES

1. Definition. A field F consists of a set F whose elements are called scalars, two operations called addition "+" and multiplication "." and two distinguished elements 0 and 1. The set is closed under addition and multiplication. Also, the two operations are commutative and associative and multiplication is distributive *w.r.t.* addition. That is, we have the following six axioms:

a) $\forall \alpha \epsilon F$, $\alpha+0 = \alpha$, $\alpha.1 = \alpha$*

b) $\forall \alpha \epsilon F \ \exists \ \beta \epsilon F \ni \alpha+\beta = 0$. Then β is indicated $-\alpha$.

c) $\forall \alpha \epsilon F$, $\alpha \neq 0 \ \exists \ \gamma \epsilon F \ni \alpha.\gamma = 1$. Then γ is indicated α^{-1}.

d) $\alpha, \beta \epsilon F \implies \alpha+\beta = \beta+\alpha \epsilon F$, $\alpha.\beta = \beta.\alpha \epsilon F$.

---

* The symbol "." is often dropped for convenience.

e) $\alpha,\beta,\gamma\epsilon F \implies \alpha+(\beta+\gamma) = (\alpha+\beta)+\gamma,\ \alpha.(\beta.\gamma) = (\alpha.\beta).\gamma.$

f) $\alpha,\beta,\gamma\epsilon F \implies \alpha\,.\,(\beta+\gamma) = \alpha.\beta+\alpha.\gamma.$

2. Example. The set of all real numbers forms a field called the *real field* denoted by R. Also, the set of all complex numbers forms a field called the *complex field* denoted by C. However, the set of all positive numbers does not form a field because axiom (b) does not hold.

3. Definition. A *vector space* (also referred to as a *linear space* or a *linear vector space)* V over a field F, denoted by (V,F), or simply by V when no ambiguity might arise, is a set of elements called *vectors* and two operations called *vector addition* and *scalar multiplication* satisfying the following nine axioms:

a) $\mathbf{x},\mathbf{y}\epsilon V \implies \mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}\ \epsilon V.$

b) $\mathbf{x},\mathbf{y},\mathbf{z}\epsilon V \implies \mathbf{x}+(\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}.$

c) $\forall \mathbf{x}\epsilon V \exists$ a unique vector $\mathbf{0}\epsilon V \ni \mathbf{x}+\mathbf{0} = \mathbf{x}.$

d) $\forall \mathbf{x}\epsilon V \exists$ a unique vector $\mathbf{y}\epsilon V \ni \mathbf{x}+\mathbf{y} = \mathbf{0}.$

e) $\mathbf{x}\epsilon V \implies 1.\mathbf{x} = \mathbf{x}$ where $1\epsilon F.$

f) $\alpha\epsilon F,\ \mathbf{x}\epsilon V \implies \alpha.\mathbf{x}\epsilon V$

g) $\alpha,\beta\epsilon F,\ \mathbf{x}\epsilon V \implies \alpha.(\beta.\mathbf{x}) = (\alpha.\beta).\mathbf{x}$

h) $\alpha\epsilon F,\ \mathbf{x},\mathbf{y}\epsilon V \implies \alpha.(\mathbf{x}+\mathbf{y}) = \alpha.\mathbf{x}+\alpha.\mathbf{y}$

i) $\alpha,\beta\epsilon F,\ \mathbf{x}\epsilon V \implies (\alpha+\beta).\mathbf{x} = \alpha.\mathbf{x} + \beta.\mathbf{x}$

4. Example. Consider a field F. Let F^n denote all n-tuples of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix},\ x_i\epsilon F\ ,\ i = 1,2,\ldots,n. \tag{1}$$

If vector addition and scalar multiplication are defined as

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n+y_n \end{bmatrix}, \ x_i, y_i \epsilon F \ , i = 1,2, \ldots, n \tag{2}$$

and

$$\alpha.\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha x_n \end{bmatrix}, \ \alpha \epsilon F, \ x_i \epsilon F, i = 1,2, \ldots, n, \tag{3}$$

then F^n forms a vector space over F, (F^n,F). Examples of this vector space are (R^n,R) , the space of n-dimensional real vectors (the n-dimensional Euclidean space) and (C^n,C), the space of n-dimensional complex vectors.

5. Definition. Let (V,F) be a vector space and let $W \subset V$. Then (W,F) is called a *subspace* of (V,F) if (W,F) is a vector space.

6. Example. The real vector space (R^n,R) is a subspace of real vector space (R^{n+m},R) for any integers $m,n \geqslant 1$. Also, the real vector space (R^n,R) is a subspace of the complex vector space (C^n,C).

A.3 LINEAR INDEPENDENCE AND RANK

1. Definition. A set of vectors $\mathbf{v}_1,\mathbf{v}_2, \ldots, \mathbf{v}_n$ belonging to a vector space (V,F) are said to be *linearly independent* if $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = 0$, where $\alpha_1, \alpha_2, \ldots, \alpha_n \epsilon F$ implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. If at least one $\alpha_i \epsilon F$ is nonzero, then the vectors are said to be *linearly dependent*. The sum $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ is referred to as a *linear combination* of vectors $\mathbf{v}_1,\mathbf{v}_2, \ldots, \mathbf{v}_n$. Δ

Note that if $\mathbf{v}_i = 0$ for any $i=1,2, \ldots, n$, then the vectors $\mathbf{v}_1,\mathbf{v}_2, \ldots, \mathbf{v}_n$ will be linearly dependent.

2. Definition. An $m \times n$ matrix A has *rank* r , denoted $\rho(A) = r$, if and only if it has at least one nonsingular $r \times r$ submatrix but has no nonsingular submatrix of order higher than r.

3. Definition. An $m \times n$ matrix A is said to have *full rank* if $\rho(A) = \min(m,n)$. Δ

Note that a square matrix has full rank if and only if it has a nonzero determinant.

The following theorem relates the concepts of linear independence and rank.

4. Theorem. Let vectors $v_1, v_2, \ldots, v_k$ in $(C^n C)$, where $k < n$, form columns of matrix A. Then these vectors are linearly independent if and only if $\rho(A) = k$. Δ

A.4 BASIS AND DIMENSION

1. Definition. A set of vectors is said to *span* a vector space V if any vector in V can be expressed as a linear combination of vectors in that set.

2. Definition. A set of vectors in a vector space V is said to form a *basis* of V if

a) they span V, and

b) they are linearly independent.

Consider a vector space (V,F) with a basis $\{e_1, e_2, \ldots, e_n\}$. Then for any vector $v \epsilon (V,F)$, unique elements $\alpha_1, \alpha_2, \ldots, \alpha_n \epsilon F$ exist such that V has a representation

$$v = \sum_{i=1}^{n} \alpha_i e_i \tag{1}$$

(see Problem A.2). The scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ are then called the *components* of V *w.r.t.* the basis $\{e_1, e_2 \cdots, e_n\}$.

3. Example. The vectors $e_1 = (2,0,0)'$, $e_2 = (0,1,1)'$, $e_3 = (-1,2,0)'$ and $x = (-3,0,4)'$ span R^3, however, they do not form a basis for R^3. The vectors e_1, e_2 and e_3 form a basis for R^3 since they are linearly independent. The vector x will then have components $\alpha_1 = -2.5$, $\alpha_2 = 4$ and $\alpha_3 = -2$ *w.r.t.* this basis because

$$x = \sum_{i=1}^{3} \alpha_i e_i. \quad \Delta \tag{2}$$

It can be proved [A.2] that all bases of a vector space have the same number of elements. This number is called the *dimension* of the vector space.

4. Definition. A vector space is said to be *finite dimensional* if its bases have a finite number of elements. Otherwise, it is an *infinite - dimensional* vector space.

A.5 LINEAR OPERATORS

1. Definition. Let X and Y be linear spaces and let A be a function with domain $D(A)$ in X and range (or range space) $R(A)$ in Y. Then Y is called a *linear operator* if $D(A)$ is a subspace of X and if

$$A(\alpha \mathbf{x}_1+\beta \mathbf{x}_2) = \alpha A(\mathbf{x}_1)+\beta A(\mathbf{x}_2) \quad . \tag{1}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are vectors in $D(A)$, and α and β are scalars. Δ

Note that $R(A)$ is a subspace of Y. If $D(A)=X$, we say that A is a linear operator on (or from) X into Y. It is customary to write $A\mathbf{x}$ instead of $A(\mathbf{x})$ whenever it is convenient.

Let X and Y be both finite-dimensional spaces, *i.e.*, let $Y = R^m$ if $X = R^n$, and $Y = C^m$ if $X = C^n$ where m and n are arbitrary positive integers. Then the $m \times n$ matrix

$$(a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2}$$

defines a linear operator A on X into Y as follows: if $\mathbf{x} = (x_1,x_2, \ldots ,x_n)'$, $\mathbf{y} = (y_1,y_2, \ldots ,y_m)'$ and

$$\sum_{j=1}^{n} a_{ij}.x_j = y_i \ , \ i=1,2, \ldots ,m \ , \tag{3}$$

then $A\mathbf{x} = \mathbf{y}$. Conversely, $A\mathbf{x} = \mathbf{y}$ means that the y_i's are defined in terms of the x_j as in (3). The following theorem can be readily proved.

2. Theorem. The rank of matrix A is equal to the dimension of its range space.

3. Definition. Consider a matrix A. Then vectors $x \epsilon (V,F)$ such that $Ax = 0$ form a vector space $N(A)$, a subspace of (V,F), called the *null space* or the *kernel* of matrix A , *i.e.*,

$$N(A) = \{x \epsilon V \mid Ax = 0\} \tag{4}$$

4. Definition. The dimension of the null space of matrix A is called the *nullity* of A denoted by $\gamma(A)$.

5. Definition. A subspace W is said to be *invariant under* A if $x \epsilon W$ implies $Ax \epsilon W$. Δ

Some properties of range space and null space are summarized in Problems A.3 and A.4.

6. Definition. A vector space V is said to be the *direct sum* of two subspaces W_1 and W_2 , denoted $V = W_1 \oplus W_2$, if W_1 and W_2 are subspaces of V and if every $x \epsilon V$ can be uniquely represented as $x = x_1 + x_2$ where $x_1 \epsilon W_1$, and $x_2 \epsilon W_2$. (See Problem A.5.)

7. Example. Vector space R^3 is the direct sum of the two subspaces R^1 and R^2, *i.e.*,

$$R^3 = R^1 \oplus R^2 \tag{5}$$

A.6 INNER PRODUCT AND NORM

1. Definition. The *inner product* of two vectors $x,y \epsilon (V,F)$ is a complex number denoted by (x,y) satisfying the following properties:

a) $(x,x) > 0 \ \forall x \neq 0 \epsilon V$ (1a)

b) $(x,x) = 0 \iff x = 0 \epsilon V$ (1b)

c) $\overline{(x,y)} = (y,x) \ \forall \ x,y \epsilon V.$ (1c)

d) $(\alpha x + \beta y, z) = \bar{\alpha}(x,z) + \bar{\beta}(y,z) \ \ \forall \ x,y,z \epsilon V \ \alpha,\beta \epsilon F.$ (1d)

where bar denotes the complex conjugate. An inner product is sometimes also referred to as a *scalar product* or a *dot product*.

2. Example. An inner product for the vector space (C^n, C) is

$$(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} \bar{x}_i y_i \tag{2}$$

where x_i and y_i are the components of x and y, respectively, *w.r.t.* some basis. It can be easily verified that the above inner product satisfies the conditions of Definition 1.

3. Definition. A vector space (V,F) is called an *inner-product space* if it is possible to define an inner product associating a scalar (x,y) with every pair of vectors $\mathbf{x},\mathbf{y} \in V$.

4. Definition. Two vectors x, and y are said to be *orthogonal* if

$$(\mathbf{x},\mathbf{y}) = 0. \tag{3}$$

A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are said to be *orthonormal* if

$$(\mathbf{x}_i,\mathbf{x}_j) = \delta_{ij}, \quad i,j = 1,2,\ldots,n \tag{4}$$

where δ_{ij} is the *Kronecker delta, i.e.*

$$\delta_{ij} = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases} \tag{5}$$

5. Example. Vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in $(R^n$5R) where

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \leftarrow \textit{ith position} \tag{6}$$

are orthonormal. Further, they are linearly independent and form a basis for (R^n,R), called an *orthonormal basis.*

6. Definition. Consider a matrix A. The *conjugate transpose* of A, denoted by $\mathbf{A}^*$, is another matrix defined by

$$(\mathbf{Ax},\mathbf{y}) = (\mathbf{x},\mathbf{A}^*\mathbf{y}), \quad \forall x,y \tag{7}$$

Note that if **A** is a real matrix, then its transpose and its conjugate transpose will be the same, *i.e.*, $\mathbf{A}' = \mathbf{A}^*$. Some of the properties of conjugate transpose are given in Problem A.8.

7. Definition. Matrix **A** is said to be *symmetric* if $\mathbf{A} = \mathbf{A}'$, *skew symmetric* if $\mathbf{A} = -\mathbf{A}'$, *hermitian* if $\mathbf{A} = \mathbf{A}^*$, *normal* if $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$, and *unitary* or *orthogonal* if $\mathbf{A}^*\mathbf{A}=\mathbf{I}$ (identity matrix).

8. Definition. Consider a vector space (V,F) where $F = R$ or $F = C$. It is said to be a *normed vector space* if there exists a function, denoted by $\|.\|$ and called a *norm* on V, which maps V into R_+ (the set of nonnegative real numbers) and satisfies the following postulates:

a) $\|\mathbf{x}\| = 0 \iff \mathbf{x} = 0$ (8a)

b) $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|, \quad \forall \alpha\epsilon F, \mathbf{x}\epsilon V$ (8b)

c) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \forall \mathbf{x},\mathbf{y}\epsilon V$ (the triangle inequality) (8c)

Then $\|\mathbf{x}\|$ is called the *norm* of **x**. Δ

The norm of a vector is a nonnegative real number which is a measure of its "length". Different norms may be defined for a given vector space. For example, the norms $\|.\|_1$, $\|.\|_2$, $\|.\|_p$ and $\|.\|_\infty$ are defined as follows. Let **x** have components $x_1, x_2, \ldots, x_n$ *w.r.t.* some basis. Then

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \tag{9}$$

$$\|\mathbf{x}\|_2 = \left[\sum_{i=1}^{n}|x_i|^2\right]^{1/2} = (\mathbf{x},\mathbf{x})^{1/2} \tag{10}$$

$$\|\mathbf{x}\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p} \tag{11}$$

$$\|\mathbf{x}\|_\infty = \max_i |x_i|. \tag{12}$$

The norm $\|.\|_2$ is also called the *Euclidean norm*. Note that if orthonormal basis is used, the Euclidean norm of a vector in the two-dimensional vector space is indeed the same as its length.

One of the major uses of norm is that it reduces the concept of convergence of vector sequences to that of real number sequences. More precisely, given a norm $\|.\|$ on a vector space (V,F) we say a sequence of vectors $x_1, x_2, \ldots, x_n \epsilon V$ converges to a vector $x \epsilon V$ in that norm if and only if the sequence of nonnegative real numbers $\|x - x_i\|$ converges to zero.

9. Definition. Two norms are said to be ***equivalent*** if any sequence of vectors which converges in one norm, also converges in the other. (See Problem A.10.)

10. Definition. A normed vector space (V,F) is said to be ***complete*** if for every sequence of vectors $x_1, x_2, \cdots$ in V such that $\lim_{\substack{i\to\infty \\ j\to\infty}} \|x_i - x_j\| = 0$, the sum $\lim_{h\to\infty} \sum_{k=1}^{n} x_k$ converges to a vector $x \epsilon V$. Δ

Complete normed vector spaces are called ***Banach spaces***, and complete inner-product spaces are called ***Hilbert spaces***.

11. Example. The space of continuous functions mapping the interval $[-\Delta,0]$ into R^n, where Δ is a finite positive number, is denoted by $C([-\Delta,0],R^n)$. Define a norm for this space as follows:

$$\|f(t)\| = \sup_{-\Delta \leq t \leq 0} |f(t)| \tag{13}$$

This space is a Banach space. Δ

Consider an $n \times n$ matrix A mapping (C^n,C) into itself. Then for any vector $x \epsilon C^n$, Ax will also be a vector in C^n. The norm of matrix A can be defined in terms of the vector norms $\|x\|$ and $\|Ax\|$ as follows.

12. Definition. The norm of a matrix A, denoted by $\|A\|$, is the minimum value of k such that

$$\|Ax\| \leq k\|x\| \quad \forall x \qquad \Delta \tag{14}$$

Using the above definition, the reader can verify that a matrix norm has the following properties:

a) $\|A\| = 0 \iff A = 0$ (15a)

b) $\|Ax\| \leqslant \|A\| \, \|x\|, \quad \forall x$ (15b)

c) $\|A\| = \max_{\|x\|=1} \|Ax\|$ (15c)

d) $\|A+B\| \leqslant \|A\| + \|B\|$ (15d)

e) $\|AB\| \leqslant \|A\| \, \|B\|.$ (15e)

The *Euclidean norm* of an $n \times n$ matrix A is defined as

$$\|A\|_E = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right]^{\frac{1}{2}} \tag{16}$$

This norm is easy to calculate and provides a bound for $\|A\|$. (See Problem A.10.)

