THE REDUCTION OF HIGH ORDER LINEAR DYNAMIC SYSTEMS

THE REDUCTION OF HIGH ORDER LINEAR DYNAMIC SYSTEMS

by

JAMES C. WISMATH, B. Eng.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Engineering

McMaster University

ACKNOWLEDGEMENTS

I wish to express my gratitude to Dr. N.K. Sinha for his encouragement, guidance and many valuable contributions throughout this project.

I would also like to thank my colleagues for their helpful suggestions in several discussions.

Financial assistance was received from the Ontario Government which is recognized with gratitude.

MASTER OF ENGINEERING
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario.

TITLE: The Reduction of High Order Linear Dynamic Systems.

AUTHOR: James Cochrane Wismath, B. Eng. (McMaster University).

SUPERVISOR: N.K. Sinha, B.Sc. Eng. (Banaras),
Ph.D. (University of Manchester).

NUMBER OF PAGES: viii; 108.

SCOPE AND CONTENTS: Three existing techniques are selected as the most promising methods of system reduction. These methods are analysed and then applied to a realistic high order system. A second order model of the actual system is derived using each of the reduction techniques. The step response of the actual system and each of the models is obtained for comparison purposes. The reduction methods are compared with a view to application, limitations and accuracy. A new method for system reduction is also proposed which fashions a low order model after the response characteristics of the actual system.

TABLE OF CONTENTS

Table of Contents Cont'd

Table of Contents Cont'd

LIST OF ILLUSTRATIONS

List of Illustrations Cont'd.

CHAPTER 1

INTRODUCTION

Basic studies of control systems frequently make it desirable to represent a complex system by a low order model. The design, analysis and optimization of systems can commonly be accomplished with greater ease if some model which approximates the system is derived. This situation is much more evident in modern control systems because of their increasingly comprehensive nature and complexity. Processes such as nuclear reactors, high speed rolling mills, jet aircraft controllers, as well as spacecraft systems usually specify fine tolerances of operational limits. This emphasizes the problem of optimization which cannot, in many cases, be practically applied to complex systems.

This need for a low order approximation of a complex system has resulted in several reduction techniques which have been conceived by different approaches to the reduction problem. Since these techniques are basically different, some assessment as to the quality of models that they will produce is necessary.

The quality of a low order model can be judged by comparing the characteristics of the actual system with those of the model. These characteristics are usually

obtained from the response of the system to a step input. Features of the step response which characterize a system include the initial slope, rise time, peak value, time of peak value, settling time and the steady state value. While all the features of a high order system cannot be maintained in a reduced model, it is generally desirable to duplicate, as closely as possible, the overall response of the actual system with emphasis sometimes being placed on one or other region of the response curve.

From the various reduction techniques examined[5,6,7,8,11] three methods that appeared the most rewarding were selected for analysis and application to a high order system. In order that a fair comparison of the methods could be made, a realistic system, which would not be particularly suited for reduction by any one method or another, was required. A proposed design for an aircraft was selected as the test system, since the transfer function of the system contained an even distribution of the poles in the s-plane.

The reduction methods were applied to the test system and in each case a second order model was derived, for the purpose of consistency. Each model was then tested with a step input and response curves were obtained for different periods of time so that the model could be assessed over both the transient and steady state portions of the response. A comparison of the reduction techniques was made with a view to application, limitations and

accuracy.

It became evident during the comparison above, that another approach to the reduction problem might well yield a more suitable model. A method was developed on the basis of this approach which gave the expected results along with the added attractions of ease in application and requirement of very little information about the complex system. In fact, much less knowledge of the system was required for this new method than for any of the methods previously examined.

The material contained in the following chapters follows the order of the preceding discussion. The principle and application of the three selected reduction techniques is discussed in Chapter 2. Chapter 3 deals with the test system which was chosen and its various characteristics. The reduction methods are then applied to the test system in Chapter 4 to obtain three second order models. A comparison of the three methods is made in Chapter 5 which serves to introduce the new approach to system reduction which is proposed in Chapter 6. A method of reduction based on this approach is developed for various systems and then applied to the test system. The conclusions of this work are then drawn in Chapter 7.

CHAPTER 2

METHODS OF REDUCTION

2.0 Introduction

Three methods of reducing high order linear systems were selected because these approaches to the reduction problem appeared to be the most promising. Each method is based on a different concept, and it is these concepts along with the methods of application that will be discussed in this chapter. The methods are examined in turn, beginning in each case with a brief description of the principle involved and then a detailed explanation of the application is given.

In order to simplify matters, the three reduction methods are designated as follows:

Method 1 - Reduction by the continued fraction expansion of the transfer function.[5]

Method 2 - Retention of dominant eigenvalues from the high order system in a low order system.[6,7]

Method 3 - A geometrical approach to the problem of system order reduction.[8]

2.1 Method 1

Method 1 is based upon the expansion of the transfer function of a high order system into a continued fraction which is then truncated according to the order desired for the reduced system. This truncated continued fraction is the transfer function which represents the low order model of the original system. The principle of the method is an expansion about the poles at the origin.

The method is applied by arranging both the numerator and denominator polynomials of the system transfer function in ascending powers of s. The transfer function is then expanded into the following continued fraction form:

$$\frac{C(s)}{R(s)} = \cfrac{1}{H_1 + \cfrac{1}{\cfrac{H_2}{s} + \cfrac{1}{H_3 + \cfrac{1}{\cfrac{H_4}{s} + \ddots}}}}$$

By considering the final value theorem, it is clear that the quotients are in the order of decreasing significance of their contributions to the response value as steady state is approached. Equivalently, as the number of quotients is increased, so higher frequencies are accounted for.

This model does not retain the actual dominant poles of the original system, but rather shifts the poles,

so that they are able to duplicate more closely the response of the original system.

The order of the reduced system is determined by the number of quotients obtained in the expansion. An nth order reduced system requires that 2n quotients be obtained, and the continued fraction may be truncated at this point. Having obtained the required number of quotients, the truncated continued fraction is converted back to the regular transfer function form and this expression represents the transfer function of the reduced system.

The continued fraction expansion can be interpreted in a physical sense as a reconstruction of the block diagram for the system. The new block diagram, however, consists of nested pairs of feedforward and feedback paths, as shown in Figure 2.1. Truncation of the continued fraction corresponds to the removal of the inner pairs of paths. The number of nested pairs remaining represents the size of the reduced system. The formation of these blocks is terminated when the desired order for the reduced system is reached. Figure 2.2 shows the nested pairs that are required for a second order reduced model of a larger system.

2.2 Method 2

The principle of this method is to neglect the

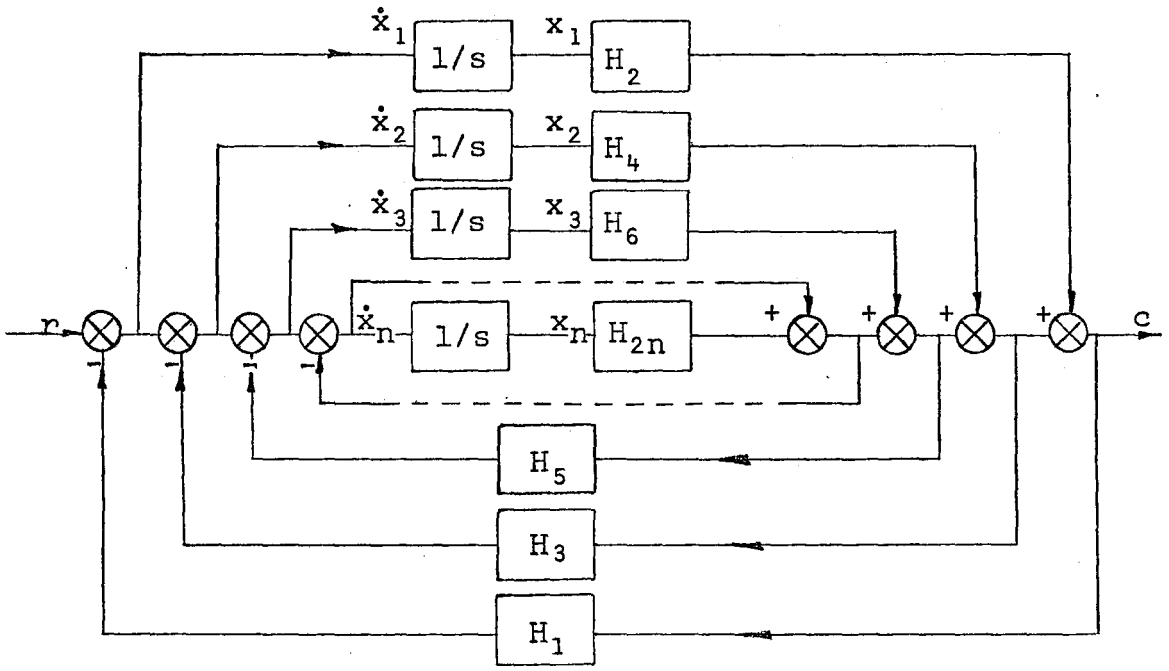


Figure 2.1 Block diagram corresponding to a continued fraction expansion

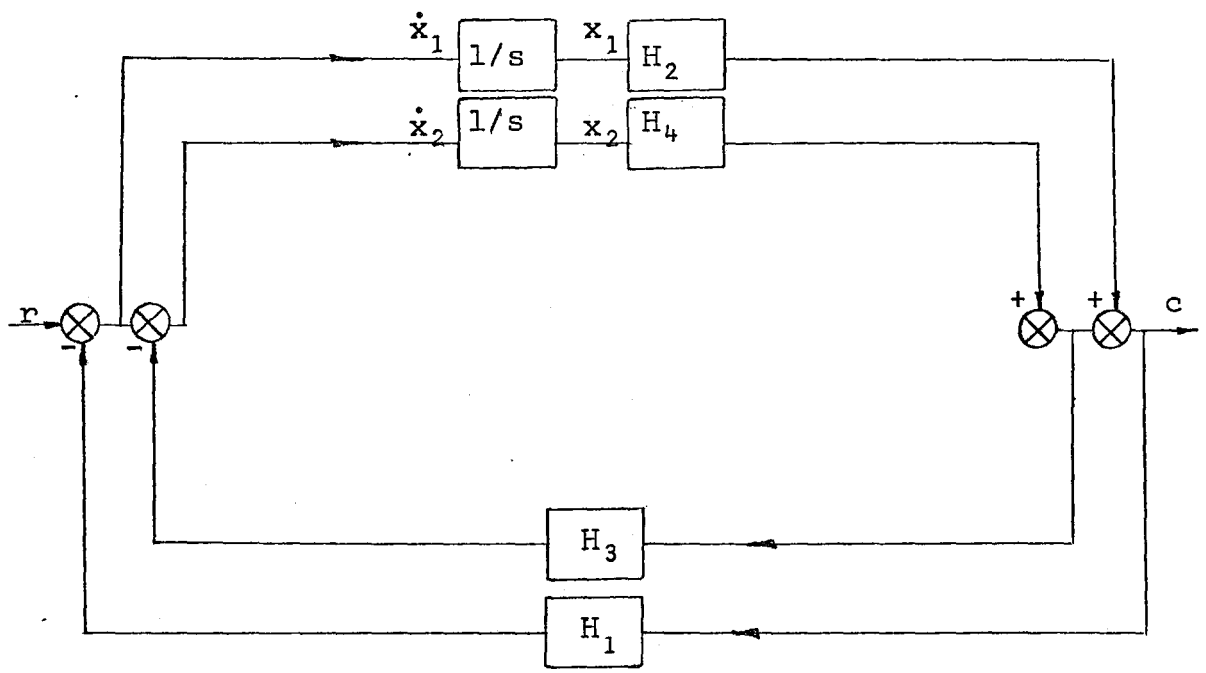


Figure 2.2 Block diagram corresponding to a truncated fraction expansion.

eigenvalues of the original system which are farthest from the $j\omega$ axis of the s-plane and retain only the dominant eigenvalues in the reduced system. Relationships from the time solution of the original system equations are used to develop a reduced system which maintains both the correct proportion of the eigenvectors and the desired eigenvalues. In a further paper[7] the author suggests an additional step in the reduction process which will improve the steady state response of the model. The following detailed outline of the method will show how the reduction is actually accomplished.

In an nth order system, which is to be represented by an ℓth order model, the first ℓ eigenvalues (assuming that they are ordered according to their distance from the $j\omega$ axis of the s-plane) are first determined. Then, ℓ variables x_r, x_s, x_p.... are selected for the reduced system, such that they form large percentages of the first ℓ original eigenvectors. Mathematically this requirement assures that the determinant of the matrix made up of the reduced eigenvectors is not singular; or physically, the requirement means that quantities chosen to represent the system are as different as possible.

The problem now, of finding a reduced system which has the first ℓ eigenvalues selected from the original system and the correct eigenvectors with respect to these

eigenvalues is overcome by considering the time solution of the equation. The nth order original system may be expressed in the form

$$\dot{\underline{x}} = A\underline{x} + B\underline{u}(t) \qquad \text{.....(2.2-1)}$$

and the ℓth order reduced system as

$$\dot{\underline{x}} = A^*\underline{x} + B^*\underline{u}(t)$$

For simplicity, assume that initial conditions are zero, the input is a unit step function and the eigenvalues of A are real and distinct.

Then, the solution for Eqn. (2.2-1) can be written as

$$\underline{x} = \int_0^t \Phi(t-\tau)Bu(\tau)d\tau$$

where $\Phi(t)$ is the transition matrix of the system. In this case,

$$\Phi(t-\tau) = \Gamma\exp[\Lambda(t-\tau)]\Gamma^{-1}$$

where Λ = the diagonal matrix of eigenvalues

and Γ = the modal matrix, composed of the eigenvectors of A (arranged in the order of corresponding eigenvalues).

i.e. $$\Gamma = \begin{bmatrix} x_1^1 & x_1^2 & --x_1^n \\ x_2^1 & x_2^2 & --x_2^n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_n^1 & x_n^2 & --x_n^n \end{bmatrix}$$

Then, by denoting

$$\Gamma^{-1} = \begin{bmatrix} \phi_1^1 & \phi_1^2 & --\phi_1^n \\ \phi_2^1 & \phi_2^2 & --\phi_2^n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \phi_n^1 & \phi_n^2 & --\phi_n^n \end{bmatrix}$$

the time solution of Eqn. (2.2-1) can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \frac{-1+e^{\lambda_1 t}}{\lambda_1} \begin{bmatrix} x_1^1 \\ x_2^1 \\ \cdot \\ \cdot \\ x_n^1 \end{bmatrix} (\phi_1^1 b_1 + \phi_1^2 b_2 + --+ \phi_1^n b_n) + --$$

$$+ \frac{-1+e^{\lambda_n t}}{\lambda_n} \begin{bmatrix} x_1^n \\ x_2^n \\ \cdot \\ \cdot \\ x_n^n \end{bmatrix} (\phi_n^1 b_1 + \phi_n^2 b_2 + --+ \phi_n^n b_n) \qquad .....(2.2\text{-}2)$$

Now, if only the first ℓ time constants of Eqn. (2.2-2) are retained, the time solution becomes

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x \end{bmatrix} = \varepsilon_1 \begin{bmatrix} x_1^1 \\ x_2^1 \\ \cdot \\ \cdot \\ x^1 \end{bmatrix} + \varepsilon_2 \begin{bmatrix} x_1^2 \\ x_2^2 \\ \cdot \\ \cdot \\ x^2 \end{bmatrix} + \cdots + \varepsilon_\ell \begin{bmatrix} x_1^\ell \\ x_2^\ell \\ \cdot \\ \cdot \\ x^\ell \end{bmatrix} \quad \text{.....(2.2-3)}$$

where

$$\varepsilon_i = \frac{-1+e^{\lambda_i t}}{\lambda_i}(\phi_i^1 b_1 + \phi_i^2 b_2 + \cdots + \phi_i^n b_n)$$

In addition, if only the ℓ variables x_r, x_s, x_p, ... are retained, the equations become

$$\begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix} = \varepsilon_1 \begin{bmatrix} x_r^1 \\ x_s^1 \\ x_p^1 \\ \cdot \\ \cdot \end{bmatrix} + \varepsilon_2 \begin{bmatrix} x_r^2 \\ x_s^2 \\ x_p^2 \\ \cdot \\ \cdot \end{bmatrix} + \cdots + \varepsilon_\ell \begin{bmatrix} x_r^\ell \\ x_s^\ell \\ x_p^\ell \\ \cdot \\ \cdot \end{bmatrix}$$

or

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_\ell \end{bmatrix} = \begin{bmatrix} x_r^1 & x_r^2 & -- & x_r^\ell \\ x_s^1 & x_s^2 & -- & x_s^\ell \\ x_p^1 & x_p^2 & -- & x_p^\ell \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}^{-1} \begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix}$$

$$= \Lambda_0^{-1} \begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix} \qquad \text{.....(2.2-4)}$$

The equations for the remaining variables can be expressed as

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_{r-1} \\ x_{r+1} \\ \cdot \\ \cdot \\ x_{s-1} \\ x_{s+1} \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} x_1^1 & - - - & x_1^\ell \\ x_2^1 & - - - & x_2^\ell \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_{r-1}^1 & - - - & x_{r-1}^\ell \\ x_{r+1}^1 & - - - & x_{r+1}^\ell \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_{s-1}^1 & & x_{s-1}^\ell \\ x_{s+1}^1 & & x_{s+1}^\ell \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_n^1 & & x_n^\ell \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_\ell \end{bmatrix} = \Lambda_1 \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_\ell \end{bmatrix}$$

Now, if the ℓ variables x_r, x_s, x_p... of

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} a_1^1 & a_1^2 & -- & a_1^n \\ a_2^1 & a_2^2 & -- & a_2^n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_n^1 & a_n^2 & & a_n^n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

are kept, then

$$\begin{bmatrix} \dot{x} \\ \dot{x}_s \\ \dot{x}_p \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} a_r^r & a_r^s & a_r^p & -- \\ a_s^r & a_s^s & a_s^p & -- \\ a_p^r & a_p^s & a_p^p & -- \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{bmatrix} \begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix} + \begin{bmatrix} a_r^1 & a_r^2 - a_r^{r-1} & a_r^{r+1} - a_r^n \\ a_s^1 & a_s^2 - a_s^{r-1} & a_s^{r+1} - a_s^n \\ a_p^1 & a_p^2 - a_p^{r-1} & a_p^{r+1} - a_p^n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_{r-1} \\ x_{r+1} \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

$$= A_0^* + A_1^* \Lambda_1 \Lambda_0^{-1} \begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix}$$

or, $$\begin{bmatrix} \dot{x}_r \\ \dot{x}_s \\ \dot{x}_p \\ \cdot \\ \cdot \end{bmatrix} = A^* \begin{bmatrix} x_r \\ x_s \\ x_p \\ \cdot \\ \cdot \end{bmatrix} \quad \text{where } A^* = A_0^* + A_1^* \Lambda_1 \Lambda_0^{-1}$$

This is the required A* matrix for the reduced system, which will have the same eigenvalues and vectors as the first ℓ eigenvalues and vectors of the A matrix from the original system. Since A* has the same eigenvalues and vectors as the matrix A in the original system, the reduction procedure can be somewhat simplified. The matrix A* can be more easily determined as

$$A^* = U_1 \lambda_1 U_1^{-1}$$

where U_1 is an ℓ x ℓ matrix consisting of elements of the 1st ℓ dominant eigenvectors of A which correspond to the retained variables x_r, x_s, x_p ...., and ℓ is an ℓ x ℓ diagonal matrix consisting of the ℓ dominant eigenvalues of A.