A.7 JORDAN CANONICAL FORM OF MATRICES

1. Definition. The scalar λ is called an *eigenvalue* of the square matrix A if there exists a nonzero vector u such that

$$Au = \lambda u. \tag{1}$$

The vector u is then called an *eigenvector* of A corresponding to eigenvalue λ.

Note that if u is an eigenvector of A, then so is αu for any scalar α. The eigenvectors u_i of a matrix are often *normalized*, *i.e.*, the constants of proportionality are chosen such that $\|u_i\|=1$

2. Definition. The *characteristic polynomial* of an $n \times n$ matrix A , denoted by $\Delta(\lambda)$, is $det\,(A-\lambda I)$, *i.e.*

$$\Delta(\lambda) = det\,(A-\lambda I). \quad \Delta \tag{2}$$

Note that the characteristic polynomial of an $n \times n$ matrix is an *nth*-order polynomial in λ. Thus there are n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ which are the roots of this polynomial. Also note that the eigenvalues may be distinct or some of them may be repeated.

3. Example. Consider the square matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \tag{3}$$

The characteristic polynomial of this matrix is

$$\Delta(\lambda) = \det(A-\lambda I) = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} = \lambda^2-2\lambda-3 \tag{4}$$

The eigenvalues of A are the roots of $\Delta(\lambda)$, *i.e.*, $\lambda_1 = -1$, $\lambda_2 = 3$. The eigenvectors of A are found as follows:

$$(A-\lambda_1 I)u_1 = 0 \Rightarrow \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} u_1 = 0 \Rightarrow u_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{5}$$

$$(A-\lambda_2 I)u_2 = 0 \Rightarrow \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} u_2 = 0 \Rightarrow u_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6}$$

which are normalized as

$$u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{7}$$

4. Theorem. If a matrix has distinct eigenvalues, then its eigenvectors will be linearly independent.

Proof. Consider an $n \times n$ matrix A with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Call the corresponding eigenvectors $u_1, u_2, \ldots, u_n$. That is

$$Au_i = \lambda_i u_i \,, \; i = 1,2, \ldots, n. \tag{8}$$

Assume that the eigenvectors are not linearly independent. Then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0 \tag{9}$$

Let

$$\alpha_k \neq 0. \tag{10}$$

Then from (9) we have

$$(\mathbf{A}-\lambda_1\mathbf{I})\,(\mathbf{A}-\lambda_2\mathbf{I})\;\cdots\;(\mathbf{A}-\lambda_{k-1}\mathbf{I})\,(\mathbf{A}-\lambda_{k+1}\mathbf{I})\;\cdots\;(\mathbf{A}-\lambda_n\mathbf{I})\times \sum_{i=1}^{n}\alpha_i\,\mathbf{u}_i=0 \tag{11}$$

Using (8), (11) yields

$$\alpha_k(\lambda_k-\lambda_1)(\lambda_k-\lambda_2)\;\cdots\;(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\;\cdots\;(\lambda_k-\lambda_n)\mathbf{u}_k=0 \tag{12}$$

Since, $\mathbf{u}_k \neq 0$ and $\lambda_k \neq \lambda_i$ for $i \neq k$, (12) yields $\alpha_k = 0$. This contradicts the original assumption (10). Thus the eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are linearly independent. Δ

Note that the class of matrices whose eigenvectors are linearly independent contains the class of matrices with distinct eigenvalues. For example the matrix

$$\mathbf{A}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{13}$$

has repated eigenvalues $\lambda_1 = \lambda_2 = 1$ but its eigenvectors are linearly independent:

$$\mathbf{u}_1=\begin{bmatrix}1\\ 0\end{bmatrix},\quad \mathbf{u}_2=\begin{bmatrix}0\\ 1\end{bmatrix} \tag{14}$$

In fact, if λ is a repeated eigenvalue of a square matrix $\mathbf{A}$, then the number of linearly independent eigenvectors corresponding to λ will be equal to the nullity of $\mathbf{A}-\lambda\mathbf{I}$. Thus for the matrix $\mathbf{A}$ in (14) we have

$$\mathbf{A}-\lambda\mathbf{I}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \tag{15}$$

and $\gamma(\mathbf{A}-\lambda\mathbf{I}) = 2$ which indicates the existence of two linearly independent eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in (15), corresponding to $\lambda=1$. But for the matrix

$$\mathbf{B}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} \tag{16}$$

with eigenvalues $\lambda_1 = \lambda_2 = 1 = \lambda$, we have

$$\mathbf{B}-\lambda\mathbf{I}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \tag{17}$$

whose nullity $\gamma(\mathbf{B}-\lambda\mathbf{I})=1$ indicates the existence of only one (linearly independent) eigenvector

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{18}$$

5. Definition. A nonzero vector $\mathbf{u}$ for which $(\mathbf{A}-\lambda\mathbf{I})^k\mathbf{u} = \mathbf{0}$ but $(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{u} \neq \mathbf{0}$, $k = 1,2,\ldots,$ is called a ***generalized eigenvector*** of order k of $\mathbf{A}$ associated with eigenvalue λ. Δ

From the above definition, the generalized eigenvectors corresponding to eigenvalue λ can be found as follows:

$$k=1 \Rightarrow (\mathbf{A}-\lambda\mathbf{I})\mathbf{u}_1 = \mathbf{0} \tag{19a}$$

$$k=2 \Rightarrow (\mathbf{A}-\lambda\mathbf{I})^2\mathbf{u}_2 = \mathbf{0} = (\mathbf{A}-\lambda\mathbf{I})\mathbf{u}_1 \tag{19b}$$

$$\Rightarrow (\mathbf{A}-\lambda\mathbf{I})\mathbf{u}_2 = \mathbf{u}_1$$

Similarly,

$$(\mathbf{A}-\lambda\mathbf{I})^k\mathbf{u}_k = \mathbf{0} \Rightarrow (\mathbf{A}-\lambda\mathbf{I})\mathbf{u}_k = \mathbf{u}_{k-1} \tag{19c}$$

Note that the generalized eigenvector of order 1 , *i.e.*, $\mathbf{u}_1$, is the same as an eigenvector of $\mathbf{A}$ associated with eigenvector λ.

It can be shown that the generalized eigenvectors of $\mathbf{A}$ associated with each eigenvalue are linearly independent. (See Problem A-15.) Note that the total number of generalized eigenvectors (including eigenvectors) associated with each eigenvalue is equal to the multiplicity of that eigenvalue. It can be proved that the generalized eigenvectors associated with different eigenvalues form a linearly independent set [A.4]. A generalized eigenvector $\mathbf{u}_1$ of matrix $\mathbf{B}$ in (16) is found from $(\mathbf{B}-\lambda\mathbf{I})\mathbf{u}_1 = \mathbf{u}$ where $\mathbf{u}$ is given by (18). Thus

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{20}$$

6. Example. Consider the matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \tag{21}$$

with eigenvalues $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 2$. The eigenvector $\mathbf{u}_1$ corresponding to λ_1 is found from $(\mathbf{C}-\lambda_1\mathbf{I})\mathbf{u}_1 = \mathbf{0}$ which yields $\mathbf{u}_1 = [1,0,0]'$. For the eigenvalues λ_2 and λ_3, the only nonzero vector resulting from $(\mathbf{C}-2\mathbf{I})\mathbf{u} = \mathbf{0}$ is $\mathbf{u}_2 = [0,1,0]'$. Thus a generalized eigenvector $\mathbf{u}_3$ corresponding to $\lambda_2 = \lambda_3 = 2$ exists which is found from $(\mathbf{C}-2\mathbf{I})\mathbf{u}_3 = \mathbf{u}_2$ as $\mathbf{u}_3 = [0,0,1]'$. Note that $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are linearly independent.

7. Theorem. Consider an $n \times n$ matrix $\mathbf{A}$ with eigenvalues λ_2, $i = 1,2,\ldots,p$. Let each λ_i have multiplicity m_i. Then a similarity transformation matrix $\mathbf{M}$ exists which transforms $\mathbf{A}$ into its *Jordan Canonical form* $\mathbf{J}$; that is,

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & & \mathbf{0} \\ & \mathbf{J}_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ \mathbf{0} & & & & \mathbf{J}_p \end{bmatrix} \tag{22}$$

where each $\mathbf{J}_i$ is an $m_i \times m_i$ matrix corresponding to λ_i and has the general form

$$\mathbf{J}_i = \left[\begin{array}{cccc|cccc} \lambda_i & 1 & & & & & & \\ & \cdot & \cdot & & & & & \\ & & \cdot & \cdot & & \mathbf{0} & & \\ & & & \cdot\ 1 & & & & \\ \mathbf{0} & & & \lambda_i & & & & \\ \hline & & & & \lambda_i & & & \mathbf{0} \\ & \mathbf{0} & & & & \cdot & & \\ & & & & & & \cdot & \\ & & & & \mathbf{0} & & & \lambda_i \end{array}\right] \tag{23}$$

Each block of $\mathbf{J}_i$ is called a *Jordan block*. The proof of this theorem may be found in reference [A.5]. It can also be shown [A.4] that the columns of $\mathbf{M}$ are the generalized eigenvectors of A. Δ

Note that $\sum_{i=1}^{p} m_i = n$. Also note that the Jordan form of an $n \times n$ matrix A is always an $n \times n$ matrix which is in an upper triangular form. If A has m linearly independent eigenvectors, then its Jordan form J will have $n-m$ ones above the diagonal. In the special case where all the eigenvectors of A are linearly independent, then $p = n$; $m_i = 1$, $i = 1,2,\ldots,p$ and J reduces to $r = diag\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$. In this case, we say that matrix A has been *diagonalized*. The similarity transformation matrix M in this case becomes the *modal matrix* of A. Note that the columns of the modal matrix are the eigenvectors.

8. Example. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix} \tag{24}$$

It is easy to verify that

$$M = \begin{bmatrix} u_1, u_2, u_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix} \tag{25}$$

transforms A to its Jordan canonical form

$$J = M^{-1}AM = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad \Delta \tag{26}$$

For any $n \times n$ real matrix A, the characteristic polynomial (24) will have real coefficients. Thus any complex eigenvalues of A will occur in conjugate pairs, *i.e.*, if

$$\lambda_i = \sigma_i + j\omega_i \quad , \omega_i \neq 0 \tag{27}$$

is an eigenvalue of A, then

$$\lambda_{i+1} = \bar{\lambda}_i = \sigma_i - j\omega_i \tag{28}$$

will also be an eigenvalue of A. Further, if

$$\mathbf{u}_i = \mathbf{v}_i + j\mathbf{w}_i \tag{29}$$

is an eigenvector of A associated with λ_i, then

$$\mathbf{u}_{i+1} = \bar{\mathbf{u}}_i = \mathbf{v}_i - j\mathbf{w}_i \tag{30}$$

will be an eigenvector of A associated with λ_{i+1}. Thus in such a case, the modal matrix M and the diagonal or the Jordan canonical form of A will have some complex entries. However, it is also possible to find a block diagonal *real* matrix which is similar to A. The following theorem, whose proof may be found in reference [A.4], establishes this fact.

9. Theorem. Consider an $n \times n$ real matrix A with eigenvalues

$$\left.\begin{aligned} \lambda_i &= \sigma_i + j\omega_i \\ \lambda_{i+1} &= \bar{\lambda}_i = \sigma_i - j\omega_i \end{aligned}\right\} i = 1, 3, \ldots, m-1 \tag{31}$$

$$\lambda_l = \bar{\lambda}_l\,, \quad l = m+1, m+2, \ldots, p. \tag{32}$$

Let each λ_i have multiplicity m_i. Then a similarity transformation matrix M exists which transforms A into its *block-Jordan form, i.e.*,

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & & & & & & \\ & \mathbf{J}_3 & & & & & \mathbf{0} & & \\ & & \ddots & & & & & & \\ & & & \mathbf{J}_{n-1} & & & & & \\ & & & & \mathbf{J}_{n+1} & & & & \\ & & & & & \mathbf{J}_{n+2} & & & \\ & & \mathbf{0} & & & & \ddots & & \\ & & & & & & & & \mathbf{J}_p \end{bmatrix} \tag{33}$$

For $i = 1, 3, \ldots, m-1$, each $\mathbf{J}_i$ is a $2m_i \times 2m_i$ real matrix corresponding to λ_i which has the general form where

$$
\mathbf{J}_i = \left[\begin{array}{ccccc|cccc}
\Lambda_i & \mathbf{I} & & & & & & & \\
& \Lambda_i & \mathbf{I} & & & & & & \\
& & \ddots & & & & \mathbf{0} & & \\
& & & \ddots & \mathbf{I} & & & & \\
& & & & \Lambda_i & & & & \\
\hline
& & & & & \Lambda_i & & & \\
& & & & & & \ddots & \mathbf{0} & \\
& & \mathbf{0} & & & & & \ddots & \\
& & & & & & \mathbf{0} & & \Lambda_i
\end{array}\right] \tag{34}
$$

and

$$
\Lambda_i = \begin{bmatrix} \sigma_i & \omega_i \\ -\omega_i & \sigma_i \end{bmatrix} \tag{35}
$$

For $i = m+1, m+2, \ldots, p$ each $\mathbf{J}_i$ is an $m_i \times m_i$ real matrix as in (23). The modal matrix $\mathbf{M}$ is an $n \times n$ real matrix whose columns are the real and imaginary parts of the eigenvectors and the generalized eigenvectors of $\mathbf{A}$:

$$
\mathbf{M} = \left[\mathbf{v}_1, \mathbf{w}_1, \mathbf{v}_3, \mathbf{w}_3, \ldots, \mathbf{v}_{m-1}, \mathbf{w}_{m-1}, \mathbf{u}_{m+1}, \mathbf{u}_{m+2}, \ldots, \mathbf{u}_n\right] \tag{36}
$$

A.8 FUNCTIONS OF A SQUARE MATRIX

Consider a finite polynomial $f(\cdot)$ of a scalar x:

$$
f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0 \tag{1}
$$

Then a corresponding polynomial of a square matrix $\mathbf{A}$ may be defined as

$$
f(\mathbf{A}) = a_m \mathbf{A}^m + a_{m-1} \mathbf{A}^{m-1} + \cdots + a_1 \mathbf{A} + a_0 \mathbf{I} \tag{2}
$$

If function $f(x)$ is analytic, then it can be uniquely expressed in a convergent MacLaurin series

$$f(x) = \sum_{k=0}^{\infty} a_k x^k / k! \tag{3}$$

where

$$a_k = \left. \frac{d^k f(x)}{dx^k} \right|_{x=0} \tag{4}$$

The concept of the function of a matrix can be extended to any analytic function as formalized by the following definition.

1. Definition. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ $(p \leq n)$ and let $f(x)$ be a function which is analytic in an open set containing $\lambda_1, \lambda_2, \ldots, \lambda_p$ with a MacLaurin series (3). Then $f(\mathrm{A})$ is defined as

$$f(\mathrm{A}) = \sum_{k=0}^{\infty} a_k \mathrm{A}^k / k! \tag{5}$$

2. Example. It is known that

$$e^{ax} = 1 + ax + \frac{a^2 x^2}{2!} + \cdots + \frac{a^k x^k}{k!} + \cdots \tag{6}$$

Thus

$$e^{\mathrm{A}t} = \mathrm{I} + \mathrm{A}t + \frac{\mathrm{A}^2 t^2}{k!} + \cdots + \frac{\mathrm{A}^k t_k}{2!} + \cdots \tag{7}$$

3. Theorem. Given an $n \times n$ matrix A, if T is any nonsingular $n \times n$ matrix then

$$\hat{\mathrm{A}} = \mathrm{T}^{-1}\mathrm{A}\mathrm{T} \Rightarrow \mathrm{f}(\hat{\mathrm{A}}) = \mathrm{T}^{-1}\mathrm{f}(\mathrm{A})\mathrm{T} \tag{8}$$

Proof. The proof follows from the following observation:

$$\hat{\mathrm{A}}^k = \hat{\mathrm{A}}\hat{\mathrm{A}} \cdots \hat{\mathrm{A}} = \left(\mathrm{T}^{-1}\mathrm{A}\mathrm{T}\right)\left(\mathrm{T}^{-1}\mathrm{A}\mathrm{T}\right) \cdots \left(\mathrm{T}^{-1}\mathrm{A}\mathrm{T}\right) = \mathrm{T}^{-1}\mathrm{A}^k\mathrm{T} \tag{9}$$

4. Theorem. *(Cayley-Hamilton Theorem).* Let A be a square matrix with characteristic polynomial $\Delta(\lambda) = \det(\mathrm{A} - \lambda\mathrm{I})$. Then $\Delta(\mathrm{A}) = 0$.

Proof. We will prove the theorem for the case where A has distinct eigenvalues. For the general case see Problem A.16. Let the $n \times n$ matrix A have distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and the characteristic polynomial

$$\Delta(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1 \lambda + \alpha_0 \tag{10}$$

We have

$$\mathbf{A} = \mathbf{M}\Lambda\mathbf{M}^{-1} \tag{11}$$

where M is the modal matrix of A and

$$\Lambda = diag\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right] \tag{12}$$

Also, from Theorem 3 we have

$$\mathbf{A}^k = \mathbf{M}\Lambda^k\mathbf{M}^{-1}, \ k = 0, 1, 2, \ldots, n \tag{13}$$

Therefore,

$$\begin{aligned} \Delta(\mathbf{A}) &= \alpha_n \mathbf{A}^n + \alpha_{n-1}\mathbf{A}\mathbf{A}^{n-1} + \ + \ldots \alpha_1 \mathbf{A} + \alpha_0 \mathbf{I} \\ &= \mathbf{M}\left[\alpha_n \Lambda^n + \alpha_{n-1}\Lambda^{n-1} + \ \cdots \ + \alpha_1 \Lambda + \alpha_0 \mathbf{I}\right]\mathbf{M}^{-1} \end{aligned} \tag{14}$$

It is easy to see that

$$\Delta(\mathbf{A}) = \mathbf{M}\begin{vmatrix} \Delta(\lambda_1) & & & 0 \\ & \Delta(\lambda_2) & & \\ & & \ddots & \\ 0 & & & \Delta(\lambda_n) \end{vmatrix}\mathbf{M}^{-1} = \mathbf{M}\mathbf{0}\mathbf{M}^{-1} = \mathbf{0}$$

$$\Delta \tag{15}$$

The Cayley-Hamilton theorem has some important applications which will be discussed here. Using the theorem it can be shown that any with - order polynomial

$$p(\mathbf{A}) = b_m \mathbf{A}^m + b_{m-1}\mathbf{A}^{m-1} + \ \cdots \ + b_1 \mathbf{A} + b_0 \mathbf{I} \tag{16}$$

of an $n \times n$ matrix A $(m \geqslant n)$ can be equivalently expressed as

$$p(A) = a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \cdots + a_1A + a_0I \tag{17}$$

This is true because of $m \geqslant n$ we can write.

$$p(\lambda) = \Delta(\lambda)D(\lambda)+R(\lambda) \tag{18}$$

where $R(\lambda)$ is the remainder polynomial of order $n-1$ resulting from the division of $p(\lambda)$ by the characteristic polynomial $\Delta(\lambda)$ of A. Thus

$$p(A) = \Delta(A)D(A)+R(A) = \Delta(A)0+R(A) = R(A) \tag{19}$$

If the $n\times n$ matrix A has an inverse, it can similarly be shown that

$$A^{-1} = c_{n-1}A^{n-1} + c_{n-2}A^{n-2} + \cdots + c_1A + c_0I \tag{20}$$

This follows from the fact that

$$\Delta(A) = \alpha_nA^n + \alpha_{n-1}A^{n-1} + \cdots + \alpha_1A + \alpha_0I = 0 \tag{21}$$

Thus

$$A^{-1}\Delta(A) = \alpha_nA^{n-1} + \alpha_{n-1}A^{n-2} + \cdots + \alpha_1I + \alpha_0A^{-1} = 0 \tag{22}$$

which implies that

$$A^{-1} = \frac{1}{\alpha_0}\left[\alpha_nA^{n-1} + \alpha_{n-1}A^{n-2} + \cdots + \alpha_1I\right] \tag{23}$$

Note that $\alpha_0 = det(A) \neq 0$ if A^{-1} exists.

The above results can be extended to the case where the polynomial $p(A)$ in (17) is analytic [A.6]. This can be exploited for the computation of $exp(At)$ which is important in the analysis of *l.t.i.* continuous-time systems. That is, we have

$$exp(At) = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k = \sum_{i=0}^{n-1} c_iA^i \tag{24}$$

where c_i, $i = 0, 1, \ldots, n-1$ are linearly independent functions of time [A.7]. Further, for any eigenvalue λ_k of A we have

$$exp(\lambda_kt) = \sum_{i=0}^{n-1} c_i\lambda_k^i,\ k = 1,2,\ldots,n. \tag{25}$$

which can be used to calculate c_i, $i=0,1,\ldots,n-1$.

5. Example. Let

$$\mathbf{A} = \begin{bmatrix} -3 & 1 \\ 0 & -2 \end{bmatrix} \tag{26}$$

Then

$$e^{\mathbf{A}t} = c_0\mathbf{I} + c_1\mathbf{A} \tag{27}$$

Note that scalars c_o and c_1 are functions of time. Also

$$e^{\lambda_i t} = c_0 + c_1\lambda_i \quad , i = 1,2 \tag{28}$$

where $\lambda_1 = -3$ and $\lambda_2 = -2$ are eigenvalues of A. Equations (28) imply that

$$c_0 = -2e^{-3t} + 3e^{-2t} \; , \; c_1 = -e^{-3t} + e^{-2t} \tag{29}$$

Thus

$$\begin{aligned} e^{\mathbf{A}t} &= (-2e^{-3t} + 3e^{-2t})\mathbf{I} + (-e^{-3t} + e^{-2t})\mathbf{A} \\ &= \begin{bmatrix} e^{-3t} & -e^{-3t} + e^{-2t} \\ 0 & e^{-2t} \end{bmatrix} \end{aligned} \tag{30}$$

Note that $e^{\mathbf{A}t}|_{t=0}$ and $\dfrac{d}{dt} e^{\mathbf{A}t}|_{t=0} = \mathbf{A}$.

A.9 QUADRATIC FORMS

1. Definition. Consider the n-dimensional column vector $\mathbf{x} = \begin{bmatrix} x_1, x_2, \ldots, x_n \end{bmatrix}'$ and the $n \times n$ matrix Q. Then the scalar function

$$\mathbf{Q}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = (\mathbf{x}, \mathbf{Q}\mathbf{x}) \tag{1}$$

is called a *quadratic form.* Δ

The matrix Q in quadratic form can be assumed hermitian with no loss of generality. (See Problem A.22.) Note that a quadratic form is always real. (See Problem A-19.)

2. Definition. A hermitian matrix Q (or its associated quadratic form) is said to be *positive definite,* denoted $\mathbf{Q} > 0$, if $(\mathbf{x}, \mathbf{Q}\mathbf{x}) > 0 \;\forall\; \mathbf{x} \neq 0$, *negative definite,* denoted $\mathbf{Q} < 0$, if $(\mathbf{x}, \mathbf{Q}\mathbf{x}) \;\forall\; \mathbf{x} \neq 0$ and *non-negative definite* or *positive semi-definite,* denoted $\mathbf{Q} \geqslant 0$, if

$(\mathbf{x},\mathbf{Qx}) \geqslant 0 \ \forall \mathbf{x}$. Note that if $\mathbf{Q} > 0$, then $\mathbf{Q} < 0$.

3. Theorem. A hermitian matrix $\mathbf{Q}$, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, can always be diagonalized by a unitary (or orthogonal) matrix $\mathbf{P}$, *i.e.*, unitary matrix $\mathbf{P}(\mathbf{P}^{-1}=\mathbf{P}^*)$ always exists such that

$$\mathbf{P}^{-1}\mathbf{QP} = \mathbf{P}^*\mathbf{QP} = \Lambda = diag\left[\lambda_1,\lambda_2, \ldots, \lambda_n\right] \tag{2}$$

Proof. The proof will be given for the case of distinct eigenvalues. For the general proof the reader is referred to Reference [A.2] or Reference [A.5]. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct then the eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of $\mathbf{Q}$ are orthonormal (see Problem A-12) and linearly independent. Thus the matrix $\mathbf{P} = \left[\mathbf{u}_1,\mathbf{u}_2, \ldots, \mathbf{u}_n\right]$ diagonalizes $\mathbf{Q}$. To show that $\mathbf{P}$ is a unitary matrix, note that

$$\mathbf{P}^*\mathbf{P} = \begin{bmatrix} \mathbf{u}^*_1 \\ \mathbf{u}^*_2 \\ \cdot \\ \cdot \\ \mathbf{u}^*_n \end{bmatrix} [\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n] = \begin{bmatrix} \mathbf{u}^*_1\mathbf{u}_1 & \mathbf{u}^*_1\mathbf{u}_2 & \cdots & \mathbf{u}^*_1\mathbf{u}_n \\ \mathbf{u}^*_2\mathbf{u}_1 & \mathbf{u}^*_2\mathbf{u}_2 & \cdots & \mathbf{u}^*_2\mathbf{u}_n \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \mathbf{u}^*_n\mathbf{u}_1 & \mathbf{u}^*_n\mathbf{u}_2 & \cdots & \mathbf{u}^*_n\mathbf{u}_n \end{bmatrix} \tag{3}$$

But $\mathbf{u}_i^*\mathbf{u}_j = (\mathbf{u}_i,\mathbf{u}_j) = \delta_{ij}$ because the eigenvectors are orthonormal. Thus $\mathbf{P}^*\mathbf{P} = \mathbf{I}$ and since $\mathbf{P}^{-1}$ exists due to linear independence of the eigenvectors, the proof is complete.

4. Theorem. A hermitian matrix is positive definite if and only if all its eigenvalues are positive.

Proof: Using Theorem 3, we can write

$$(\mathbf{x},\mathbf{Qx}) = \mathbf{x}^*\mathbf{Qx} = \mathbf{x}^*\mathbf{P}\Lambda\mathbf{P}^*\mathbf{x} = \mathbf{y}^*\Lambda\mathbf{y} = \sum_{i=1}^{n} \lambda_i \left|y_i\right|^2 \tag{4}$$

where $\mathbf{y} = \mathbf{P}^*\mathbf{x}$. Note that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real (see Problem A-12). Thus

$$\mathbf{x}^*\mathbf{Q}x > 0 \ \text{ for all } \mathbf{x} \neq 0 \iff \lambda_i > 0,\ i = 1,2, \ldots, n \quad \Delta \tag{5}$$

It can similarly be proved that

$$Q < 0 \Longleftrightarrow \lambda_i < 0, i = 1,2, \ldots, n \tag{6}$$

and

$$Q \geqslant 0 \Longleftrightarrow \lambda_i \geqslant 0, i = 1,2, \ldots, n \tag{7}$$

5. Definition. The *minors* of an $n \times n$ matrix Q are the determinants of submatrices of Q formed by deleting an equal number of its rows and columns. The *principal minors* of Q are the minors which result from deleting pairs of identically numbered rows and columns of Q. The *mth leading principal minor* of Q, denoted $det(Q_m)$, is the principal minor formed by deleting the last (n-m) pairs of rows and columns of Q.

6. Theorem. *(Sylvester's Theorem).* Let Q Be a hermitian matrix.

(a) Q is positive definite if and only if all its leading principal minors $det(Q_m)$, $m = 1,2, \ldots, n$, are positive.

(b) Q is nonnegative definite if and only if all its principal minors are nonnegative.

Proof. For the proofs of parts (a) and (b), see References [A.1] and [A.7], respectively. Note that $det(Q_m) \geqslant 0$, $m = 1,2, \ldots, n$ does *not* imply $Q \geqslant 0$.

7. Example. Consider the hermitian matrix

$$Q = \begin{bmatrix} q_1 & q_2 \\ \bar{q}_2 & q_3 \end{bmatrix} \tag{8}$$

a) $Q > 0 \Longleftrightarrow q_1 > 0$ and $det(Q) = q_1 q_3 - |q_2|^2 > 0$.

b) $Q \geqslant 0 \Longleftrightarrow q_1 \geqslant 0, q_3 \geqslant 0$ and $det(Q) \geqslant 0$.

Thus for $q_1 = q_2 = 0$, $q_3 < 0$, matrix Q will not be nonnegative definite.

A.10 DIFFERENTIATION WITH RESPECT TO A VECTOR

Letx be an n -dimensional column vector with components $x_1, x_2, \ldots, x_n$. Consider a scalar-valued function of **x**, $f(x_1, x_2, \ldots, x_n) \triangleq f(\mathbf{x})$. The *gradient vector* of $f(\mathbf{x})$, denotedby $\frac{df}{d\mathbf{x}}$, is defined as the following row vector provided that all the partial derivatives exist:

$$\frac{df}{d\mathbf{x}} = \left\{ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right\} \tag{1}$$

Now consider a vector-valued function of $\mathbf{x}$, $\mathbf{f}(\mathbf{x})$, where $\mathbf{f}$ is itself an m-dimensional row vector. Then each component $f_i(\mathbf{x})$, $i=1,2,\ldots,$ m is a scalar-valued function of $\mathbf{x}$ whose gradient is an n-dimensional column vector similar to (3). Thus the gradient of $\mathbf{f}$ is

$$\nabla_{\mathbf{x}}\mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{2}$$

which is an $m \times n$ matrix known as the *Jacobian matrix* of $\mathbf{f}(\mathbf{x})$. The Jacobian matrix plays an important role in the concept of linearization of nonlinear systems.

A.11 DIGITAL COMPUTER IMPLEMENTATION

Many of the calculations related to the concepts discussed in this section are programmed on the digital computer. A package including programs to determine linear independence of vectors, inner product of vectors, rank of a matrix, norm of a vector or a matrix, eigenvalue and eigenvectors of a matrix, characteristics polynomial of a matrix, exponential of a matrix e^{At}) and matrix polynomials as well as many other programs related to linear system analysis has been developed by the authors [A.9] and is available.

PROBLEMS

A.1.a) For an $m\times n$ matrix A and an $n\times k$ matrix **B** show that

1. $\rho(A) = \rho(AA') = \rho(A'A) = \rho(A')$

2. $\rho(AB) \leqslant \rho(A)$, $\rho(AB) \leqslant \rho(A)$

b) For arbitrary square matrix A and nonsingular matrices **B** and C show that

$$\rho(A) = \rho(BA) = \rho(AC) = \rho(BAC)$$

A.2. Let the vectors $e_1, e_2, \ldots, e_n$ in (V,F) form a basis for the vector space. Show that any vector $v\in(V,F)$ has unique components $\alpha_1, \alpha_2, \ldots, \alpha_n$ in F *w.r.t.* this basis.

A.3. Show that

a) The range space of any matrix is a vector space.

b) The null space of any matrix is a vector space.

c) $R(A)$ and $N(A)$ are invariant subspaces under A.

d) $N(A-\lambda_i I)$ is an invariant subspace under A for any eigenvalue λ_i of A.

e) Dimension of $N(A-\lambda_i I)$ is equal to the multiplicity of λ_i, an eigenvalue of A.

f) $\rho(A)+\gamma(A) = n$ for any $m\times n$ matrix A.

A.4. Let A be a linear operator on X into Y , where X and Y both have finite dimension n. Show that $R(A)=Y$ if and only if A^{-1} exists.

A.5. If $V=W_1\oplus W_2$, show that

$$dim(V) = dim(W_1)+dim(W_2)$$

A.6. Consider an inner product $(.,.)$ for a vector space (V,F). Show that

a) $(\mathbf{x},\mathbf{x})$ is real $\forall \mathbf{x} \epsilon V$.

b) $(\mathbf{x},\mathbf{0}) = 0 \quad \forall \mathbf{x} \epsilon V$.

c) $(\mathbf{x},\alpha\mathbf{y}) = \alpha(\mathbf{x},\mathbf{y}) \quad \forall \mathbf{x},\mathbf{y} \epsilon V, \forall \alpha \epsilon F$.

A.7. Show that for any inner product $(\cdot,\cdot)$, the Schwarz inequality

$$|(\mathbf{x},\mathbf{y})| \leqslant (\mathbf{x},\mathbf{x}).(\mathbf{y},\mathbf{y})$$

holds. Furthermore, show that equality holds if and only if $\mathbf{x}=\mathbf{0}$, $\mathbf{y}=\mathbf{0}$ *or* $\mathbf{x}=\alpha\mathbf{y}$ for some scalar α.

A.8. Show for arbitrary matrices **A** and **B**

a) $(\mathbf{A}^*)^* = \mathbf{A}$

b) $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$

c) $(\mathbf{A}+\mathbf{B})^* = \mathbf{A}^*+\mathbf{B}^*$

d) $(\alpha\mathbf{A})^* = \bar{\alpha}\mathbf{A}^*$ for any complete scalar α.

A.9. Verify that the norms $\|.\|_1$, $\|.\|_2$, $\|.\|_p$ and $\|.\|_\infty$ indeed satisfy the postulates of Definition A.2.28. Further, show that in finite-dimensional vector spaces, these norms are equivalent.

A.10. For a matrix $\mathbf{A}=(a_{ij})$ show that

a) $\|\mathbf{A}\|_1 = \max\limits_{j} \left[\sum\limits_{i=1}^{n} |a_{ij}| \right]$

b) $\|\mathbf{A}\|_\infty = \max\limits_{i} \left[\sum\limits_{j=1}^{n} |a_{ij}| \right]$

c) $\|\mathbf{A}\| \leqslant \left[\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} |a_{ij}|^2 \right]^{1/2}$

d) $\|\mathbf{A}\|_2 = \lambda_{\max}$

e) $\lambda_{min}||x||_2 \leqslant ||Ax||_2 \leqslant \lambda_{max}||x||_2$

where λ_{min}^2 and λ_{max}^2 are, respectively, the minimum and the maximum eigenvalue λ of A^*A .

f) $|\lambda| <= ||A||$ for any eigenvalue λ of A.