By a similar approach, the B matrix from the original system can be reduced to a simpler form, obtained according to the equations following:

$$B^* = \Gamma^*[\Gamma^{-1}B]$$

$$\text{where } \Gamma^* = \begin{bmatrix} x_r^1 & x_r^2 & -- & x_r^\ell \\ x_s^1 & x_s^2 & -- & x_s^\ell \\ x_p^1 & x_p^2 & -- & x_p^\ell \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

$$\text{and } \Gamma^{-1}B = \text{the first } \ell \text{ rows of } \begin{bmatrix} x_1^1 & x_1^2 & -- & x_1^n \\ x_2^1 & x_2^2 & -- & x_2^n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_n^1 & x_n^2 & & x_n^n \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$$

After selecting the first ℓ dominant eigenvalues and the ℓ variables to be retained in the reduced system, the matrices A* and B* are evaluated by using the relationships above.

The modification suggested by the author in a later paper[7], (which improves the steady state response of the model) is described as a matrix transformation. Consider the reduced system as previously defined, with only a single input. The new model is proposed in the form

$$\dot{\underline{z}} = DA^*D^{-1}\underline{z} + DB^*u$$

where z is an ℓ-vector consisting of the ℓ-vectors x_r, x_s, x_p... and D is a diagonal matrix defined as

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_\ell \end{bmatrix}$$

$$\text{where } d_j = \frac{[A^{-1}B]_{j*}}{[A^{*-1}B^*]_j} \quad \text{if } [A^{*-1}B^*]_j \neq 0$$

$$= 1 \quad \text{if } [A^{*-1}B^*]_j = 0 \Bigg| \; j = 1,2,\text{--}\ell.$$

and $[A^{*-1}B^*]_j$ is the jth element of the ℓ vector $A^{*-1}B^*$, and $[A^{-1}B]_{j*}$ is the element of the n vector $A^{-1}B$ which corresponds to the jth state retained in the simplified system. The reduced model is equivalent to the following system:

$$\dot{\underline{x}} = A^*\underline{x} + B^*u$$

$$\underline{z} = D\underline{x}$$

In summary then, after selecting the eigenvalues and state variables for the reduced system, the matrices A*, B* and D are determined in order to arrive at the simplified model.

2.3 Method 3

Anderson presents a method of system reduction based on a geometrical consideration of the reduction problem as developed from the theory of linear vector spaces. A simplified explanation of this method will be given first, to show the idea behind Anderson's technique.

The state space equations for the given high order system are solved at regular time intervals up to some limit at which only insignificant response changes occur. These solutions are then substituted in the state space equations which represent the unknown low order system and the parameters are evaluated which will give these solutions. A best value for these parameters is based on a 'least squares fit' between the solutions of the two sets of state space equations. Thus, given all the information about a high order system, a low order model is developed which is based on the requirement that it match the response of the high order system, in a least squares sense, over a given time interval.

The way in which the reduction is actually accomplished may be better explained by examining the problem from a geometrical point of view and then using the theory of linear vector spaces to obtain the desired results.

A linear, time-invariant dynamic system may be

described by the vector equation

$$\dot{\underline{z}} = F\underline{z} + G\underline{u}$$

$$\text{and} \quad \underline{y} = H\underline{z} \qquad \text{.....(2.3-1)}$$

where $\underline{z}$ is an n-vector of state variables, $\underline{u}$ is an m-vector of inputs and $\underline{y}$ is a p-vector of outputs. If none of the elements of the vector $\underline{y}$ is a linear combination of the others, Eqn. (2.3-1) can be represented as

$$\dot{\underline{x}} = A\underline{x} + B\underline{u} \qquad \text{.....(2.3-2)}$$

in which the first p elements of the vector $\underline{x}$ are the elements of the vector $\underline{y}$.

If the inputs $\underline{u}$ are held constant between successive measurements, Eqn. (2.3-2) can be solved to obtain the state transition equation

$$\underline{x}\left[(k+1)T\right] = \Phi(T)\underline{x}(kT) + \Delta(T)\underline{u}(kT) \qquad \text{.....(2.3-3)}$$

in which T is the sampling period

$\Phi(T)$ is the transition matrix

$\Delta(T)$ is the driving matrix

and $$\Phi(T) = \sum_{k=o}^{\infty} \frac{A^k T^k}{k!}$$

and $$\Delta(T) = \sum_{k=o}^{\infty} \frac{A^k T^{k+1}}{(k+1)!}.B$$

The reduction problem is now to replace Equations

(2.3-2) or (2.3-3) by a low order system. For an rth order system with q inputs, Eqn.(2.3-2) can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdot\cdot & a_{1r} \\ a_{21} & a_{22} & \cdot\cdot & a_{2r} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{r1} & a_{r2} & & a_{rr} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_r \end{bmatrix} + \begin{bmatrix} b_{11} & \cdot\cdot & b_{1q} \\ b_{21} & & \\ \cdot & & \\ \cdot & & \\ b_{r1} & \cdot\cdot & b_{rq} \end{bmatrix} \begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ uq \end{bmatrix} \quad .....(2.3\text{-}4)$$

Let ℓ be the desired order of the reduced system and m be the number of retained inputs.

In Eqn. (2.3-4), if $\underline{u}$ is held constant between sampling instants, the variables $\underline{\dot{x}}$ and $\underline{x}$ can be evaluated at n instants with time intervals of T seconds.

Now, each of the low order state space equations can be represented similarly to the first equation at every sampling instant. The first line of Eqn. (2.3-4) can thus be written at each of the n sampling instants to give the following set of equations:

$$c_1 x_{11} + c_2 x_{12} + \text{---} + c_\ell x_{1\ell} = W_1$$

$$\vdots$$

$$c_1 x_{n1} + c_2 x_{n2} + \text{---} + c_\ell x_{n\ell} = W_n$$

where $\underline{W}^T = [\dot{x}_1(0) \quad \dot{x}_1(T) \text{ --- } \dot{x}_1(kT)]$

$\underline{C} = [a_{11}\ a_{12} \cdots a_{1\ell}\ b_{11} \cdots b_{1m}]$

$\underline{X}_{11}^T = [x_1(0) \quad x_1(T) \text{ --- } x_1(kT)]$

$\vdots$

$\underline{X}_{1\ell}^T = [x_\ell(0) \quad x_\ell(T) \text{ --- } x_\ell(kT)]$ .....(2.3-5)

and ℓ = order of the reduced system

m = number of inputs to the system

T = sampling interval

$n = k+1$ = number of sampling instances.

The exact values for $\underline{W}$ and $\underline{x}_1$, $\underline{x}_2$ -- $\underline{x}_\ell$ can be obtained from the solution of the equations describing the actual system, and the problem then is to solve the linear Eqns. (2.3-5) for the coefficients c_1, c_2, --- c_ℓ. In general, this is not possible, since numbers for c_1, c_2, --- c_ℓ cannot be found which will satisfy all the equations of the system.

As an approximation, however, coefficients

c_1, c_2, --- c_ℓ may be searched for that will minimize the sum of the squared errors between the actual values of the R.H.S. of Eqns. (2.3-5) and the values obtained when suitable coefficients are used. This problem can be readily solved when it is interpreted geometrically

Let $\underline{X}_1$, $\underline{X}_2$, --- $\underline{X}_\ell$ be the ℓ vectors

$$\underline{X}_1^T = [x_{11}\ x_{21}\ ---\ x_{n1}]$$

$$\underline{X}_2^T = [x_{12}\ x_{22}\ ---\ x_{n2}]$$

$$\vdots$$

$$\underline{X}_\ell^T = [x_{1\ell}\ x_{2\ell}\ ---\ x_{n\ell}]$$

and $$\underline{W}^T = [w_1\ w_2\ ---\ w_n]$$

where ℓ = reduced system order

n = number of sampling instances.

Then, for the coefficients which could be selected, some values D_1, D_2 --- D_n will result in the R.H.S. of Eqns. (2.3-5). In vector notation,

$$\underline{D} = c_1\underline{x}_1 + c_2\underline{x}_2 + --- + c_\ell\underline{x}_\ell$$

and the problem is to choose coefficients c_1, c_2, --- c_ℓ so that the inner product $\langle \underline{W}-\underline{D}, \underline{W}-\underline{D} \rangle$ is minimized. If the vectors $\underline{X}_1, \underline{X}_2$, --- $\underline{X}_\ell$ are considered to form a subspace

$L(\underline{X}_1, \underline{X}_2, ---, \underline{X}_\ell)$ spanned by the vectors $X_1, X_2, ---, X_\ell$, then the orthogonal projection of the vector $\underline{W}$ on that subspace is the vector in L which is closest to $\underline{W}$. This can be seen by expanding the vector $\underline{W}$ as the sum $\underline{g} + \underline{h}$, where $\underline{g}$ is in the subspace L and $\underline{h}$ is orthogonal to this space. The inner product to be minimized becomes

$$\langle(\underline{g}+\underline{h}-\underline{D}),(\underline{g}+\underline{h}-\underline{D})\rangle$$

$$= \langle\underline{g},\underline{g}\rangle - 2\langle\underline{g},\underline{D}\rangle + \langle\underline{h},\underline{h}\rangle + \langle\underline{D},\underline{D}\rangle$$

as both $\underline{g}$ and $\underline{D}$ lie in the subspace L and $\underline{h}$ is orthogonal to it. The inner product is clearly minimized when $\underline{D}$ is equal to $\underline{g}$, the projection of $\underline{W}$ on the subspace L.

The coefficients $c_1, c_2, ---, c_\ell$ must therefore be chosen in such a way that $\underline{D}$ is the projection of the vector $\underline{W}$ on the subspace L spanned by the vectors $X_1, X_2, ---, X_\ell$.

If we express the vector $\underline{W}$ as $\underline{g} + \underline{h}$, and minimize the error by letting $\underline{D} = \underline{g}$, then $\underline{W} = \underline{D} + \underline{h}$.

But, $\underline{D} = c_1\underline{x}_1 + c_2\underline{x}_2 + ---+ c_\ell\underline{x}_\ell$

Then, using the requirements for orthogonality between the subspace L and $\underline{h}$, we have

$$\langle \underline{h}, \underline{x}_1 \rangle = \langle (\underline{W}-\underline{D}), \underline{x}_1 \rangle = \langle \underline{W}, \underline{x}_1 \rangle - c_1 \langle \underline{x}_1, \underline{x}_1 \rangle - \cdots - c_\ell \langle \underline{x}_1, \underline{x}_\ell \rangle = 0$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\langle \underline{h}, \underline{x}_\ell \rangle = \langle (\underline{W}-\underline{D}), \underline{x}_\ell \rangle = \langle \underline{W}, \underline{x}_\ell \rangle - c_1 \langle \underline{x}_\ell, \underline{x}_1 \rangle - \cdots - c_\ell \langle \underline{x}_\ell, \underline{x}_\ell \rangle = 0$$

The required coefficients $c_1, c_2, \cdots, c_\ell$ are then obtained by solving these equations using Cramer's rule. Since the solution to this problem exists and is unique[10], the Grammian from the above equation is non-zero.

The coefficients in Eqn. (2.3-5) will thus give the smallest deviation (in a least squares sense) from the actual values for the right hand side of this equation. The elements of the A and B matrices for the reduced system are, of course, these coefficient values so that the reduced system is now determined.

The three selected methods of reduction have been discussed in detail as to their application. It remains now to find a suitable system to which these methods may be applied. The next chapter deals fully with the choice of such a system.

CHAPTER 3.

THE TEST SYSTEM

3.0 Introduction

This chapter is concerned with the 'test system' which was selected as an example of a high order system. Characteristics of this system which are required in later chapters, will be outlined here and the methods used to find certain of the characteristics will be discussed.

For the sake of reality and validity, it was decided that an authentic system model containing a reasonable number of complex poles, fairly distributed in the s-plane, should be selected.

Reduction techniques invariably neglect (to some extent) the poles of the system which are located far from the origin since these poles represent small time constants and their effects are relatively short-lived. Some authors have taken examples in which there are two sets of poles, one near the $j\omega$-axis, and the other far away from it. This is not very realistic. To avoid biassed results, then, a transfer function with poles distributed over the entire left-half plane was preferred.

3.1 Description of the Test System

The system chosen[1] is one of the designs studied for the current super-sonic transport aircraft. Figure 3.1 shows the block diagram of the system with variable parameters K1, K2, τ, ζ and ω_n. To obtain reasonable pole-zero locations, the following parameter values were selected (consistent with design description and maintaining stability):

$$K1 = 0.2$$

$$K2 = 1.0$$

$$\omega_{n1} = 2.5 \text{ [r/s]}$$

$$\tau_1 = 12$$

$$\zeta = 0.707$$

3.2 System Representation in the s-plane.

With these parameter values, the transfer function of the test system becomes

$$\frac{C(s)}{R(s)} = \frac{375000(s + 0.0833)}{s^7 + 83.635s^6 + 4097.4035s^5 + 70341.905s^4 + 853703.3s^3 + 2814271s^2 + 3310875s + 281250}$$

The characteristic equation of the system is given by

F(s) = Denominator of Transfer Function(s)

Originally, the roots of the characteristic

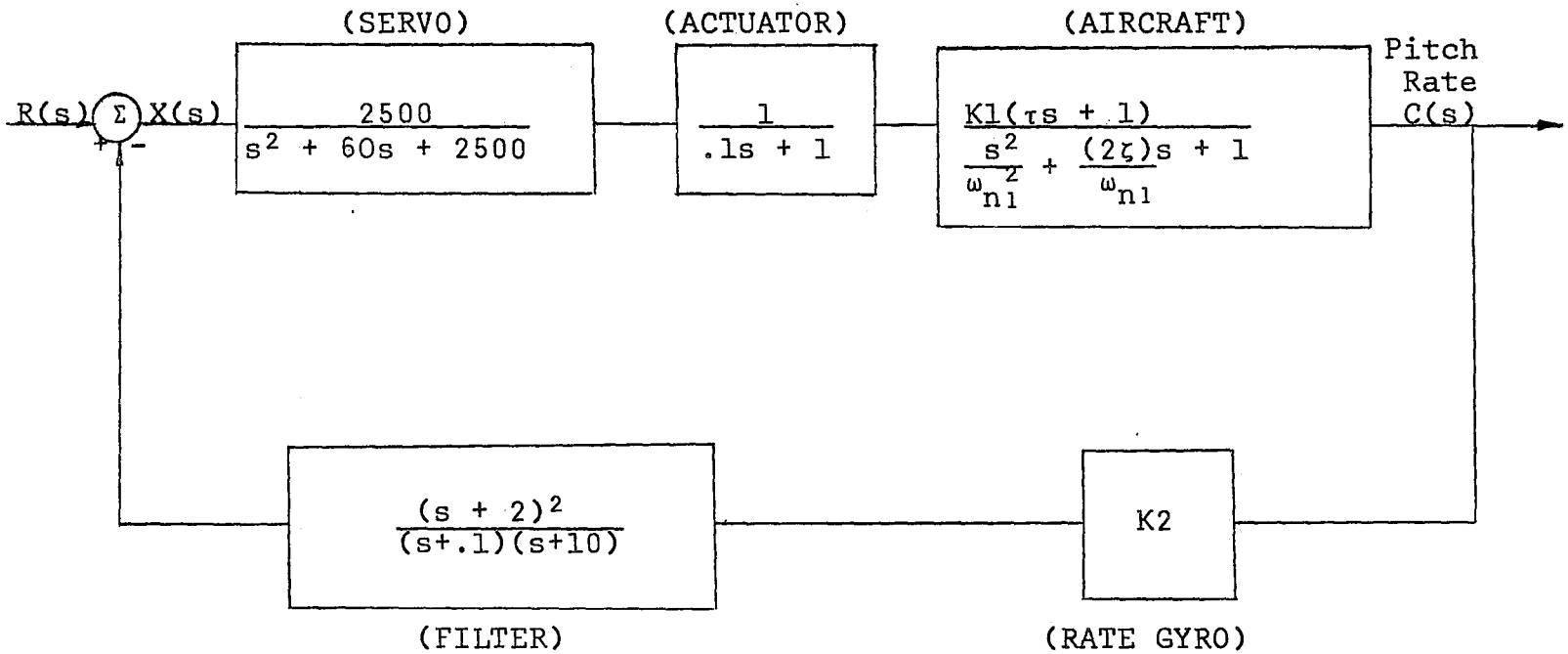


Figure 3.1 Block Diagram for a Super Sonic Transport Airplane Design.

equation were obtained by solving the polynomial using the Newton-Raphson technique on a CDC6400 computer. Errors resulted, however, due to 'rounding-off' within the computer and a more successful approach is described in the next section.

3.3 State Space System Representation

From the transfer function, the state space equations were developed (see Appendix A.1) in the form