A.11. Let λ be an eigenvalue of A. Prove that

a) λ is also an eigenvalue of A' and $P^{-1}AP$ for any nonsingular matrix P.

b) λ is an eigenvalue of A and A^*

c) λ^k is an eigenvalue of A^k for any positive or negative integer k.

d) $f(\lambda)$ is an eigenvalue of $f(A)$ where $f(\cdot)$ is any polynomial.

A.12. Show that the eigenvalues of a hermitian matrix are real and the eigenvectors associated with distinct eigenvalues are orthogonal.

A.13. If $\Delta(\lambda) = \lambda^n+\alpha_{n-1}\lambda^{n-1}+ \cdots + \alpha_1\lambda+\alpha_0$ is the characteristic polynomial of A, show that

a) $\alpha_0 = det(A) = \lambda_1\lambda_2 \cdots \lambda_n$

b) $-\alpha_{n-1} = tr(A) = \lambda_1+\lambda_2+ \cdots +\lambda_n$

where the *trace* of the square matrix A, is the sum of the elements on the main diagonal of A.

A.14. Show that a normal matrix A can always be diagonalized by a similarity transformation $Q^{-1}AQ$ where Q is unitary.

A.15. Prove that the generalized eigenvectors u_i, $i=1,2, \ldots ,k$ of matrix A associated with each eigenvalue λ are linearly independent. Hint: Assume $\sum_{i=1}^{k} \alpha_i u_i = 0$. Premultiply by $(A-\lambda I)^{j-1}$ to show $\alpha_j = 0$, $j = k-1,k-2, \ldots ,1$.

A.16. Prove Cayley-Hamilton theorem for the case of matrix **A** with repeated eigenvalues. Hint: Use Jordan canonical form of **A** to show that $\Delta(\mathbf{A}) = \mathbf{M}\Delta(\mathbf{J})\mathbf{M}^{-1}$ and $\Delta(\mathbf{J}) = \mathbf{0}$.

A.17. If

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

show that

$$e^{\mathbf{J}t} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & e^{\lambda t} \end{bmatrix}$$

A.18. Prove that $(e^{\mathbf{A}t})^* = e^{\mathbf{A}^*}$ and $(e^{A^*})^{-1} = e^{-\mathbf{A}t}$.

A.19. Show that if **A** is hermitian, then $(\mathbf{x},\mathbf{A}\mathbf{x})$ is real.

A.20. If **A** is a skew symmetric matrix $(\mathbf{A} = -\mathbf{A}')$, show that $exp(\mathbf{A}t)$ will be unitary, *i.e.*, $exp(\mathbf{A}'t)exp(\mathbf{A}t) = \mathbf{I}$.

A.21. Show that $exp(\mathbf{A}t)exp(\mathbf{B}t) = exp(\mathbf{A}+\mathbf{B})t$ if and only if **A** and **B** commute, *i.e.*, **AB**=**BA**.

A.22. Show that in a quadratic form $\mathbf{x}^*\mathbf{Q}\mathbf{x}$ there is no loss of generality in assuming that **Q** is hermitian, *i.e.*, show that hermitian matrix **P** always exists such that $\mathbf{x}^*\mathbf{Q}\mathbf{x} = \mathbf{x}^*\mathbf{P}\mathbf{x}$.

A.23. Determine α such that the following matrix will be
a) positive definite , b) positive semi-definite.

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 4 & 1 \\ 0 & 1 & \alpha \end{bmatrix}$$

REFERENCES

[A.1] A. E. Taylor and D. C. Lay, *Introduction of Functional Analysis,* Second edition, Wiley and Sons, 1980

[A.2] R. Bellman, *Introduction to Matrix Analysis,* McGraw-Hill, 1960

[A.3] D. M. Wiberg, *State Space and Linear Systems,* Schaum's Series, McGraw-Hill, 1971

[A.4] C. T. Chen, *Introduction to Linear System Theory*, Holt, Rinehart, and Winston, 1970

[A.5] F. R. Gantmacher, *The Theory of Matrices*, Volume I, Chelsea, 1959

[A.6] P. DeRusso, R. Roy and C. Close, *State Variables for Engineers*, Wiley, 1965

[A.7] M. Vidyasagar, "A characterization of e^{-At} and a constructive proof of the controllability criterion," *IEEE Trans. Aut. Control,* Vol. AC-16, pp. 370-371, 1971

[A.8] H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, Dover, 1961

[A.9] M. Jamshidi and M. Malek-Zavarei, *Linear Control Systems - A Computer-Aided Approach*, Pergamon Press, 1986

APPENDIX B

REVIEW OF LAPLACE, z AND MODIFIED z TRANSFORMS

B.1 INTRODUCTION

The purpose of this appendix is to review the Laplace, z and Modified z transformations. These transformations are fundamental tools in the study of linear time-invariant systems. They convert differential and/or difference equations describing *l.t.i.* systems into algebraic equations.

It is assumed that the reader already has a basic understanding of these transformations. We will review only those aspects of these subjects which have immediate relevance to the study of *l.t.i.* systems. For a more detailed treatment of these vast topics the reader is referred to the references listed at the end of this appendix.

B.2 REVIEW OF THE LAPLACE TRANSFORM

The Laplace transform is a fundamental tool for studying *l.t.i.* continuous-time systems. It converts a differential equation describing a *l.t.i.* system into an algebraic equation. Thus it considerably reduces the computational effort required for the analysis and design of systems described by such equations.

In this section the Laplace transformation concept will be presented and some properties of the Laplace transform with relevance to the material of the book will be reviewed. For a more detailed and vigorous treatment of this topic the reader is referred to references [B.1] to [B.4]. We will deal only with *unilateral* or *one-sided* Laplace transform. We assume that all time functions of interest vanish for $t < t_0$. Further, we assume that the initial time $t_0 = 0$ with no loss of generality, since in the following treatment t can always be replaced by $t - t_0$.

1. Definition. Let $f(t)$ be an integrable function of time defined on the internal $[0,\infty)$. Then the (unilateral) Laplace transform of $f(t)$, denoted by $L[f(t)] \underline{\Delta} F(s)$ is defined by the integral

$$F(s) = \int_{0-}^{\infty} f(t)e^{-st}dt \tag{1}$$

for all values of the complex frequency s for which the integral exists. Δ

It is clear from (1) that if $F(s)$ exists for some $s=\sigma_t+j\omega_c$, then it will also exist for all s such that $Re(s)>\sigma_c$. The greatest lower bound on σ_c for which integral (1) exists is called the ***abscissa of convergence*** for the function $f(t)$. The region in the complex plane to the right of the abscissa of convergence is called the ***region*** or the ***domain*** of convergence of $F(s)$. (See Figure 1.)

The Laplace integral (1) can be extended to other values of $s = \sigma + j\omega$. That is, $F(s)$ can be considered a well-defined function for all values of s except for those which cause integral (1) to diverge.

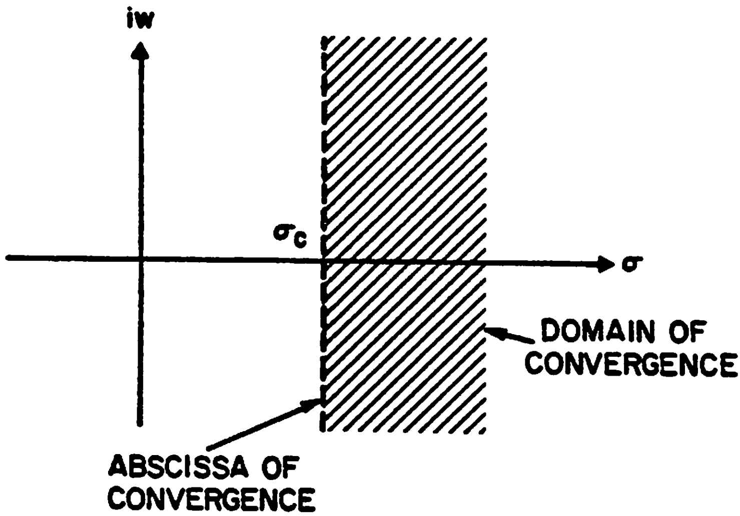

Figure 1. Domain of Convergence of the Laplace Transform in the s-plane

2. Example. Consider the function $f(t)=e^{\alpha t}$ where α is a real or a complex number. Then

$$L[f(t)] \underline{\Delta} F(s) = \int_{0-}^{\infty} e^{\alpha t} e^{-st} dt = \frac{e^{-(s-\alpha)t}}{-(s-\alpha)} \Big|_{0-}^{\infty} \tag{2}$$

Thus $F(s)$ exists, *i.e.* it is finite, when $Re(s-\alpha)>0$, and it is infinite otherwise. That is,

$$F(s) = L[e^{\alpha t}] = \frac{1}{s-\alpha} \text{ for } Re(s) > Re(\alpha) \tag{3}$$

and $Re(\alpha)$ is the abscissa of convergence for $e^{\alpha t}$. However, we will regard the Laplace transform of $e^{\alpha t}$ to be $\frac{1}{s-\alpha}$, which is well-defined for all s except $s=\alpha$.

B.2.1 Properties of the Laplace Transform

Here we will discuss the important properties of the Laplace transform. Further properties will be introduced in the Problems at the end of this chapter.

3. Uniqueness. If two time functions $f_1(t)$ and $f_2(t)$ have the same function of the complex frequency s, say $F(s)$, as their Laplace transform, then $f_1(t)$ and $f_2(t)$ can differ only trivially. Δ

Uniqueness is a fundamental property of the Laplace transform. Given a time function $f(t)$, it allows us to use a *unique* function $F(s)$ as its Laplace transform. Conversely, given a Laplace transform $F(s)$, there is a *unique* time function $f(t)$ over the interval $[0,\infty)$ (except for trivialities) such that $F(s) = L[f(t)]$. This is written as

$$f(t) = L^{-1}[F(s)] \tag{4}$$

meaning that $f(t)$ is the *inverse Laplace transform* of $F(s)$.

4. Example. Consider the following time functions:

$$f_1(t) = \begin{cases} 1 & \text{for } t>0 \\ 0.5 & \text{for } t=0 \\ 0 & \text{for } t<0 \end{cases}, \quad f_2(t) = \begin{cases} 1 \text{ for } t>0 \\ 1 \text{ for } t=0 \\ 0 \text{ for } t<0 \end{cases} \tag{5}$$

For our purposes, the difference between these functions is trivial. Both functions are referred to as the *unit step function* and have the same Laplace transform.

$$F_1(s) = F_2(s) \stackrel{\Delta}{=} \int_{0-}^{\infty} e^{-st} dt = \frac{1}{s} \tag{6}$$

5. Linearity. If $F_1(s)$ and $F_2(s)$ are the Laplace transforms of the time functions $f_1(t)$ and $f_2(t)$, respectively, then

$$L\left[\alpha_1 f_1(t)+\alpha_2 f_2(t)\right] = \alpha_1 F_1(s) + \alpha_2 F_2(s) \tag{7}$$

for arbitrary constants α_1 and α_2. Δ

The linearity property of the Laplace transform follows immediately from the linearity of the Laplace integral (1).

6. Example. Consider

$$f(t) = \sin \alpha t = \frac{1}{2j}\left(e^{j\alpha t} - e^{-j\alpha t}\right) \tag{8}$$

Then, using (3) and the linearity property of the Laplace transform yields

$$L[\sin \alpha t] = \frac{1}{2j}\left[\frac{1}{s-j\alpha} - \frac{1}{s+j\alpha}\right] = \frac{\alpha}{s^2+\alpha^2} \tag{9}$$

7. Differentiation Rule. Let $F(s)$ be the Laplace transform of $f(t)$. Then,

$$L\left[\frac{d^n}{dt^n} f(t)\right] = s^n F(s) - s^{n-1} f(0-) - s^{n-2} f^{(1)}(0-)$$

$$- \cdots - s f^{(n-2)}(0-) - f^{(n-1)}(0-) \tag{10}$$

where $f^{(i)}$ indicates the ith derivative of f w.r.t. t. Δ

The differentiation rule can be verified by performing the Laplace integral (1) for $\frac{d}{dt} f(t)$ by parts:

$$L\left[\frac{d}{dt} f(t)\right] = f(t)e^{-st}\Big|_{0-}^{\infty} - \int_{0-}^{\infty} f(t)(-se^{-st})dt$$

$$= -f(0-) + s\int_{0-}^{\infty} f(t)e^{-st}dt = sF(s) - f(0-) \tag{11}$$

Application of (11) $n-1$ times implies (10).

8. Example. Consider $\frac{d}{dt} \sin \alpha t = \alpha \cos \alpha t$. Thus, (10) for $n=1$ and (9) yield

$$L\left[\frac{d}{dt}\sin\alpha t\right] = L\left[\alpha\cos\alpha t\right] = s\left[\frac{\alpha}{s^2+\alpha^2}\right]$$

Using the linearity of the Laplace transform results

$$L[\cos\alpha t] = \frac{s}{s^2+\alpha^2} \tag{12}$$

9. Integration Rule. If $L[f(t)] = F(s)$, then

$$L\left[\int_{0-}^{t}\int_{0-}^{\tau_1}\cdots\int_{0-}^{\tau_{n-1}} f(\tau_n)d\tau_n d\tau_{n-1}\cdots d\tau_1\right] = \frac{1}{s^n}F(s) \quad \Delta \tag{13}$$

Equation (13) can be verified again by using integration by parts:

$$L\left[\int_{0-}^{t} f(\tau)d\tau\right] \triangleq \int_{0-}^{\infty}\left[\int_{0-}^{t} f(\tau)d\tau\right]e^{-st}\,dt$$

$$= \left[\int_{0-}^{t} f(\tau)d\tau\right]\frac{e^{-st}}{-s}\Big|_{0-}^{\infty} - \int_{0-}^{\infty} f(t)\left(\frac{e^{-st}}{-s}\right)dt$$

$$= 0 + \frac{1}{s}\int_{0-}^{\infty} f(t)e^{-st}dt = \frac{1}{s}F(s) \tag{14}$$

Repeated application of (14) implies (13).

10. Time Function Translation. If $L[f(t)] = F(s)$, then for any positive scalar α,

$$L[f(t-\alpha)] = e^{-\alpha s}F(s) \tag{15}$$

Further, the scalar α may be negative if $f(t-\alpha)$ vanishes for $t<0$. Δ

This property follows easily from the Laplace transform defining integral:

$$L[f(t-\alpha)] \triangleq \int_{0-}^{\infty} f(t-\alpha)e^{-st}dt = \int_{\alpha-}^{\infty} f(t-\alpha)e^{-st}dt$$

$$\int_{0-}^{\infty} f(\tau)e^{-s(\tau+\alpha)}d\tau = e^{-s\alpha}\int_{0-}^{\infty} f(\tau)e^{-s\tau}d\tau = e^{-s\alpha}F(s) \tag{16}$$

where $\tau = t-\alpha$. Note that $f(t-\alpha)$ is the original function delayed by α as illustrated in Figure 2.

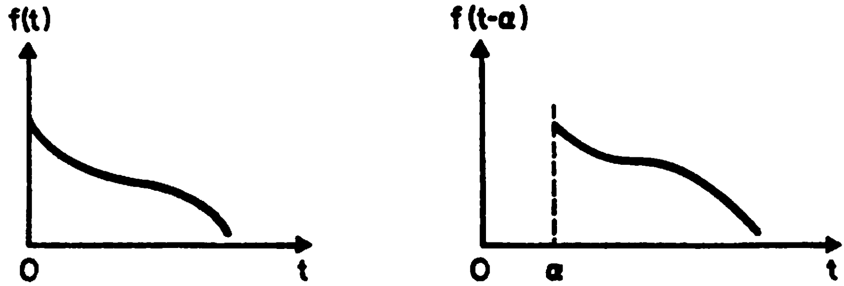

Figure 2. Function $f(t)$ Delayed by $a>o$

11. Example. Consider a pulse function defined by

$$p(t) = \begin{cases} \frac{1}{a} & 0 < t \leqslant a \\ 0 & elsewhere \end{cases} \tag{17}$$

as shown in Figure 3. Note that

$$p(t) = \frac{1}{a}[u(t) - u(t-a)] \tag{18}$$

where $u(t)$ is the unit step function. Using (6) and the linearity and time function translation properties of the Laplace transform we obtain

$$\begin{aligned} L[p(t)] &= \frac{1}{a}\{L[u(t)] - L[u(t-a)]\} \\ &= \frac{1}{a}[\frac{1}{s} - e^{-sa}\frac{1}{s}] = \frac{1-e^{-sa}}{as} \end{aligned} \tag{19}$$

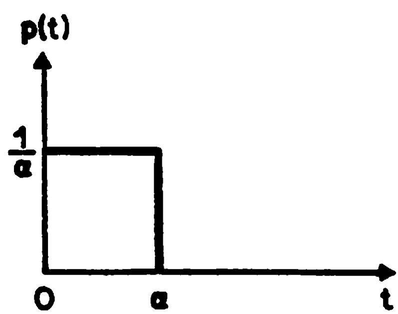

Figure 3. A Pulse Function

A *unit impulse*, or a *Dirac delta function*, at $t=0$, denoted by $\delta(t)$, is defined as

$$\delta(t) = \lim_{a\to 0} p(t) \tag{20}$$

Thus,

$$L[\delta(t)] = \lim_{a\to 0} \frac{1-e^{-sa}}{as} = 1 \tag{21}$$

Also, by the definition of the derivative, (20) implies that $\delta(t) = du(t)/dt$ and by the differentiation rule of the Laplace transform

$$L[\delta(t)] = s\ L[u(t)] - u(0-) = s(\frac{1}{s})-0 = 1 \tag{22}$$

which verifies (21).