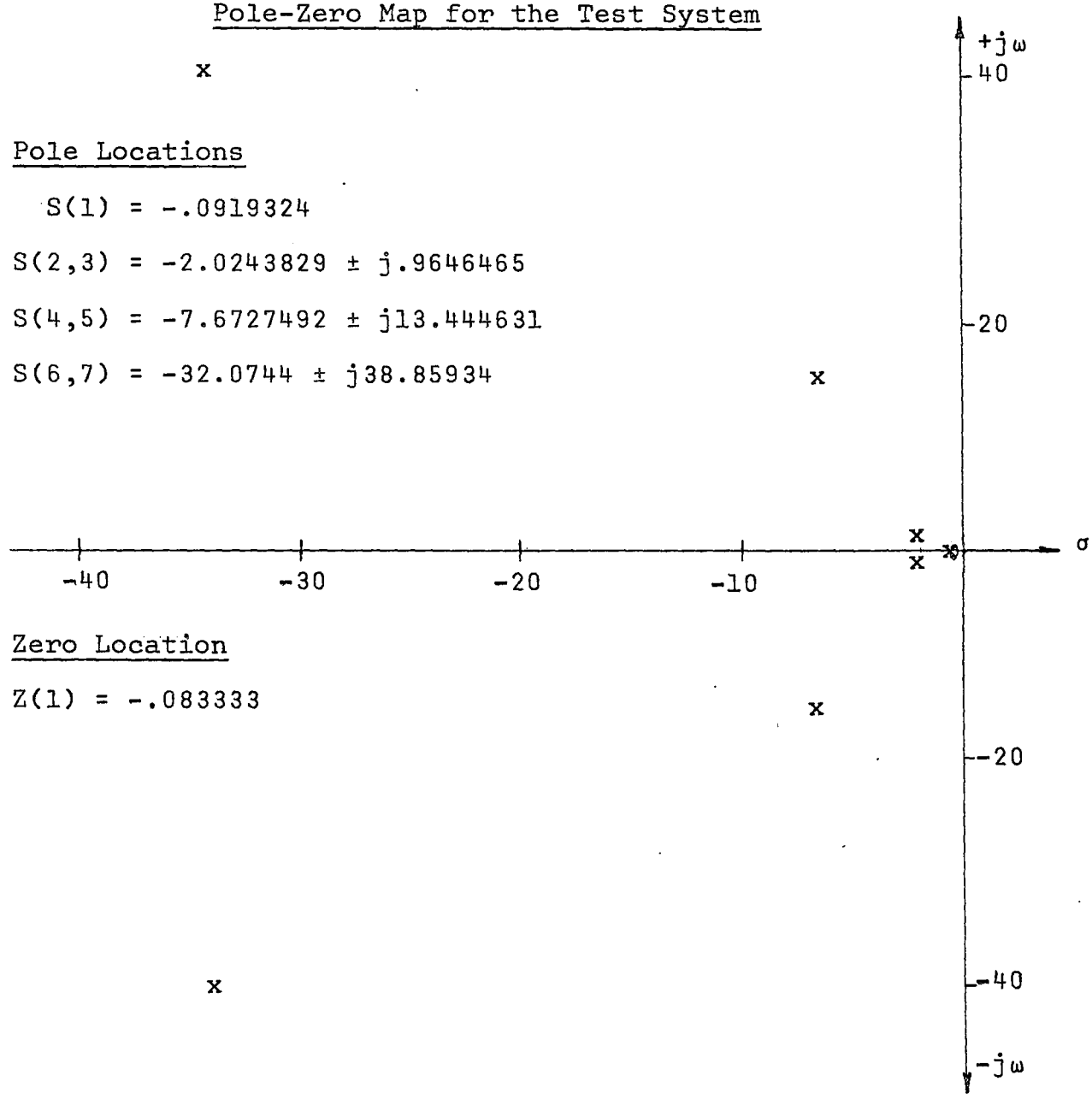
$$\dot{\underline{x}} = A\underline{x} + B\underline{u} \qquad .....(3.3\text{-}1)$$

$$c = H\underline{x} \qquad .....(3.3\text{-}2)$$

It is assumed throughout, that the variables x, y, u, and c are understood to be functions of time, and for simplicity (t) is omitted from the state space equations. Now, the eigenvalues of A will be the required roots of the characteristic equation. Using the iterative method of Eberlein[3], these eigenvalues were obtained to a satisfactory degree of accuracy on the computer (see Appendix A.2).

The eigenvectors of A, which are required later, were also obtained and appear in Appendix A.2. The roots of the characteristic equation represent the poles of the system and the pole-zero pattern is displayed in Figure 3.2.

Pole-Zero Map for the Test System



Pole Locations

S(1) = -.0919324

S(2,3) = -2.0243829 ± j.9646465

S(4,5) = -7.6727492 ± j13.444631

S(6,7) = -32.0744 ± j38.85934

Zero Location

Z(1) = -.083333

Figure 3.2 Pole-Zero Pattern for the Test System.

3.4 Solution of the State Space Equations

State space equations in the form of equation (3.3-1) have the solution

$$\underline{x}(t) = e^{At}\left[\underline{x}(o) + \int_0^t e^{-A\tau} B u(\tau) d\tau\right] \quad .....(3.4\text{-}1)$$

This equation can be solved by direct integration and then substitution of the various required values for t. Another approach to the solution of this equation, which is particularly adaptable to machine computation, is outlined in Appendix B. Using the series expansion for the exponential terms and then an iterative procedure, which requires solving the equation for a very small interval ($\Delta T << t$), $\underline{x}$ is calculated with continuously updated values for $\underline{x}(0)$.

Direct evaluation of x(t) for any significant time t involves round-off errors in the computer giving unsatisfactory results. In addition to improved accuracy, the method outlined above also lends itself well to the generation of data cards for plotting the system response. A graph of the system response is, of course, a natural conclusion to the problem.

A general computer program for the solution of equations in the form of equation (3.3-1) is also included in Appendix B.

3.5 Response to a Step Input

Using the method outlined above, the response to a unit step input is plotted in figures 3.3 and 3.4 for 5 second and 20 second time bases, respectively.

Features of the response which will be used later are:

1. (10-90)% Rise Time = 1 second.
2. Steady State Value = 0.11111 (verified by calculation).
3. Initial Slope = 0.0 (obtained by the initial value theorem).
4. Maximum overshoot = 0.120689 (obtained from the plotted response).
5. Time to reach the first maximum (t_p) = 2.9 sec. (obtained from the plotted response).