12. Laplace Transform Translation. If $L[f(t)]=F(s)$, then

$$L\left[e^{\alpha t}f(t)\right] = F(s-\alpha) \tag{23}$$

where α is any complex scalar.

13. Initial Value Property. If $f(t)$ approaches a limit as t approaches zero from the right, then

$$f(0+) = \lim_{s\to\infty} sL[f(t)] \tag{24}$$

Note that the limit on the right side of (24) may exist without the existence of $f(0+)$.

14. Final Value Property. If $f(t)$ approaches a limit as t approaches ∞, then

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sL[f(t)] \qquad \Delta \tag{25}$$

Again, note that the limit on the right side of (25) may exist without the existence of the limit on the left side.

15. Example. Consider $f(t) = e^{-\alpha t}\cos\omega t$ where α is a real positive number. Using (12) and (22) we have

$$L[f(t)] \triangleq F(s) = \frac{s+\alpha}{(s+\alpha)^2+\omega^2} \tag{26}$$

Now $\lim_{s\to\infty} sF(s) = 1$ which is equal to $f(0+)$. Also $\lim_{s\to 0} sF(s) = 0$ which is equal to $\lim_{t\to\infty} f(t)$.

16. Definition. Let time functions $f_1(t)$ and $f_2(t)$ be defined on the interval $(-\infty,\infty)$. Then the *convolution* of f_1 and f_2, denoted $f_1{*}f_2$, is a time function defined by

$$f_1{*}f_2(t) = \int_0^t f_1(\tau)f_2(t-\tau)d\tau \tag{27}$$

provided that the integral exists for all t.

Note that if $f_1(s)$ and $f_2(s)$ vanish for $s<0$, then their convolution will be

$$f_1{*}f_2(t) = \int_0^t f_1(\tau)f_2(t-\tau)d\tau \tag{28}$$

See Problem B.4.

17. Convolution Property. If $F_1(s)$ and $F_2(s)$ are the Laplace transforms of $f_1(t)$ and $f_2(t)$, respectively, then

$$L\left[f_1{*}f_2(t)\right] = F_1(s)F_2(s) \qquad \Delta \tag{29}$$

See Problem B.5. This is an important property of the Laplace transform often used in the study of *l.t.i.* continuous-time systems.

For some further properties of the Laplace transform see Problem B.6.

B.2.2 Inversion of the Laplace Transform

The importance of the Laplace transform lies in the fact that there is a unique time function $f(t)$ for each Laplace transform $F(s)$. The following theorem provides a procedure for determining the inverse Laplace transform of $F(s)$.

18. Theorem. Let $F(s) = L[f(t)]$ and let σ_c be the abscissa of convergence for the function $f(t)$. Then

$$f(t) \triangleq L^{-1}[F(s)] = \frac{1}{2\pi j} \lim_{\omega\to\infty} \int_{\sigma-j\omega}^{\sigma+j\omega} F(s)e^{st}ds \tag{30}$$

for all $\sigma>\sigma_c$.

The proof of this fact may be found in reference [B.2] or [B.5]. Note that by assumption $f(t) = 0$ for $t<0$; thus the integral, (30) yields $f(t)$ for $t>0$. Also, $f(.)$ could be discontinuous at t; in fact the left-hand side of (30) may be replaced by $\frac{1}{2}[f(t+0) + f(t-0)]$.

The integral in (30) is, in general, difficult to compute. Some authors have developed numerical methods of evaluating this integral [B.6 - B.10]. These methods provide the inverse Laplace transform function in a tabular form rather than analytical form. An alternative method of determining the inverse Laplace transform is by expanding it into easily invertible components and applying the linearity property of the Laplace transform. This procedure is formalized below.

19. Definition. A function $F(s)$ is said to be *rational* if the function is expressed as the ratio of two polynomials. Further, if the degree of the numerator polynomial is m and that of its denominator is n, it is called *proper* if $m \leqslant n$, *strictly proper* if $m<n$ and *improper* if $m>n$.

20. Theorem. Let $F(s)=L[f(t)]$ be a strictly proper rational function with real coefficients, *i.e.*,

$$F(s) = \frac{Q(s)}{P(s)} = \frac{q_m s^m + q_{m-1}s^{m-1} + \cdots + q_1 s + q_0}{p_n s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0}$$

$$= \frac{q_m}{p_n} \frac{\prod_{i=1}^{m}(s-z_j)}{\prod_{i=1}^{n}(s-p_i)} \tag{31}$$

where the coefficients q_j, $j = 1,2,\ldots,m$ and p_i, $i = 1,2,\ldots,n$ are real. Then complex numbers K_{ij} exist such that $F(s)$ can be written as

$$F(s) = \sum_{i=1}^{p}\sum_{j=1}^{m_i} \frac{K_{ij}}{(s-p_i)^j} \tag{32}$$

where m_i is the multiplicity of root p_i of $P(s)$ and p is the number of distinct roots of $P(s)$. The complex numbers K_{ij} are determined as follows:

$$K_{i,m_i-j+1} = \frac{1}{(j-1)!} \frac{d^{j-1}}{ds^{j-1}} (s-p_i)^{m_i} F(s) \mid_{s=p_i},$$
$$j = 1,2, \ldots, m_i \quad ; \; i = 1,2, \ldots, p \tag{33}$$

Then the inverse Laplace transform of $F(s)$ is

$$f(t) = \sum_{i=1}^{p} \sum_{j=1}^{m_i} \frac{K_{ij} t^{j-1}}{(j-1)!} e^{\lambda_i t} \qquad \Delta \tag{34}$$

The proof of this theorem is straightforward but tedious [B.1].

The roots z_j, $j = 1,2, \ldots, m$ of $Q(s)$ are called the *zeros* of $F(s)$ and the roots p_i, $i = 1,2, \ldots, n$ of $P(s)$ are called the *poles* of $F(s)$. Clearly, the abscissa of convergence for $f(t)$ is $\max_i \{Re(p_i)\}$. The complex numbers K_{ij}, $j = 1,2, \ldots, m_i$ are called the *residues* corresponding to the role p_i. The poles, the zeros and the residues either are real or occur in complex conjugate pairs because the coefficients p_i and q_j are assumed to be real. (See Problem B.8.) Note that $\sum_{i=1}^{p} m_i = n$. Also note that in the case where all the poles of $F(s)$ are distinct, *i.e.* where $m_i = 1$, $i = 1,2, \ldots, p$, then $p = n$ and (32) becomes

$$F(s) = \sum_{i=1}^{n} \frac{K_i}{s-p_i} \tag{35}$$

where

$$K_i = (s-p_i) F(s)|_{s=p_i}, \quad i = 1,2, \ldots, n \tag{36}$$

and (34) becomes

$$f(t) = \sum_{i=1}^{n} K_i e^{\lambda_i t} \tag{37}$$

21. Example. Consider

$$F(s) = \frac{s}{(s+1)^3 (s+2)} \tag{38}$$

It can be written as

$$F(s) = \frac{K_{11}}{s+1} + \frac{K_{12}}{(s+1)^2} + \frac{K_{13}}{(s+1)^3} + \frac{K_{21}}{s+2} \tag{39}$$

where

$$K_{13} = (s+1)^3 F(s)\Big|_{s=-1} = -1 \tag{40}$$

$$K_{12} = \frac{d}{ds}(s+1)^3 F(s)\Big|_{s=-1} = \frac{2}{(s+2)^2}\Big|_{s=-1} = 2 \tag{41}$$

$$K_{11} = \frac{1}{2}\frac{d^2}{ds^2}(s+1)^3 F(s)\Big|_{s=-1} = \frac{-2}{(s+1)^3}\Big|_{s=-1} = -2 \tag{42}$$

$$K_{21} = (s+2)F(s)\Big|_{s=-2} = 2 \tag{43}$$

Thus

$$F(s) = \frac{-2}{s+1} + \frac{2}{(s+1)^2} - \frac{1}{(s+1)^3} + \frac{2}{s+2} \tag{44}$$

and

$$L^{-1}[F(s)] = -2e^{-t} + 2te^{-t} - \frac{t^2}{2}e^{-t} + 2e^{-2t}, \quad t \geqslant 0 \qquad \Delta \tag{45}$$

If $p_i = \sigma_i + j\omega_i$ and $p_{i+1} = \bar{p}_i = \sigma_i - j\omega_i$ are a pair of complex conjugate poles of $F(s)$, then the corresponding pair of residues K_i and K_{i+1} will also be complex conjugate. This arrangement can be combined into a more compact form as follows:

$$\frac{K_i}{s-p_i} + \frac{K_{i+1}}{s-p_{i+1}} = \frac{K_i}{s-\sigma_i-j\omega_i} + \frac{\bar{K}_i}{s-\sigma_i+j\omega_i} = \frac{A_i(s-\sigma_i)+jB_i\omega_i}{(s-\sigma_i)^2+\omega_i^2} \tag{46}$$

where $A_i = 2\,Re(K_i)$ and $B_i = 2Im(K_i)$. Then A_i and B_i may be found from

$$B_i + jA_i = 2jK_i = \frac{1}{\omega_i}\left[(s-\sigma_i)^2+\omega_i^2\right]F(s)\Big|_{s=\sigma_i+j\omega_i} \tag{47}$$

The inverse Laplace transform of (46) is

$$e^{\sigma_i t}(A_i \cos\omega_i t + B_i \sin\omega_i t), \qquad t \geqslant 0 \tag{48}$$

or

$$2|K_i|e^{\sigma_i t}\cos(\omega_i t + \angle K_i), \qquad t \geqslant 0 \tag{49}$$

22. Example. Consider $F(s) = \dfrac{2s^2+25s+50}{s^3+10s^2+50s}$. It can be written as

$$\frac{2s^2+25s+50}{s[(s+5)^2+25]} = \frac{K_1}{s} + \frac{A_2(s+5)+jB_2(5)}{(s+5)^2+5^2} \tag{50}$$

where

$$K_1 = sF(s)|_{s=0} = 1 \tag{51}$$

and

$$B_2+jA_2 = \frac{1}{5}\left[(s+5)^2+5^2\right]F(s)|_{s=-5+j5} = \frac{2s^2+25s+50}{5s}|_{s=-5+j5} \tag{52}$$

which yields $A_2=1$ and $B_2=2$. Thus

$$L^{-1}[F(s)] = 1+e^{-5t}(\cos 5t+2\sin 5t),\ t \geqslant 0 \quad \Delta \tag{53}$$

If $F(s)$ is not a strictly proper rational function, it can always be written as a polynomial in s plus a strictly proper rational function. Then the linearity property of the Laplace transform can be applied to find the inverse Laplace transform. Note that

$$L^{-1}\left[a_k s^k+a_{k-1}s^{k-1}+\cdots+a_1 s+a_0\right]$$

$$= a_k\delta^{(k)}(t)+a_{k-1}\delta^{(k-1)}(t)+\cdots+a_1\delta^{(1)}(t)+a_0\delta(t) \tag{54}$$

where $\delta(t)$ is the unit impulse and

$$\delta^{(i)}(t) \overset{\Delta}{=} \frac{d^i}{dt^i}\delta(t). \qquad \Delta$$

The Laplace transforms of some common functions of time as well as the inverse Laplace transforms of some common functions of s are collected in Table 1. Note that the time function $f(t)$ is assumed to vanish for $t<0$.

Table 1. Laplace Transform Pairs

$f(t)$	$F(s)$
$\delta(t)$	1
$u(t)$ (unit step function)	$\frac{1}{s}$
$u(t-\alpha)$	$\frac{e^{-\alpha s}}{s}$
$r(t) = tu(t)$ (unit ramp)	$\frac{1}{s^2}$
t^n	$\frac{n!}{s^{n+1}}$
$t^n e^{\alpha t}$	$\frac{n!}{(s-\alpha)^{n+1}}$
$-1-\alpha t+e^{\alpha t}$	$\frac{\alpha^2}{s^2(s-\alpha)}$
$1-e^{\alpha t}+\alpha t e^{\alpha t}$	$\frac{\alpha^2}{s(s-\alpha)^2}$
$\frac{-\beta}{\alpha^2} + \frac{\beta/\alpha+\alpha t-\beta t}{\alpha} e^{\alpha t}$	$\frac{s-\beta}{s(s-\alpha)^2}$
$\sin\omega t$	$\frac{\omega}{s^2+\omega^2}$
$\cos\omega t$	$\frac{s}{s^2+\omega^2}$
$e^{\alpha t}\sin\omega t$	$\frac{\omega}{(s-\alpha)^2+\omega^2}$
$e^{\alpha t}\cos\omega t$	$\frac{s-\alpha}{(s-\alpha)^2+\omega^2}$
$\frac{1}{\omega^2}(1-\cos\omega t)$	$\frac{1}{s(s^2+\omega^2)}$

$f(t)$	$F(s)$
$e^{-\zeta\omega_n t}\sin(\omega_n\sqrt{1-\zeta^2}t)$	$\dfrac{\omega_n\sqrt{1-\zeta^2}}{s^2+2\zeta\omega_n s+\omega_n^2}$
$\dfrac{e^{\alpha t}}{\alpha^2+\omega^2}+\dfrac{\sin(\omega t-\rho)}{\omega\sqrt{\alpha^2+\omega^2}}$ where $\rho=\tan^{-1}\dfrac{\omega}{\alpha}$	$\dfrac{1}{(s-\alpha)(s^2+\omega^2)}$
$e^{\alpha t}(A\cos\omega t+B\sin\omega t)$	$\dfrac{A(s-\alpha)+B\omega}{(s-\alpha)^2+\omega^2}$
$e^{\alpha t}\left[A\cos\omega t+\dfrac{B+\alpha A}{\omega}\sin\omega t\right]$	$\dfrac{As+B}{(s-\alpha)^2+\omega^2}$
$\dfrac{\sqrt{(\alpha+\beta)^2+\omega^2}}{\omega}e^{\alpha t}\sin(\omega t+\rho)$ where $\rho=\tan^{-1}\dfrac{\omega}{\alpha+\beta}$	$\dfrac{s+\beta}{(s-\alpha)^2+\omega^2}$
$2\|K\|e^{\alpha t}\cos(\omega t+\angle K)$	$\dfrac{K}{s-\alpha-j\omega}+\dfrac{\bar{K}}{s-\alpha+j\omega}$
$\dfrac{1}{\alpha^2+\omega^2}+\dfrac{e^{\alpha t}}{\omega\sqrt{\alpha^2+\omega^2}}\sin(\omega t-\rho)$ where $\rho=\cos^{-1}\dfrac{\omega}{2}$	$\dfrac{1}{s\left[(s-\alpha)^2+\omega^2\right]}$
$\dfrac{1}{\omega_n^2}-\dfrac{e^{-\zeta\omega_n t}}{\omega_n\sqrt{1-\zeta^2}}\sin\left[\omega_n\sqrt{1-\zeta^2}t+\rho\right]$ where $\rho=\cos^{-1}\dfrac{\sqrt{1-\zeta^2}}{\zeta}$	$\dfrac{1}{s\left[s^2+2\zeta\omega_n s+\omega_n^2\right]}$
$\dfrac{\beta}{\alpha^2+\omega^2}-\dfrac{e^{\alpha t}\sqrt{(\alpha+\beta)^2+\omega^2}}{\omega\sqrt{\alpha^2+\omega^2}}\sin(\omega t+\rho)$ where $\rho=\tan^{-1}\dfrac{\omega}{\alpha+\beta}-\tan^{-1}\dfrac{\omega}{\alpha}$	$\dfrac{s+\beta}{s\left[(s-\alpha))^2+\omega^2\right]}$

B.2.3 Solution of Differential Equations

An important application of the Laplace transform is to solve linear constant-coefficient differential equations. By taking the Laplace transform of both sides of such a differential equation, the Laplace transform of the unknown function can be expressed as a rational function of s. The inverse Laplace transform of this function then provides the solution. Thus the Laplace transform converts the solution of a linear constant-coefficient differential equation to that of an algebraic equation. The following example further illustrates the method.