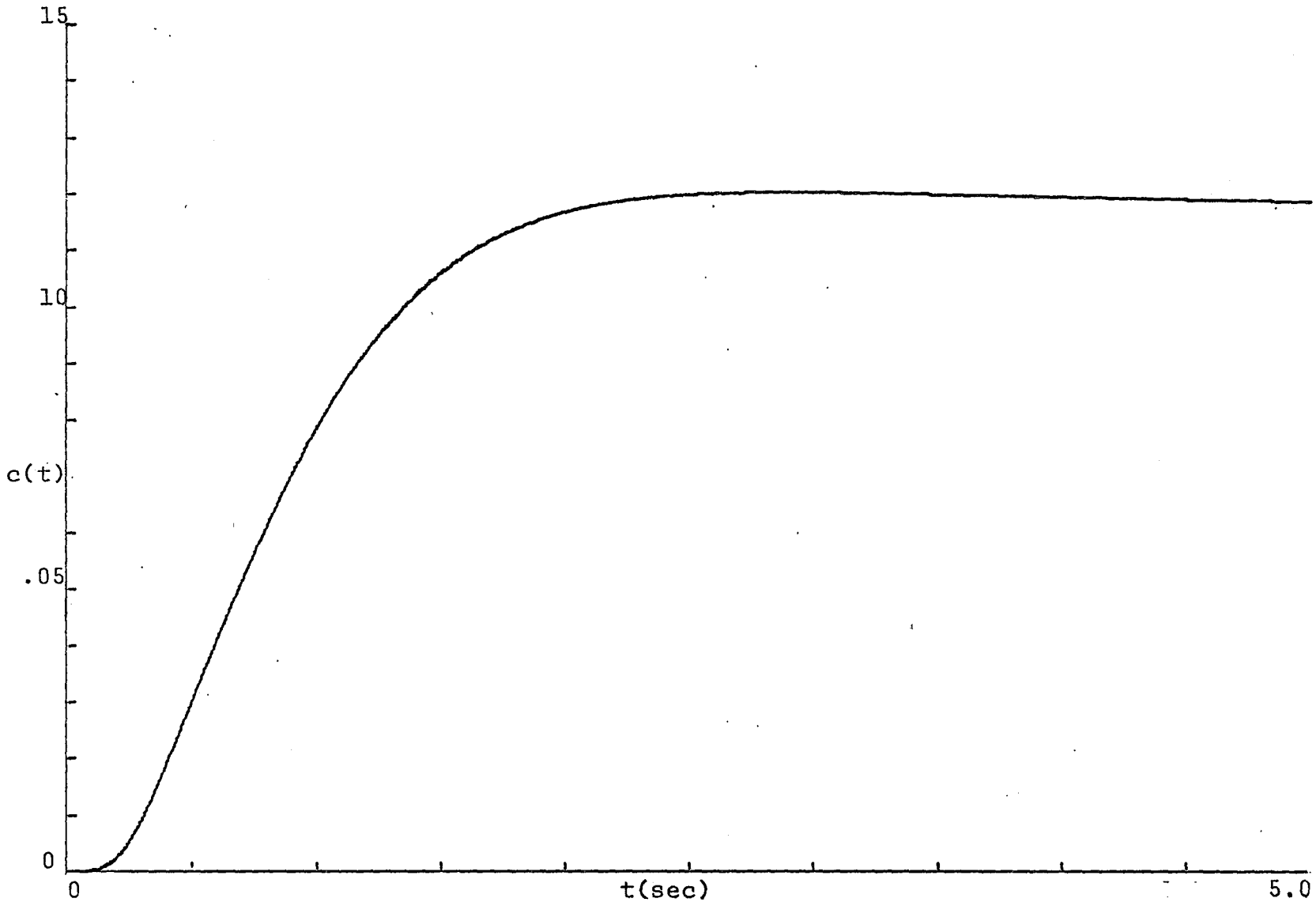


Figure 3.3 Response of the Test System to a Step Input

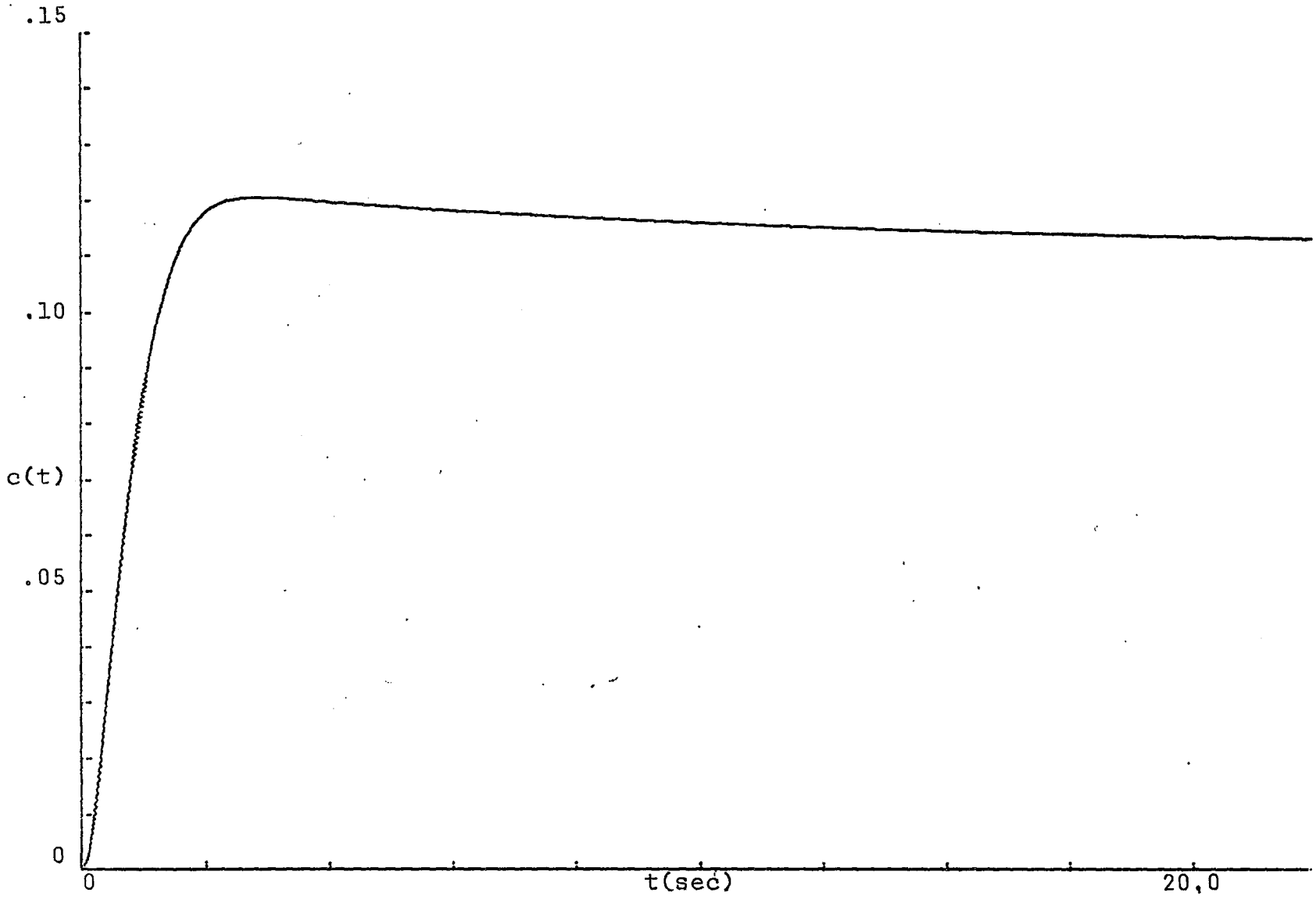


Figure 3.4 Response of the Test System of a Step Input

CHAPTER 4

APPLICATION OF THE VARIOUS METHODS OF REDUCTION TO THE TEST SYSTEM

4.0 General

This chapter deals with the application of each of the reduction methods which have been described in Chapter 3. The methods will be applied in order and in each case the test system is reduced to a second order model. The response of each to a step input is shown and comments are made on the closeness of fit to the actual response. Further comments on these responses and the various reduction methods are given in the following chapter.

4.1 The Application of Method 1 to the Test System

The application of Method 1 involves a repeated sequence of operations which are relatively easy to program for the computer. A computer program which applies this method to a general high order system was therefore prepared and is shown in Appendix C. The following reduced transfer function resulted from the continued fraction expansion of the test system transfer function.

$$\frac{C(s)}{R(s)} = \frac{0.1299s + 0.01105}{s^2 + 1.14644s + 0.09941}$$

In state space representation, the reduced system equations are

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -.09941 & -1.14644 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = \begin{bmatrix} .01105 & .1299 \end{bmatrix} \underline{x}$$

A step input was applied to the system using the computer program in Appendix B.2, and the response was evaluated. This response is shown for a twenty second time interval in Figure 4.1.

It is apparent in this graph that the response has a finite initial slope which is not present in the actual response. There is also a considerably shorter rise time, and the peak response occurs sooner than in the actual response. From approximately 3.5 seconds to steady state, however, the two response curves fit with increasing closeness.

4.2 The Application of Method 2 to the Test System

In order to obtain a second order model using this method, the two most dominant eigenvalues of the test system must first be selected. In this case, however, the second most dominant eigenvalue is part of a complex pair, and there are not, in fact, two simple dominant eigenvalues in the system. However, the most dominant root is located

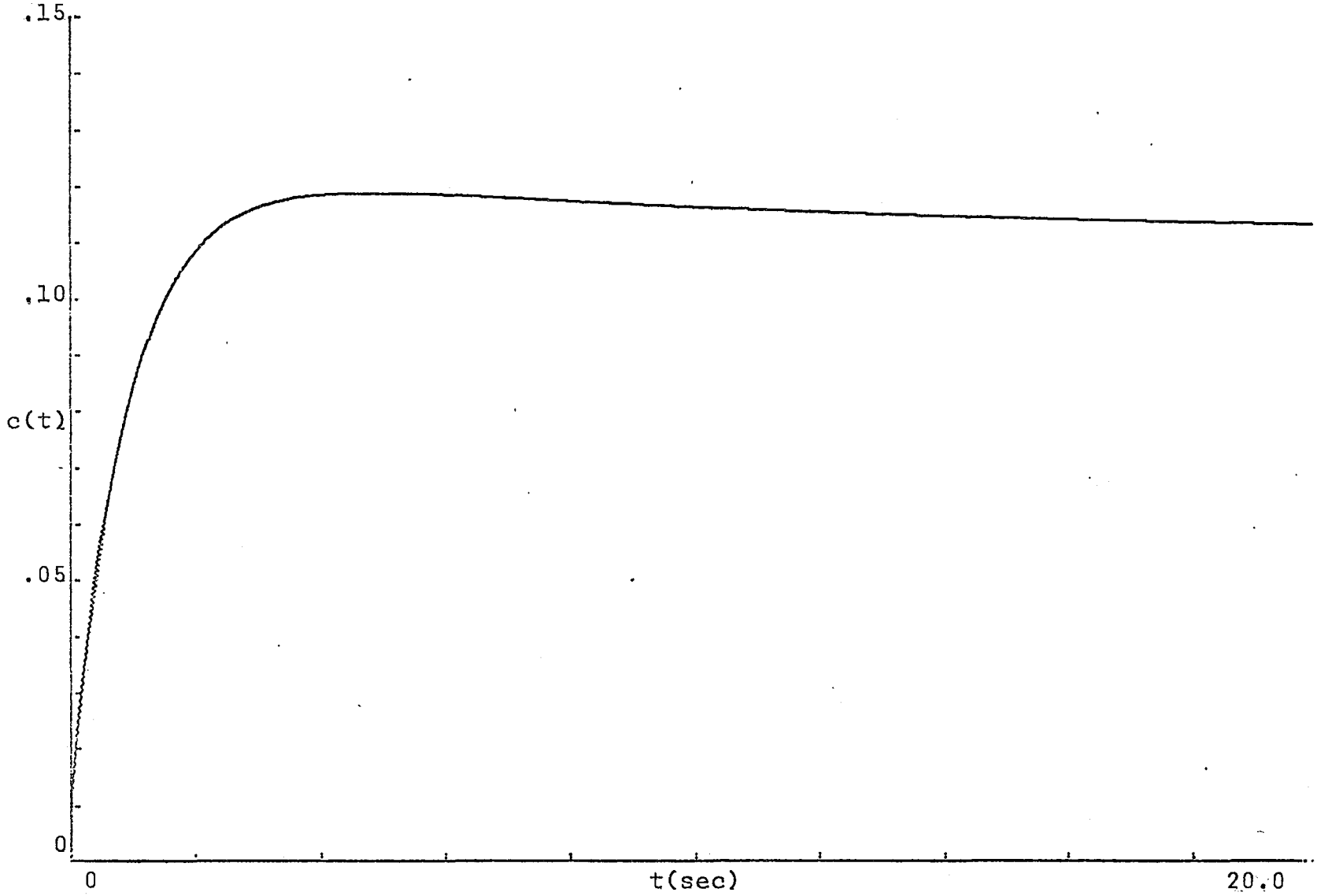


Figure 4.1 Step Response of the Model from Method 1.

very near to the zero of the system, so that the pair of complex poles located nearest to the jω-axis could reasonably be selected to represent the system. Now, two state variables from the test system must also be selected to be retained in the model. The author requires that the state variables be chosen so that the determinant of the reduced modal matrix be as large as possible. For this reason, x_1 and x_7 from the test system were retained in the reduced system. Various operations were performed in order to proceed from this point to the reduced system. The numerical results of these operations are shown in Appendix D.1 which gives a step by step progression from the reduced system matrices back to the test system matrices.

The application of this method is rather involved and some problems were encountered when the method was applied to the system. It was decided that a single computer program to perform the reduction process would be impractical because of the size and required running time. However, these operations, when performed individually, involved round-off errors, occasional human errors, and were generally awkward and tedious. One problem was the invasion of a complex matrix for which the determinants of ersion and imaginary parts were zero. This was overcome by a useful technique outlined in Appendix D.2.

From Appendix D.1 then, the second order model of the test system is given as

$$\dot{\underline{x}} = \begin{bmatrix} -3.907092 & 1.29\text{x}10^{-6} \\ -3463089.0 & -.1416739 \end{bmatrix} \underline{x} + \begin{bmatrix} -.050525 \\ 437819.7 \end{bmatrix} u$$

$$c = x_1$$

The system was subjected to a step input and the response for a twenty second interval is graphed in Figure 4.2.

The reduced system response closely approximates the actual response from approximately 16 seconds to steady state. Although the rise time is reasonably close, the maximum overshoot, settling time, and initial slope leave much to be desired.

4.3 The Application of Method 3 to the Test System

The application of this method is relatively simple using a digital computer to accomplish several matrix operations which are easily combined in one program. The means by which this method was applied to the test system is outlined in Appendix E.1 and the computer program used is given in Appendix E.2.

The author suggests that the total time interval over which response values are calculated should considerably exceed the longest time constant in the unreduced system. The longest time constant in the test system is

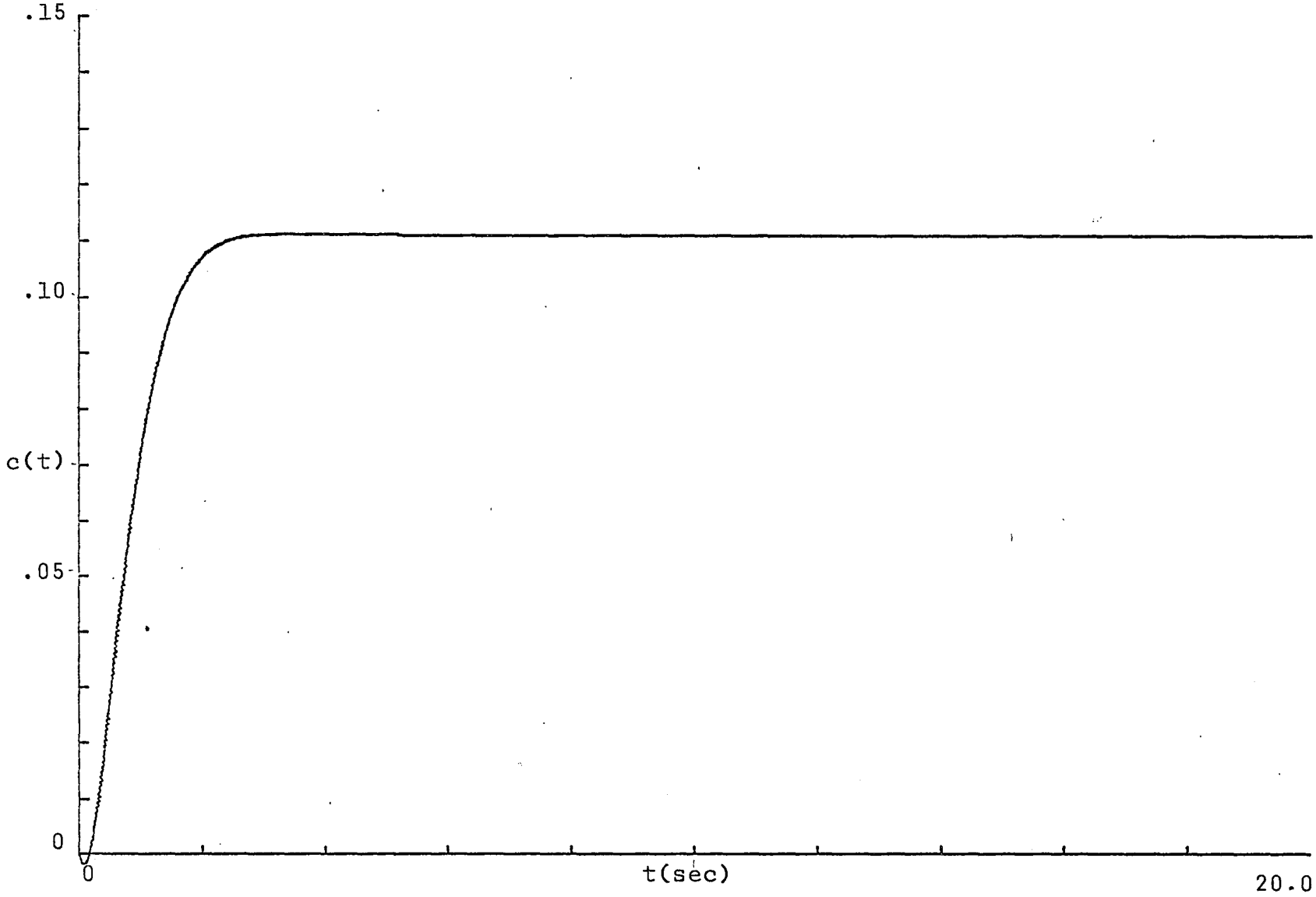


Figure 4.2 Step Response of the Model from Method 2

approximately 10 seconds, and response samples were taken every 0.04 seconds over a total interval of 20 seconds to comply with the above.

The resulting second order model, as evaluated in Appendix E.1 is

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -2.687909 & -1.902574 \end{bmatrix} \underline{x} + \begin{bmatrix} 0.0 \\ 0.309611 \end{bmatrix} u$$

$$c = x_1$$

The response of this model to a step input is shown in Figure 4.3. The initial part of the transient response closely approximates that of the actual response and the overall comparison for a twenty second interval is quite accurate. Indeed, the error between these curves at the sampling instants (for the time involved) is minimized, although the actual area between the curves may not be. A sizeable steady-state error is evident in the response, however, as well as an undesirably high overshoot.

Three different models have been obtained to represent the test system by appling the various reduction techniques. The next chapter will compare these methods and their usefulness.

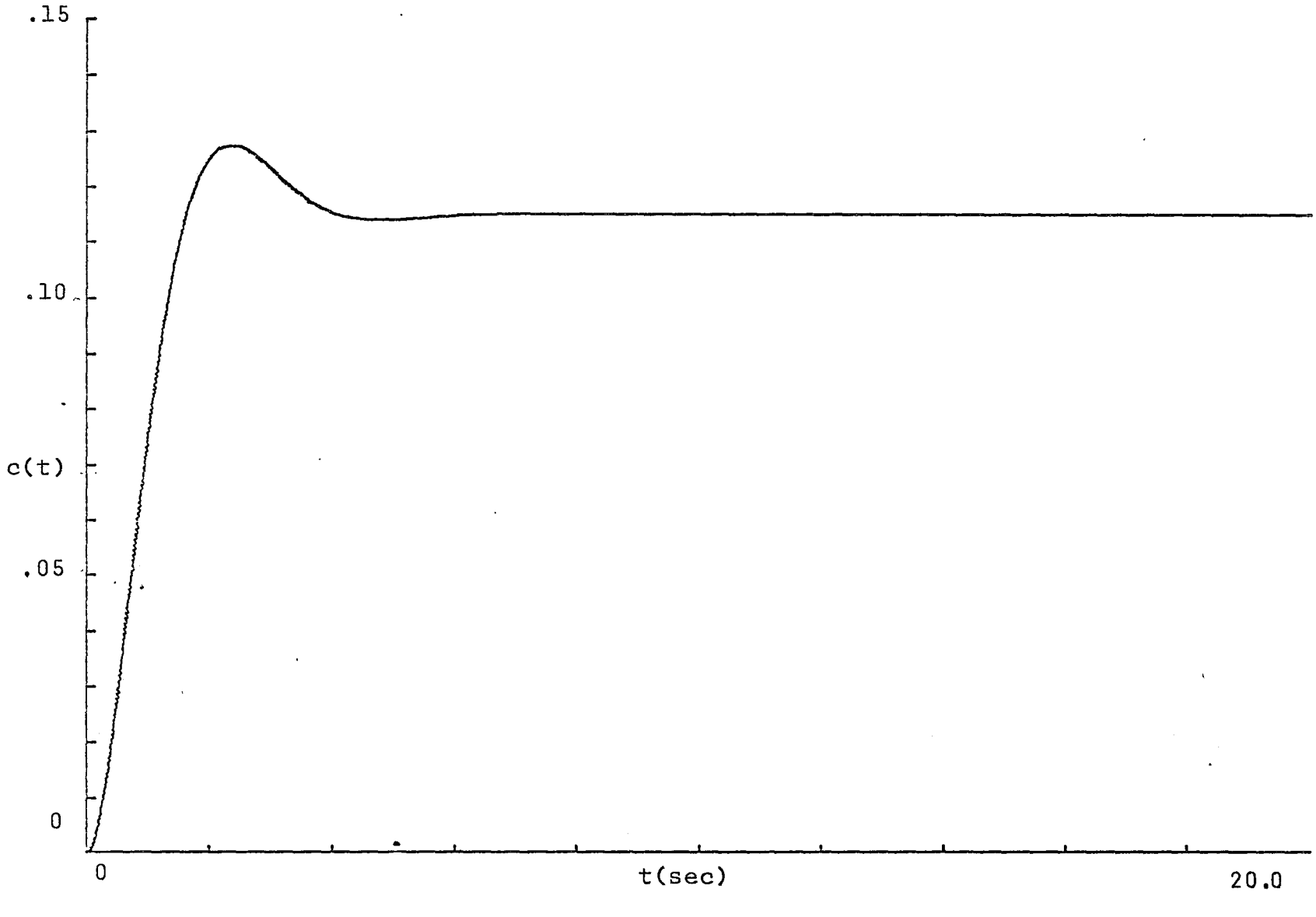


Figure 4.3 Step Response of the Model from Method 3

CHAPTER 5

COMPARISON OF REDUCTION METHODS

5.0 General

The various responses of the reduced systems are compared with the actual response in this chapter. An appraisal is made which weighs the advantages and disadvantages of the methods based on the accuracy of the reduced systems, their relative ease of application and their limitations.