23. Example. Consider the differential equation

$$\frac{d^3f}{dt^3} + 10\frac{d^2f}{dt^2} + 37\frac{df}{dt} + 52f(t) = -3u(t) + \frac{d\delta(t)}{dt},\ t \geqslant 0 \tag{55a}$$

with initial conditions

$$f(0-) = \dot{f}(0-) = 1,\ \ddot{f}(0-) = 0 \tag{55b}$$

Taking the Laplace transform of both sides of Eq. (55a) yields

$$\left[s^3F(s)-s^2f(0-)-s\dot{f}(0-)-\ddot{f}(0-)\right] + 10\left[s^2F(s)-sf(0-)-\dot{f}(0-)\right]$$
$$+ 37[sF(s)-f(0-)] + 52\,F(s) = \frac{3}{s} + s \tag{56}$$

or

$$F(s) = \frac{s^3+12s^2+47s-3}{s(s^3+10s^2+37s+52)} \tag{57}$$

which can be written in partial fraction form as

$$F(s) = \frac{-3/52}{s} + \frac{63/20}{s+4} - \frac{\frac{136}{65}(s+3)-\frac{541}{130}}{(s+3)^2+4} \tag{58}$$

Thus, the solution is

$$f(t) = L^{-1}[\mathrm{F}(s)] = -\frac{3}{52} + \frac{63}{20}\,\mathrm{e}^{-4t} -$$
$$\frac{136}{65}e^{-3t}\cos 2t - \frac{541}{130}\,\mathrm{e}^{-3t}\sin 2t,\ t \geqslant 0 \qquad \Delta \tag{59}$$

Laplace transform methods can also be used to solve differential-difference equations with constant coefficients [B.5]. This, however, does not lead to a closed form solution of such equations. To illustrate this, consider the following example.

24. Example. Consider the first-order differential-difference equation

$$\dot{f}(t)-f(t-1)=0\,,\quad t\geqslant 0 \tag{60}$$

Assuming $f(t)$ is Laplace transformable, taking the Laplace transform of (60) yields

$$sF(s)-f(0-)-e^{-s}F(s)=0 \tag{61}$$

or, assuming $s-e^{-s}\neq 0$,

$$F(s)\stackrel{\Delta}{=}\int_{0-}^{\infty} f(t)e^{-st}dt=\frac{f(0-)}{s-e^{-s}} \tag{62}$$

Equation (62) can also be written as follows:

$$\int_{1}^{\infty} f(t)e^{-st}dt=\frac{f(0-)}{s-e^{-s}}-\int_{0-}^{1} f(t)e^{-st}dt \tag{63}$$

which expresses the transform of $f(t)$ for $t>1$ in terms of values of $f(t)$ over $[0,1]$. From (30), for $t>1$ we obtain

$$f(t)=\frac{1}{2\pi j}\lim_{\omega\to\infty}\int_{\sigma-j\omega}^{\sigma+j\omega}\left[\frac{f(0-)}{s-e^{-s}}-\int_{0-}^{1} f(t)e^{-st}dt\right]e^{st}ds \tag{64}$$

where σ is larger than the abscissa of convergence of $f(t)$. Equation (64) expresses $f(t)$ for $t>1$ in terms of values of $f(t)$ over $[0,1]$. For example, if the boundary condition for (60) is

$$f(t)=2,\quad 0\leqslant t\leqslant 1, \tag{65}$$

(64) becomes

$$f(t)=\frac{1}{2\pi j}\lim_{\omega\to\infty}\int_{\sigma-j\omega}^{\sigma+j\omega}\frac{2se^{-s}-2e^{-2s}+2e^{-s}}{s(s-e^{-s})}\,e^{st}ds \tag{66}$$

Equation (66) cannot be solved to obtain a closed-form solution for $f(t)$. However, a series expansion of $f(t)$ for $t>1$ may be obtained from (66). Δ

The detailed method of solution of differential-difference equations via the Laplace transform can be found in reference [B.5]. We will not, however, use this method to solve differential-differences equations. Rather, we will use the modified z transform method which, as we will see in Section B.4, will result in closed-form solutions.

B.3 REVIEW OF THE z TRANSFORM

In the previous section we illustrated how the Laplace transform can be used to convert the solution of a linear constant-coefficient differential equation into that of an algebraic equation. In a similar manner, the z transform can be used to convert the solution of a linear difference equation with constant coefficients to that of an algebraic equation. As such, the z transform is a fundamental tool for studying *l.t.i.* discrete-time systems.

In this section we will review the basic z transform concept as well as some of its important properties. For a more detailed and rigorous treatment of the z transform the reader is referred to references [B.11] to [B.17].

The concept of the z transform is closely related to the idea of sampling a time function explained below.

1. Ideal Sampling. Consider a function of continuous time $f(t)$ where $f(t) = 0$ for $t < 0$. An *ideal sampler* is one that takes samples of infinitesimal width of $f(t)$ at regular intervals of time (Figure 1). Thus, an ideal sampling $f^*(t)$ of $f(t)$ can be viewed as a sequence of equally spaced impulses with magnitudes equal to the values of the function $f(.)$ at the corresponding discrete times. That is,

$$f^*(t) = \sum_{k=0}^{\infty} f(kT)\delta(t-kT) \tag{1}$$

where T is called the *sampling period*. Δ

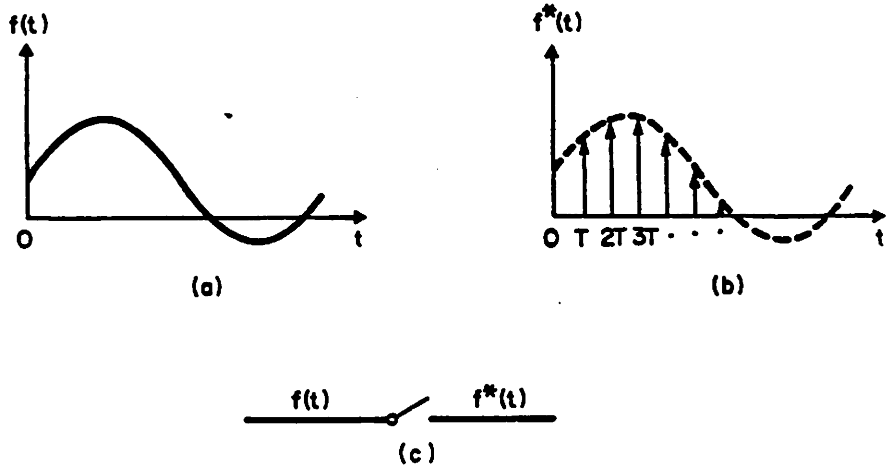

Figure 1. Ideal Sampling, a) a Function of Continuous Time, b) Its Ideal Samples, c) Symbol of Ideal Sampler

Taking the Laplace transform of both sides of (1) yields

$$L[f^*(t)] \triangleq F^*(s) = \sum_{k=0}^{\infty} f(kT)L[\delta(t-kT)]$$
$$= \sum_{k=0}^{\infty} f(kT)e^{-kTs} \tag{2}$$

Now define the complex variable z as

$$z = e^{Ts} \tag{3}$$

Thus $s = T^{-1} \ln z$ and (2) becomes

$$F^*(s)_{s=T^{-1}\ln z} = \sum_{k=0}^{\infty} f(kT)z^{-k} \tag{4}$$

which is defined as the *(one-sided) z transform* of $f(t)$. More formally, we have the following definition.

2. Definition. Let T be a positive constant. The z transform of a function $f(t)$ where $f(t)=0$ for $t<0$ and $f(t)$ is continuous at $t=kT$, $k=0,1,2, \ldots$ is

$$Z[f(t)] \stackrel{\Delta}{=} F(z) = \sum_{k=0}^{\infty} f(kT)z^{-k} \,. \tag{5}$$

Equation (5) is also the z transform of a sequence of scalars $f(kT)$. Δ

Note that the z transform exists if the infinite sum in (5) converges. The *radius of convergence* r_c of the infinite series in (5) is given by

$$r_c = \lim_{k \to \infty} |f(k)|^{1/k} \tag{6}$$

The series in (5) is analytic for $|z| > r_c$†. Thus it converges absolutely for all z in the domain $|z| > r_c$.

3. Example. Consider the unit step function $u(t)$. Then

† Some authors define radius of consequence of the series as $1/r_c$.

$$Z[u(t)] \triangleq U(z) = \sum_{k=0}^{\infty} z^{-k} \tag{7}$$

which converges to

$$\frac{1}{1-\dfrac{1}{z}} = \frac{z}{z-1} \tag{8}$$

if $|z|>1$. Thus the radius of convergence of the series in (7) is 1. Note that (8) is also the z transform of the sequence $u^*(t) = 1,1,1,\cdots$. Although (8) is the z transform of this unique sequence, it does not correspond to a unique function of continuous time since there are many functions whose samples $f(k)$ at $k=0,1,2,\ldots$ are 1 (Fig. 2).

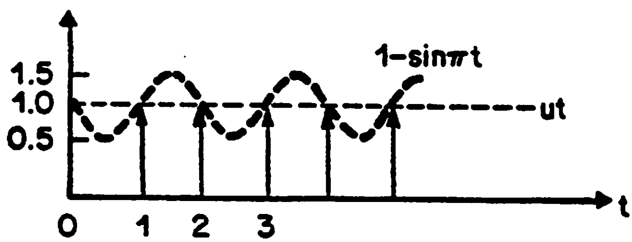

Figure 2. Different Time Functions with the Same Samples

4. Example. Consider the function

$$f(k) = \alpha^{t/T} \tag{9}$$

Then

$$\begin{aligned} Z[f(t)] &\triangleq \sum_{k=0}^{\infty} f(kT)z^{-k} = \sum_{k=0}^{\infty} \alpha^{kT} z^{-k} \\ &= \frac{1}{1-\frac{\alpha}{z}} = \frac{z}{z-\alpha} \end{aligned} \tag{10}$$

provided that $|\frac{\alpha}{z}|<1$ or $|z|>|\alpha|$. Thus the radius of convergence of the infinite series in (10) is $|\alpha|$.

B.3.1 Properties of the z Transform

Since the z transform is defined through the use of the Laplace transform, the properties of the z transform can be deduced from those of the Laplace transform. In this section we will list some useful properties of the z transform without proof. The proofs follow the same lines as in Section B.2.1. We assume that the functions considered in this section vanish for $t<0$ and that they have z transforms with finite radii of convergence. We use $F(z)$ for the z transform of $f(t)$.

5. Uniqueness. As discussed in Example 3, a given function $F(z)$ corresponds to a unique *sequence* $f(k)$. It does not, however, correspond to a unique function of continuous time $f(t)$.

6. Linearity. We have

$$Z[\alpha_1 f_1(t) + \alpha_2 f_2(t)] = \alpha_1 Z[f_1(t)] + \alpha_2 Z[f_2(t)] \tag{11}$$

for arbitrary constants α_1 and α_2.

7. **Advance.** This property of the z transform is analogous to the differentiation rule of the Laplace transform:

$$Z[f(t+mT)] = z^m F(z) - \sum_{z=0}^{m-1} f(kT)z^{m-k} \tag{12}$$

for any positive integer m.

8. Delay. This property of the z transform is analogous to the integration rule of the Laplace transform:

$$Z[f(t-mT)] = z^{-m}F(z) \tag{13}$$

for any positive integer m.

9. Initial value property. We have

$$f(0+) = \lim_{z\to\infty} F(z) \tag{14}$$

provided that the limit exists.

10. Final Value Property. If the sequence $f(k)$ tends to a limit as tends to ∞, then

$$\lim_{k\to\infty} f(kT) = \lim_{z\to 1}(z-1)F(z) \tag{15}$$

if the limit exists.

11. Convolution property. If $F_1(z)$ and $F_2(z)$ are the z transforms of $f_1(t)$ and $f_2(t)$, respectively, then

$$Z\left[\sum_{j=0}^{k} f_1(jT)f_2(k-jT)\right] = F_1(z)F_2(z) \tag{16}$$

For some further properties of the z transform see Problem B.14.

B.3.2 Inversion of the z Transform

Given a function $F(z)$ of the complex variable z, one can find a function $f(.)$ in the time domain such that $Z[f(t)] = [F(z)]$ or equivalently,

$$f(t) = Z^{-1}[F(z)] \tag{17}$$

Note, however, that unlike the case of the Laplace transform, the inverse z transform $f(t)$ is not unique. In fact an infinite number of functions $f(t)$ exist for which $F(z)$ is the z transform. These functions have the same values only at the sampling instants, *i.e.* they can all be represented by the same ideal sampling $f^*(t)$. The methods of determining the inverse z transform can be grouped into three: the power series method, the residue method and the partial fraction expansion method.

B.3.2.1 The Power Series Method

Given $F(z)$, the defining infinite series (5) may be used to determine the inverse $Z^{-1}[F(z)]$ at the given sampling instants. This is done by expanding the rational function $F(z)$ into an infinite series in z^{-k}. Then the coefficient of z^{-k} will be equal to $f(k)$. However, this procedure is, in general, tedious and does not yield a closed-form solution.

12. Example. Consider

$$F(z) = \frac{z}{z-2} \tag{18}$$

Long division results in

$$F(z) = 1 + 2z^{-1} + 4z^{-2} + 8z^{-3} + \cdots \tag{19}$$

Thus, by comparison with Eq. (5), we have

$$f(0) = 1,\ f(1) = 2,\ f(2) = 4,\ f(3) = 8,\ etc. \tag{20}$$

B.3.2.2 The Residue Method

The following theorem is analogous to the Laplace transform inversion theorem.

13. Theorem. Let $F(z) = Z[f(k)]$. Then

$$f(k) = \frac{1}{2\pi j}\int_{\Gamma} F(z)z^{k-1}dz,\ k = 0,1,2,\cdots \tag{21}$$

where integration is performed in counterclockwise direction and Γ is any closed curve enclosing the origin and lying outside the circle $|z|>r_c$ (the radius of convergence of $F(z)$) or, equivalently, Γ is any closed curve enclosing the poles of $F(z)z$, $k = 0,1,2,\cdots$. Δ

The contour integral in (21) can be evaluated by using Cauchy's residue theorem which will be given after the following definition.

14. Definition. The *residue* of a complex function $H(z)$ at a pole $z=\beta$ with multiplicity m of $H(z)$ is

$$\frac{1}{(m-1)!}\frac{d^{m-1}}{dz^{m-1}}[(z-\beta)^m H(z)]|_{z=\beta} \tag{22}$$

(Compare to B.2.33 for $j=1,2,...,m_i$.)

15. Cauchy's Residue Theorem [B.19]. Let $H(z)$ be analytic in R. Then if Γ is a closed curve in R not passing through these singularities,

$$\frac{1}{2\pi j}\int_{\Gamma} H(z)dz = \sum \textit{of the residues of } H(z) \tag{23}$$

where integration is performed in the counterclockwise direction. Δ

Thus, from (21) we have

$$f(kT) = Z^{-1}[F(z)] = \sum \textit{residues of } F(t)z^{k-1} \tag{24}$$

16. Example. Again consider

$$F(z) = \frac{z}{z-\alpha} \tag{25}$$

F(z) has a simple pole at $z=\alpha$. Thus,

$$f(kT) = Z^{-1}[F(z)] = Res\left[\frac{z}{z-\alpha}z^{k-1}\right] = z^k\ |_{z=\alpha} = \alpha^k \tag{26}$$

17. Example. Consider

$$G(z) = \frac{z}{(z-\beta)^2} \tag{27}$$

Then

$$g(kT) = Z^{-1}[G(z)] = Res\left[\frac{z}{(z-\beta)^2}\ z^{k-1}\right] = \frac{d}{dz}(z^k)|_{z=\beta} = \beta^{k-1}k \tag{28}$$

B.3.2.3 The Partial Fraction Expansion Method

The method is similar to that used in determining the inverse z transform in similar to the method used in determining the inverse Laplace transform of a rational function. Instead of $F(s)$ in the case of the Laplace transform, here we expand $\frac{F(z)}{z}$ in partial fractions. The reason for this is that it is usually more convenient to find the

Table 1. z **Transform Pairs**

$f(t)$	$F(z)$
$\delta(t)$*	1
$\delta(t-kT)$	z^{-k}
$u(t)$ (unit step function)	$\dfrac{z}{z-1}$
t	$\dfrac{Tz}{(z-1)^2}$
t^n	$\lim\limits_{\alpha\to 0} (-1)^n \dfrac{\partial^n}{\partial\alpha^n}\left[\dfrac{z}{z-e^{-\alpha T}}\right]$
e^{-at}	$\dfrac{z}{z-e^{\alpha T}}$
te^{-at}	$\dfrac{Tze^{-\alpha T}}{(z-e^{-\alpha T})^2}$
$\sin \alpha t$	$\dfrac{z \sin \alpha T}{z^2-2z\cos \alpha T+1}$
$\cos \alpha t$	$\dfrac{z(z-\cos \alpha T)}{z^2-2z\cos \alpha T+1}$
$e^{-at} \sin bt$	$\dfrac{ze^{-}\alpha f\,2\mathrm{T}\, \sin bT}{z^2-2ze^{-aT} \cos bT+e^{-2aT}}$
$e^{-at} \cos bt$	$\dfrac{z(z-e^{-\alpha T} \cos bT)}{z^2-2ze^{-\alpha T} \cos bT+e^{-2\alpha T}}$

* In discrete form we have $\delta(k) = \begin{cases} 1 & k=0 \\ 0 & k\neq 0 \end{cases}$ which is the *Kronecker delta.*

inverse z transform of a fraction with a free z in its numerator. The following example illustrates this method.

18. Example. Consider

$$F(z) = \frac{z(-2z+1)}{z^2-3z+2} \tag{29}$$

The partial fraction expansion of $F(z)/z$ yields

$$\frac{F(z)}{z} = \frac{1}{z-1} - \frac{3}{z-2} \tag{30}$$

Thus,

$$F(z) = \frac{z}{z-1} - 3\,\frac{z}{z-2} \tag{31}$$

and using (24), the inverse z transform of $F(z)$ is

$$f(k) = 1 - 3(2^k) \tag{32}$$

The z transforms of some common functions and sequences are collected in Table 1. Note that $f(k) = 0$ for $k<0$.

B.3.3 Solution of Difference Equations

As mentioned before, the z transform can be used to solve linear difference equations with constant coefficients. By taking the z transform of both sides of such a difference equation, the z transform of the unknown sequence can be expressed as a rational function of z. The inverse z transform of this function then provides the solution. Thus, the z transform converts the solution of a linear constant-coefficient difference equation to that of an algebraic equation. The following example illustrates the method.

19. Example. Consider the difference equation

$$f(k+2) + 3f(k+1) + 2f(k) = 0, \ k = 0,1,2,\cdots \tag{33}$$

subject to

$$3f(0) + f(1) = 0 \tag{34}$$

Taking the z transform of both sides of (33) yields

$$[z^2F(z) - z^2f(0) - zf(1)] + 3[zF(z) - zf(0)] + 2F(z) = 0 \tag{35}$$

or, using (34),

$$F(z) = \frac{z^2f(0)}{z^2+3z+2} \tag{36}$$

which can be written as

$$\frac{F(z)}{z} = f(0)\left[\frac{-1}{z+1} + \frac{2}{z+2}\right] \tag{37}$$

Thus the solution is

$$f(k) = Z^{-1}[F(z)] = f(0)[-(-1)^k + 2(-2)^k] \tag{38}$$

B.4 THE MODIFIED z TRANSFORM

The modified z transform is an extension of the z transform which is useful in the study of differential-difference equations [B.11], [B.14], [B.15], [B.20]. Such equations arise in the study of hybrid (mixed continuous and discrete) systems and TD systems. In this section we will introduce the modified z transform concept and discuss some of its properties. Application of the modified z-transform to the solution of differential-difference equations will be illustrated through an example.

Consider a function $f(t)$ whose samples are available at times $t=kT$, $k=0, 1, 2, \cdots$ where T, the sampling period, is a positive constant. The modified z transform can be applied to obtain the values of $f(t)$ between the samples as will be explained. Delay the function by ΔT, a function of the sampling period, as shown in Figure 1. If Δ is varied between 0 and 1, then $t=(k-\Delta)T$, $k=0, 1, 2, \cdots$ will cover all values of time and $f(t)$ can be obtained. It is more convenient to let

$\Delta = 1 - m,\ 0 \leqslant m \leqslant 1$. Then

$$t = (k-\Delta)T = (k-1+m)T\ ,\ k=0,1,2,\ldots,\ 0 \leqslant m \leqslant 1 \tag{1}$$

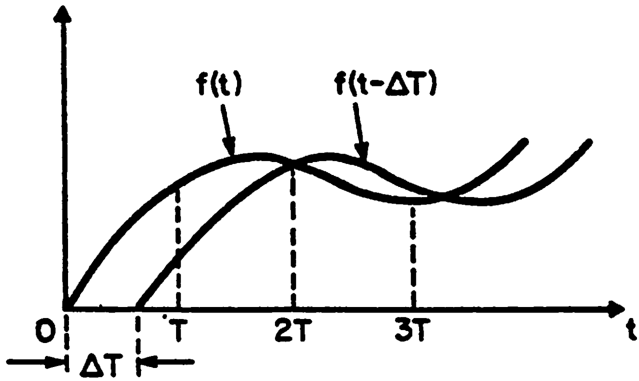

Figure 3. A Delayed Function $f(t)$ to Scan Values at Times Other Than kT, $k=0,1,2,\cdots$

1. Definition. The modified z transform of a function $f(t)$ where $f(t)=0$ for $t<0$ is defined as

$$Z_m[f(t)] \stackrel{\Delta}{=} F(z,m) = \sum_{k=0}^{\infty} f((k-1+m)T)z^{-k}\ ,\ 0 \leqslant m \leqslant 1 \tag{2}$$

Equation (2) is also the modified z transform of a sequence of scalars $f((k-1+m)T)$. Δ

Noting that $f(kT)=0$ for $k<0$, (2) can be written as

$$zF(z,m) = \sum_{k=0}^{\infty} f((k+m)T)z^{-k}\ ,\ 0 \leqslant m \leqslant 1 \tag{3}$$

Thus we have

$$zF(z,0) = \sum_{k=0}^{\infty} f(kT)z^{-k} = Z[f(t)] \tag{4}$$

That is, the z transform can be obtained as a special case of the modified z transform. If there are no discontinuities at the sampling instants, the z transform can also be obtained as follows

$$Z[(t)] = F(z,1) \tag{5}$$

Equation (2) can also be written as

$$F(z,m) = \sum_{k=0}^{\infty} f[(k+m)T-T]z^{-k} = z^{-1} \sum_{k=0}^{\infty} f[(k+m)T]z^{-k} \tag{6}$$

where the delay property of the z transform, has been used.

2. Example Consider the function $f(t)=e^{-2t}$, $t \geqslant 0$. The z transform of this function is

$$F(z) \triangleq \sum_{k=0}^{\infty} f(kT)z^{-k} = \sum_{k=0}^{\infty} e^{-2kT}z^{-k} = \frac{z}{z-e^{-2T}} = \frac{1}{1-e^{-2T}z^{-1}} \tag{7}$$

The modified z transform of $f(t)$ is found by using (6):

$$F(z,m) = z^{-1} \sum_{k=0}^{\infty} e^{-2(k+m)T}z^{-k} = z^{-1} \frac{e^{-2mT}}{1-e^{-2T}z^{-1}} = \frac{e^{-2mT}}{z-e^{-2T}} \tag{8}$$

Note that $zF(z,0) = F(z)$; however, $F(z,1) \neq F(z)$ since $f(t)$ has a discontinuity at $t=0$.

B.4.1 Properties of the Modified z Transform

The properties of the modified z transform can be derived similar to those of the z transform. Some of these properties are given below.

3. Linearity. If $F_i(z,m)$ is the modified z transform of $f_i(t)$, $i=1,2$, then

$$Z_m(\alpha_1 f_1(t)+\alpha_2 f_2(t)] = \alpha_1 F_1(z,m)+\alpha_2 F_2(z,m) \tag{9}$$

where α_1 and α_2 are arbitrary constants α_1 and α_2.

4. Delay. We have

$$Z_m[f(t-kT)] = z^{-k}F(z,m)\ ,\ k=0,1,2,\cdots \tag{10}$$

and for $0<\Delta<1$ if $f(0)=0$, we have

$$Z_m[f(t-\Delta)] = z^{-1}F(z,m-\Delta+1),\ 0\leqslant m\leqslant\Delta$$
$$= F(z,m-\Delta),\ \ \Delta\leqslant m\leqslant 1 \tag{11}$$

5. Initial value property. We have

$$\lim_{\substack{k\to 0\\ m\to 0}} f((k+m)T) = \lim_{\substack{z\to\infty\\ m\to 0}} zF(z,m) \tag{12}$$

and, in particular,

$$\lim_{\substack{k\to 0\\ 0\leqslant m\leqslant 1}} f((k+m)T) = \lim_{\substack{z\to\infty\\ 0\leqslant m\leqslant 1}} zF(z,m) \tag{13}$$

provided that the limits exists.

6. Final value property. We have

$$\lim_{k\to\infty} f(k,m)\ T = \lim_{z\to 1} (z-1)F(z,m) \tag{14}$$

if the limit exists.

7. Differentiation *w.r.t.* **m.** We have

$$Z_m\left[\frac{\partial}{\partial m} f((k-1+m)T)\right] = \frac{\partial}{\partial m} F(z,m) \tag{15}$$

8. Differentiation *w.r.t.* **t.** We have

$$Z_m\left[\frac{\partial^n f(t)}{\partial t^n}\right] = \frac{1}{T^n}\ \frac{\partial^n F(z,m)}{\partial m^n} \tag{16}$$

provided that $\lim_{t\to 0} f^{(i)}(t) = 0$ for $0\leqslant i\leqslant n-1$.

9. Summation of series. We have

$$\sum_{k=0}^{\infty} f((k+m)T) = \lim_{z\to 1} zF(z,m),\ \ 0\leqslant m\leqslant 1 \tag{17}$$

if the sum exists.

B.4.2 Inverse Modified z Transform

Given a modified z transform $F(z,m)$, the value of $f(t)$ for $t=(k-1+m)T$, $k=0,1,2,\ldots,\ 0\leqslant m\leqslant 1$ can be obtained by the inverse modified z transform. This is denoted as

$$f(t)\big|_{t=(k-1+m)T}=Z_m^{-1}[F(z,m)] \tag{18}$$

Methods similar to those for the inverse z transform, *i.e.*, the power series method, the residue method and the partial fraction expansion method, can be used for the inverse modified z transform. For example, the power series method yields

$$\begin{aligned} zF(z,m) &= f(mT)+f(1+m)T\,z^{-1}+f(2+m)Tz^{-2}+\cdots \\ &\quad +f(k+m)Tz^{-k+}\cdots,\quad 0\leqslant m\leqslant 1 \end{aligned} \tag{19}$$

where $f(k+m)$ represents $f(t)$ in the $(k+1)^{st}$ sampling period. The residue method yields the time function in a closed form:

$$f(t)\big|_{t=(k-1+m)T}=\frac{1}{2\pi j}\int_{\Gamma}F(z,m)z^{k-1}dz,\quad 0\leqslant m\leqslant 1,\ k=0,1,2,\cdots \tag{20}$$

where Γ is any closed curve enclosing the poles of $F(z,m)$. Cauchy's residue theorem (eq. (B.3.24) can be used to evaluate the contour integral (20).

10. Example. Consider the function $f(t)=e^{-2t}$, $t\geqslant 0$ whose modified z transform was found in Example 2 as $F(z,m)=\dfrac{e^{-2mT}}{z-e^{-2T}}$. Dividing the numerator of $zF(z,m)$ by its denominator results

$$zF(z,m)=e^{-2mT}+z^{-1}e^{-2(m+1)T}+z^{-2}e^{-2(m+2)T}+\ldots \tag{21}$$

and comparison with (19) yields

$$f(mT)=e^{-2mT},\ f[(1+m)T]=e^{-2(m+1)T},\ f[(2+m)T]=e^{-2(m+2)T},\ldots \tag{22}$$

as expected. We can also use (20) to obtain $f(t)$:

$$
\begin{aligned}
f(t)\big|_{t=(k-1+m)T} &= \textit{Residue of } F(z,m)z^{k-1} \\
&= \textit{Residue of } \frac{e^{-2mT}}{z-e^{-2T}}\, z^{k-1} = e^{-2mT}\, z^{k-1}\big|_{z=e^{-2T}} \\
&= e^{-2T(m+k-1)}
\end{aligned} \tag{23}
$$

which results in $f(t)=e^{-2T}$. Δ

Table 1 shows the modified z transforms of some common functions.

Table 1. Modified z Transform Pairs

$f(t)$	$F(z,m)$
$\delta(t)$	0
$\delta(t-kT)$	z^{m-1-k}
$u(t)$ (unit step function)	$\dfrac{1}{z-1}$
t	$\dfrac{mT}{z-1}+\dfrac{T}{(z-1)^2}$
t^n	$\lim\limits_{\alpha\to 0}(-1)^n \dfrac{\partial^n}{\partial\alpha^n}\left[\dfrac{e^{-\alpha mT}}{z-e^{-\alpha T}}\right]$
e^{-at}	$\dfrac{e^{-amT}}{z-e^{-aT}}$
te^{-at}	$\dfrac{Te^{-amt}[e^{-aT}+m(z-e^{-aT})]}{\left[z-e^{-aT}\right]^2}$
$\sin at$	$\dfrac{z\,\sin amT+\sin\,(1-m)aT}{z^2-2z\,\cos aT+1}$
$\cos at$	$\dfrac{z\,\cos amT-\cos(1-m)aT}{z^2-2z\,\cos aT+1}$
$e^{-at}\,\sin bt$	$\dfrac{e^{-amT}[z\,\sin bmT+e^{-aT}\,\sin(1-m)bT]}{z^2-2ze^{-aT}\,\cos bT+e^{-2aT}}$
$e^{-at}\,\cos bt$	$\dfrac{e^{-amT}[z\,\cos bmT-e^{-aT}\,\cos(1-m)bT]}{z^2-2ze^{-aT}\,\cos bt+e^{-2aT}}$

B.4.3 Solution of Differential-Difference Equations

The modified z transform can be used to solve linear differential-difference equations with constant coefficients. The procedure is similar to the application of the z transform to the solution of different equations.

Let $F(z,m)$ be the modified z transform of $f(t)$. A relation which is very useful in the solution of differential-difference equations is

$$F(z,1) = zF(z,0)-f(0) \tag{24}$$

which follows from the definition of the modified z transform *i.e.* (2). The following example illustrates the procedure.

11. Example. Consider the first-order differential-difference equation

$$\dot{f}(t)-f(t-1) = 0 \tag{25a}$$

with boundary condition

$$f(t) = 2, \quad 0 \leqslant t \leqslant 1 \tag{25b}$$

Assume that $f(t)$ is continuous for $t \geqslant 0$ and let the sampling period T be unity for convenience. Taking the modified z transform of (25a) we have

$$\frac{\partial}{\partial m} F(z,m) - z^{-1}F(z,m) = 0 \tag{26}$$

The solution to (26) is

$$F(z,m) = X(z)e^{m/z} \tag{27}$$

where $X(z)$ must be determined. Using (24) we obtain

$$X(z)e^{1/z} = zX(z)-2 \tag{28}$$

which yields

$$zX(z) = \frac{2}{1-z^{-1}e^{1/z}} \tag{29}$$

Substituting (29) in (27) we obtain

$$zF(z,m) = \frac{2e^{m/z}}{1-z^{-1}e^{1/z}} \tag{30}$$

and expansion in powers of z^{-1} yields

$$\begin{aligned} zF(z,m) &= 2e^{z/m} \sum_{k=0}^{\infty} z^{-k} e^{k/z} = 2 \sum_{k=0}^{\infty} z^{-k} e^{\frac{k+m}{z}} \\ &= 2 \sum_{k=0}^{\infty} z^{-k} \sum_{n=0}^{\infty} \frac{(k+m)^n}{z^n n!} = 2 \sum_{i=0}^{\infty} z^{-i} \sum_{j=0}^{i} \frac{(i+m-j)^j}{j!} \end{aligned} \tag{31}$$

Comparison of (31) with (3) yields

$$f(i+m) = \sum_{j=0}^{i} \frac{(i+m-j)^j}{j!}, \; 0 \leqslant m \leqslant 1 \tag{32}$$

which can be simplified as

$$f(t) = 2 \sum_{j=0}^{i} \frac{(t-j)^j}{j!}, \; i \leqslant t \leqslant i+1, \;\; i=0,1,2, \cdots \tag{33}$$

PROBLEMS

B.1) Show that a sufficient condition for the Laplace transform of $f(t)$ to exist is that $f(t)$ be of *exponential order, i.e.*

$$\lim_{t \to \infty} f(t)e^{kt} = 0$$

for some real number k.

B.2) Show that $exp(t^k)$ where $k > 1$ is not Laplace transformable.

B.3) Prove a) the initial value theorem of the Laplace transform, *i.e.*, show that $f(0+) = \lim_{s \to \infty} s\, L[f(t)]$ provided that $f(0+)$ exists.

b) the final value theorem of the Laplace transform *i.e.*, show that $\lim_{t \to \infty} f(t) = \lim_{s \to 0} sF(s)$ provided that the limit on the left exists.

B.4) Show that the convolution of two time functions is a commutative operation, *i.e.*, prove that

$$f_1 * f_2(t) = f_2 * f_1(t) \quad \text{for } all\ t$$

B.5) Prove the convolution theorem of the Laplace transform, *i.e.*, show that $L[f_1 * f_2(t)] = L[f_1(t]L[f_2(t)]$.

Hint: Since $f_1(t) = f_2(t) = 0$ for $t < 0$, then $f_1 * f_2(t) = \int_0^\infty f_1(\tau) f_2(t-\tau) d\tau$.

Use the definition of the Laplace transform and the Laplace transform translation property.

B.6) Show the following

$$L[tf(t)] = -\frac{d}{ds} F(s)$$

$$L\left[\frac{f(t)}{t}\right] = \int_s^\infty F(\alpha) d\alpha$$

$$L[f(\frac{t}{\alpha})] = \alpha F(\alpha s) \text{ for } any\ scalar\ \alpha.$$

B.7) Show that $L[t^n e^{\alpha t}] = \dfrac{n!}{(s-\alpha)^{n+1}}$ where n is any positive integer and α is any complex scalar.

B.8) If $F(s)$ is a strictly proper rational function with real coefficients show that

a) The residues of real poles of $F(s)$ are real.

b) The residues of a pair of complex conjugate poles of $F(s)$ are complex conjugate.

B.9) Show that

$$L^{-1}\left[\frac{as^2+(10a+b)s+50}{s^3+10s^2+50s}\right] = 1+(a-1)e^{-5t}\cos 5t+(a+\frac{b}{5}-1)e^{-5t}\sin 5t,\ t\geqslant 0$$

B.10. Determine the z transform of a) the sequence $f(k) = \alpha^{k+1}$, $k=0,1,2.$

b) the sequence $g(k) = \alpha^k$, $k=0,1,2,\cdots$

c) the sequence $h(k) = \alpha^{k-1}$, $k=0,1,2,\cdots$

d) the function $h(t) = e^{-\alpha t}$

B.11. Show that

$$Z[te^{-\alpha t}] = \frac{Tze^{-\alpha T}}{(z-e^{-\alpha T})^2}$$

provided that $|z|>e^{-\alpha T}$.

B.12. Prove the final value theorem of the z transform:

$$\lim_{k\to\infty} f(k) = \lim_{z\to 1}(z-1)F(z)$$

provided that F(z) is analytic for $|z|>1$.

B.13. Prove the convolution property of the z transform.

$$Z\left[\sum_{j=0}^{k} f_1(j)f_2(k-j)\right] = Z[f_1(k)]Z[f_2(k)]$$

B.14. Show the following properties of the z transform

a) $Z\left[\frac{f(t)}{t}\right] = \frac{1}{T}\int_{z}^{\infty} \frac{F(z')}{z'}\, dz' + \lim_{t \to 0} \frac{f(t)}{t}$

b) $Z[t\, f(t)] = -\, Tz \,\frac{dF(z)}{dz}$

c) $Z[e^{-aT} f(t)] = F(ze^{aT})$ provided that $|ze^{aT}| >$radius of convergence of $F(z)$.

B.15. Use the z transform method to solve the following difference equation:

$$f(k+3) + 3f(k+2) + 4f(k+1) + f(k) = 2\ , \quad k=0,1,2,\cdots$$

subject to

$$2f(2) + f(1) + 3f(0) = 0$$

B.16. Show the following properties of the modified z transform:

a) $Z_m\left[\frac{f(t)}{t}\right] = \frac{1}{T}\, z^{m-1}\int_{z}^{\infty} z^{-m}\, F(z',m)dz' + \lim_{t \to 0} \frac{f(t)}{t}$

b) $Z_m[tf(t)] = T\left[(m-1)F(z,m) - z\frac{\partial}{\partial z}\, F(z,m)\right]$

c) $Z_m[e^{-at} f(t)] = e^{-aT}(m-1)\; F(ze^{aT},m)$

B.17. Use the modified z transform to solve the following differential-difference equations:

$$3\dot{f}(t+1) + 2\dot{f}(t) + f(t) = 3+t\ , \quad t \geqslant 0$$

with boundary conditions

$$f(t) = 0, \dot{f}(t) = 0, \quad 0 \leqslant t \leqslant 1$$

REFERENCES

[B.1] G. Doetsch, *Introduction to the Theory and Application of the Laplace Transformation*, Springer, 1974

[B.2] W. R. LaPage, *Complex Variables and the Laplace Transform for Engineers*, McGraw-Hill, 1961

[B.3] M. R. Spiegel, *Theory and Problems of Laplace Transforms*, McGraw-Hill, 1965

[B.4] E. J. Watson, *Laplace Transforms and Applications*, Van Nostrand Reinhold, 1981

[B.5] R. E. Bellman and K. L. Cooke, *Differential-Difference Equations*, Academic Press, 1963

[B.6] R. E. Bellman, R. E. Kalaba and J. A. Lockett, *Numerical Inversion of the Laplace Transform*, American Elsevier, 1966

[B.7] M. L. Liou, "A novel method of evaluating transient response," *Proc. IEEE*, Vol. 54, pp. 20-23, 1966

[B.8] R. Piessens, "A bibliography on numerical inversion of the Laplace transform and applications," *J. Comput. Appl. Math.*, Vol. I, No. 2, pp. 115-128, 1975

[B.9] R. Piessens and N. D. P. Dang, "A bibliography on numerical inversion of the Laplace transform and applications: A supplement," *ibid*, Vol. 2, No. 3, pp 225-227, 1976

[B.10] A. Talbot, "The accurate numerical inversion of Laplace transforms," *J. Inst. Math. Appl.*, Vol. 23, pp. 97-120, 1979

[B.11] G. V. Lago, "Additions to *z* transformation theory for sampled-data systems," *Trans. A.I.E.E.*, Vol. 74, part II, pp. 403-408, 1955

[B.12] H. A. Helm, "The *z* transformation," *Bell System Technical Journal*, Vol, 38, pp. 177-196, 1956

[B.13] H. Freeman, and O. Lowenschuss, "Bibliography of sampled-data control systems and z transform applications," *I.R.E. Trans. Automatic Control (PGAC)*, pp. 28-30, March 1958

[B.14] E. I. Jury, *Sampled Data Control Systems*, Wiley, 1958

[B.15] J. R. Ragazzini and G. F. Franklin, *Sampled-data Control Systems*, McGraw-Hill, 1958

[B.16] E. I. Jury, *Theory and Application of the z Transform Method*, Wiley, 1964

[B.17] G. Doetsch, *Guide to the Applications of the Laplace and z Transforms*, Van Nostrand Reinhold, 1971

[B.18] E. I. Jury, and C. A. Galtieri, "A note on the inverse z Transform," *I.R.E. Trans. Circuit Theory*, Vol. CT-9, pp. 371-374, 1961

[B.19] R. V. Churchill, J. W. Brown and R. F. Verhey, *Complex Variables and Applications*, McGraw-Hill, 1974

[B.20] R. H. Baker, "The pulse transfer function and its application to sampling servo systems," *Proc. I.E.E.*, Vol. 99, Part IV, pp. 202-317, 1952

SUBJECT INDEX

Abbreviations, 12
Adaptive Smith Predictor, 414
Adjoint Matrix (see Matrix)
Admissible control, 133, 170
Admissible initial function, 170
Algorithm,
 adjoint matrix, 216
 costate prediction, 270
 costate prediction in TD systems, 323
 coupled linear TD systems, 253
 discrete differential dynamic programming, 394
 goal coordination, 264
 interaction prediction, 267
 interaction prediction in TD systems, 307
 nonserial TD systems, 293
 optimal control, 214
 serial TD systems, 281
 suboptimal control of linear time-invariant TD systems, 188
 suboptimal control of linear time-varying TD systems, 198
 traffic control, 371
Allegheny Ludlum three-stand cold rolling mill, 348
Analysis, 13, 61
Applications, 14, 345

Banach space (see Space)
Bang bang control, 173
Basic variables in three-time-scale systems, 348
Bellman's principle of optimality, 155

Bode plot, 111
for a hydraulic system, 407-408
Boundary-value problem, 181
Bridles in cold rolling mills, 348

CAD example for interaction prediction, 272
Cauchy's residue theorem, 478
Cayley-Hamilton Theorem, 442, 452
Characteristic equation, 99, 103, 107, 116, 123
Characteristic polynomial, 434, 451
Cold rolling mills, 6, 46, 96, 347
as a large-scale system, 352
optimal control design, 353
state model for, 352
Comparison sensitivity function, 228
Computational algorithm,
adjoint matrix, 216
cold rolling mills, 360
costate prediction, 323
hierarchical control of TD systems, 307
nonserial TD systems, 293
optimal control of TD systems, 214
serial TD systems, 281
Conjugate gradient in goal coordination, 263
Continuation, 181
Control,
bang bang, 173
traffic, 371
via Smith Predictor, 409
Controllability, 133
to the origin, 134, 140
Convergence,
abscissa of, 456
Convolution property, 462
Costate prediction, 260, 267-272
algorithm for 270
algorithm for TD systems, 323
computational algorithm for, 252, 253

TD systems, 318
Coupled TD systems, 42
near-optimum control of, 245

Decomposition of large-scale TD systems, 303
Digital control, 33
Digital-to-analog converter, 33
Dirac delta function, 63, 65, 461
Discrete differential dynamic programming (DDDP), 390
algorithm for, 394
Distributed delay, 119
Dynamic programming method, 166

Eigenvalue, 434-437, 451
Eigenvector, 434-437
ϵ - coupling, applied to cold rolling mill's pay-off subsystem, 364-371
generalized, 437, 451
technique, 360-363
Equilibrium state, 77

Fast variables in a three-time-scale systems, 348
Field,
complex, 426
real, 426
FORTRAN
77, 386
IV, 257
Frequency domain, 28, 114
Fundamental matrix, 61, 63-66, 73

Goal coordination, 260-264
algorithm for, 264
conjugate gradient method for, 263
dual variable, 261

global (master) problem, 261
interaction error, 261
interaction variables, 261
Lagrange variables, 261
structure for two-level problem, 262
Gradient vector, 447

HP - 9845 computer, 283, 329
Hamiltonian function, 157, 213, 320
Hamilton-Jacobi-Bellman equation, 155
Hermitian (see Matrix)
Hierarchical control, 260-278
costate prediction, 267
costate prediction for TD systems, 318-337
goal coordination, 260
interaction prediction, 264
interaction prediction for TD systems, 301-317
nonserial TD systems, 285-300
serial TD systems, 279-284
Hilbert space (see Space)
Hold circuit, 33
Housing, mill, 352
Hydraulic system, 9, 405

IBM/1130 computer, 257
IBM 370/165, 386
IBM PC, 272
Infinite dimensional vector space (see Vector space)
Initial function, 25, 63, 137, 149
Initial state, 25, 66
Inner product, 430
Input space, 25
Interaction error, 26
Interaction prediction, 260, 264-267
(see also Hierarchical control)
adjoint (compensation) vector, 266
algorithm for, 267

in TD systems, 301
interaction equality constraint, 265
Interaction variables, 261

Jacobian matrix, 41, 319, 448
Jordan block, 438
Jordan canonical form (see Matrix)

Kernel matrix (see Fundamental matrix)
Kronecker delta, 431

Lagrange multipliers, 261, 262, 304, 320, 373
Lagrangian, 263
Laplace transform, 29, 34, 455-460, 467, 489
Large-scale systems, 42, 57, 245
global (master) problem, 261
hierarchical control, 260-278
hierarchical control of TD systems, 279-337
interaction error, 261
interaction variables, 261
near-optium control for TD systems, 245
Level, hydraulic system, 405
Linear coupled TD systems, 249
Linearization of nonlinear systems, 40, 211, 302
Lipschitz condition, 209
London,
Planning and Transportation Department, 384
West three-intersection network, 385
LSSPAK/PC, a CAD package for large-scale systems, 272
Lyapunov equation, 87, 237
Lyapunov method, 79

Matrix
adjoint, 216
hermitian, 432, 446
Jordan canonical form, 434, 438

modal, 439
normal, 432
orthogonal (or unitary), 432
symmetric, 432
Maximum Principle for TD systems, 14, 157-159, 213
Mills,
cold rolling, 347
housing, 352
stand parameters, 351
Modal (see Matrix)
Model, 23
cold rolling mill, 352
oversaturated intersection, 378
Modeling, 13, 23, 43
Modified z transform (see z transform)
Multiple delays, 65, 73, 148

Near-optimum design of large-scale TD systems, 245-344
Nondelay conversion approach, 202-210
computational algorithm for pth - order suboptimal, 209
multiple constant delays, 202
Nonlinear TD systems, 115
coupled systems, 245
suboptimal control, 210
Nonserial TD systems, hierarchical control for, 285-300
Norm, 432
Normal system (see System)
Nyquist diagram, 107, 110
Nyquist stability criterion, 106

Observability, 133, 142
strong, 142
Optimal control,
computational algorithm for, 214
design of cold rolling mills, 353
for TD systems, 181-244

sensitivity, 181
time, 170
Optimality, Principle of, 155, 167
Optimization, 14, 153
parametric vs. functional, 373
time-delay systems, 155-180
Output equation, 26
Output space, 25

Parameters, mill stand, 351
Parameter variation (see Sensitivity)
Pay-off reel, in cold rolling mills, 348
Penalty term, 156
Performance index sensitivity vector, 227
Perturbation equality constraint, 304
Phase space, 26
Pointwise complete (see System)
Pontryagin's Maximum Principle, 155, 157
Pontryagin's stability method, 99
Principle of optimality, 167
Process, rolling, 348
Pseudo state in nonlinear TD systems, 213
Pulse transform, 34, 36

Quadratic form, 445

Rank, 428
Residue, 464, 477
Riccati equation,
algebraic, 161, 185, 229
differential (DMRE), 204
Riccati, generalized method, 160-165
time-invariant case, 161
time-varying case, 163
Riccati formulation, 159
nonserial TD systems, 287
serial TD systems, 281
Riccati transformulation, 159

Rolling (see Cold rolling mills)
 process, 348
Root locus method, 106

Sampling, 27
Screw down motors in cold rolling mills, 348
Sensitivity, 14
 approach, 181
 approximate modeling, 234
 optimal control for, 181
 parameter variations, 227
 with respect to approximate modeling, 233
 with respect to parameter variations, 229
Serial TD systems, 50
 hierarchical control of, 279
Singular perturbation, 43
 method, 222-227
Slow variable in a three-time-scale system, 348
Smith predictor, 409-419
 adaptive form, 414
 generation of sensitivity functions, 417
Spaceship control, 8
Speed control, 7
Stand parameters in rolling mills, 351
State regulator problem, 160
State equation,
 adjoint, 70
 forced, 66
 homogeneous (or unforced), 61, 75
State space,
 infinite-dimensional, 25
 model for cold rolling mills, 352
 representation, 24
Steepest descent (see Conjugate gradient)
Suboptimal control of TD systems, 181
 computational algorithm for time-invariant case, 188
 nondelay conversion approach, 202-210

nonlinear systems, 210
Nth order, 184
pth order, 209
singular perturbation, 222
time-invariant case, 182-190
time-varying case, 190-199
Superposition property, 68
System
continuous-time, 3, 28
deterministic, 3
discrete-time, 3, 7, 117
distributed parameter, 2
hybrid, 3. 33, 35
linear, 3
lumped-parameter, 2
nonlinear, 3
normal, 171, 172, 174
pointwise complete, 136, 138
sampled-data, 33, 34, 55
stochastic, 3
time-invariant (stationary), 4, 27
time-varying, 4
Symbols, 12
Stability, 77
in the sense of Lyapunov, 77
asymptotic, 78, 84
asymptotic in the large, 78
uniform, 78
uniform asymptotic, 84, 90, 122
Stochastic time-delay systems, 90, 93
stability of, 90
Smith Controller, 103
Spectral decomposition, 144
Space,
Banach, 134
Eucilidean, 427
linear, 426
inner-product, 431
normed, 432
null, 430

Hilbert, 433
Sylvester's theorem, 447
Schwarz inequality, 450
Sampling, 471, 476

Three-time-scale systems, 348, 352-360
Time delay algorithm, 373, 376
control via Smith predictor, 409
Time delay chart, 113, 118
Time optimal control of TD systems, 170-176
Traffic control, 371
Transfer function, 29
matrix, 31
Two-point boundary-value (TPBV) problem, 158, 159, 183, 185, 247-255
costate prediction control of TD systems, 321
interaction prediction, 265
serial TD systems, 279

Unitary matrix (see Matrix)

Variable delays, 93, 95
Vector space
finite dimensional, 429
infinite dimensional, 429

Water resources systems, 8, 52, 308, 389
hierarchical control of, 308-317
Khuzestan, 395-402
Wind tunnel system, 10

z-transfer function, 31, 36
z-transform, 37, 455, 471, 475-479, 490
modified, 34, 37, 455, 481-486

Published volumes in the North-Holland Systems and Control Series

Volume 1:
M.G. SINGH: Decentralised Control, 1981. ISBN 0 444 86198 X.
Volume 2:
R. DAYAL: An Integrated System of World Models, 1981.
ISBN 0 444 86272 2.
Volume 3:
M.A. PAI: Power System Stability, 1981. ISBN 0 444 86310 9.
Volume 4:
S.G. TZAFESTAS (ed.): Optimisation and Control of Dynamic Operational Research Models, 1982. ISBN 0 444 86380 X.
Volume 5:
J. BERNUSSOU / A. TITLI: Interconnected Dynamical Systems: Stability, Decomposition and Decentralisation, 1982.
ISBN 0 444 86504 7.
Volume 6:
M. SHUBIK (ed.): Mathematics of Conflict, 1983.
ISBN 0 444 86678 7.
Volume 7:
S.G. TZAFESTAS (ed.): Applied Digital Control, 1985.
ISBN 0 444 87882 3.
Volume 8:
N. VISWANADHAM / V.V.S. SARMA / M.G. SINGH:
Reliability of Computer and Control Systems, 1987.
ISBN 0 444 70210 5.

ISBN: 0 444 70204