The different responses of the reduced systems varied considerably in their accuracy, and in the appraisal some emphasis has been put on the transient response because of its importance. Indeed, the desire to reduce a high order system in order to evaluate an approximate optimal control for a continuous process suggests that the system will be changing frequently. In practice, high speed systems such as airplanes, paper mills and steel mills require controllers which will spend much time in transient states, making corrections quickly, but with strict limits on overshoot and oscillations. It is therefore necessary that the transient response of the reduced model duplicates that of the actual system as closely as possible, maintaining certain other criteria.

5.1 Comparison of Accuracy

Figures 5.1 to 5.9 show the step response of the different reduced systems compared to that of the actual system, over various time intervals. Figures 5.1 to 5.3 show the accuracy of each reducing method during the transient part of the response while Figures 5.4 to 5.6 show the actual and reduced system responses over a 100 second interval.

The transient response of the system reduced by Method 3 most closely fits the actual response, although the reduced system gives a considerably higher overshoot. There is also an undershoot in the response which is not in the actual response, and a steady state error is evident.

The system reduced by Method 1 has a step response which very closely fits the actual from approximately 4.5 seconds to steady state. The initial slope of the response, which is actually zero, has not been maintained by the reduced model, and the maximum overshoot as well as the time of maximum overshoot are considerably altered.

The response of the system reduced by Method 2, while agreeing closely at steady state, is a relatively poor fit through the transient time. The initial slope of the model's response is negative, resulting in a period of negative response which, although short in duration and

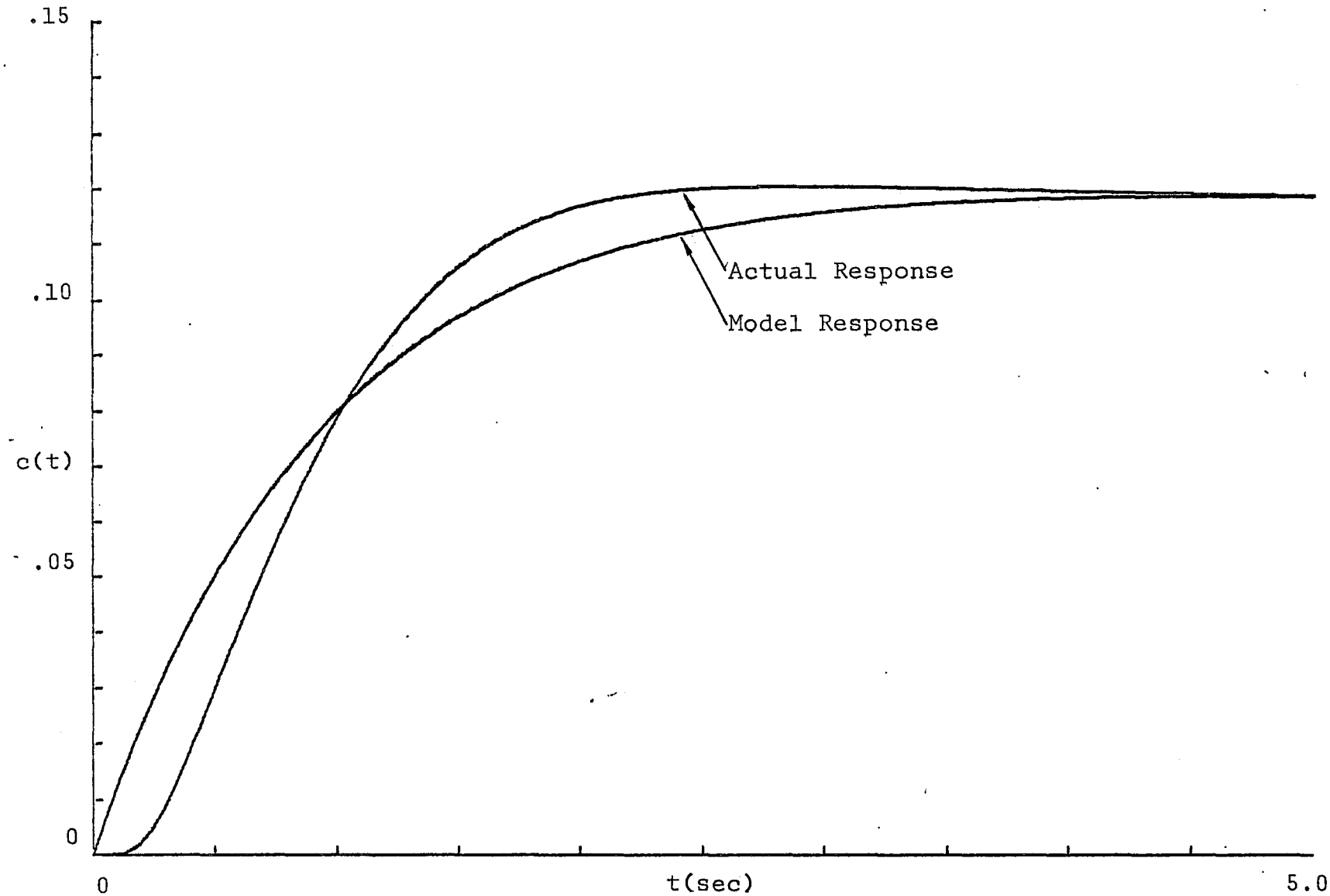


Figure 5.1 Transient Response Accuracy using Method 1.

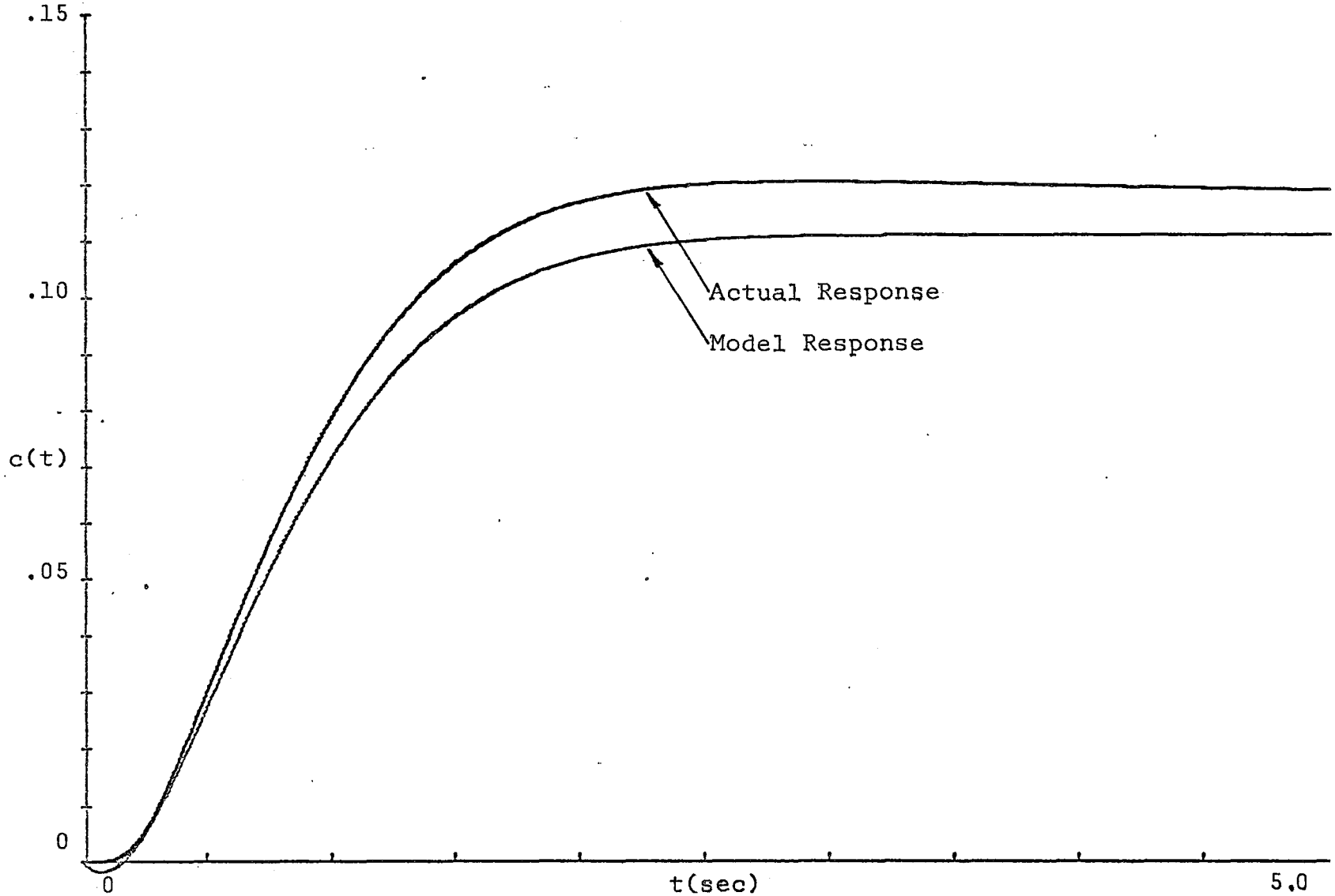


Figure 5.2 Transient Response Accuracy using Method 2.

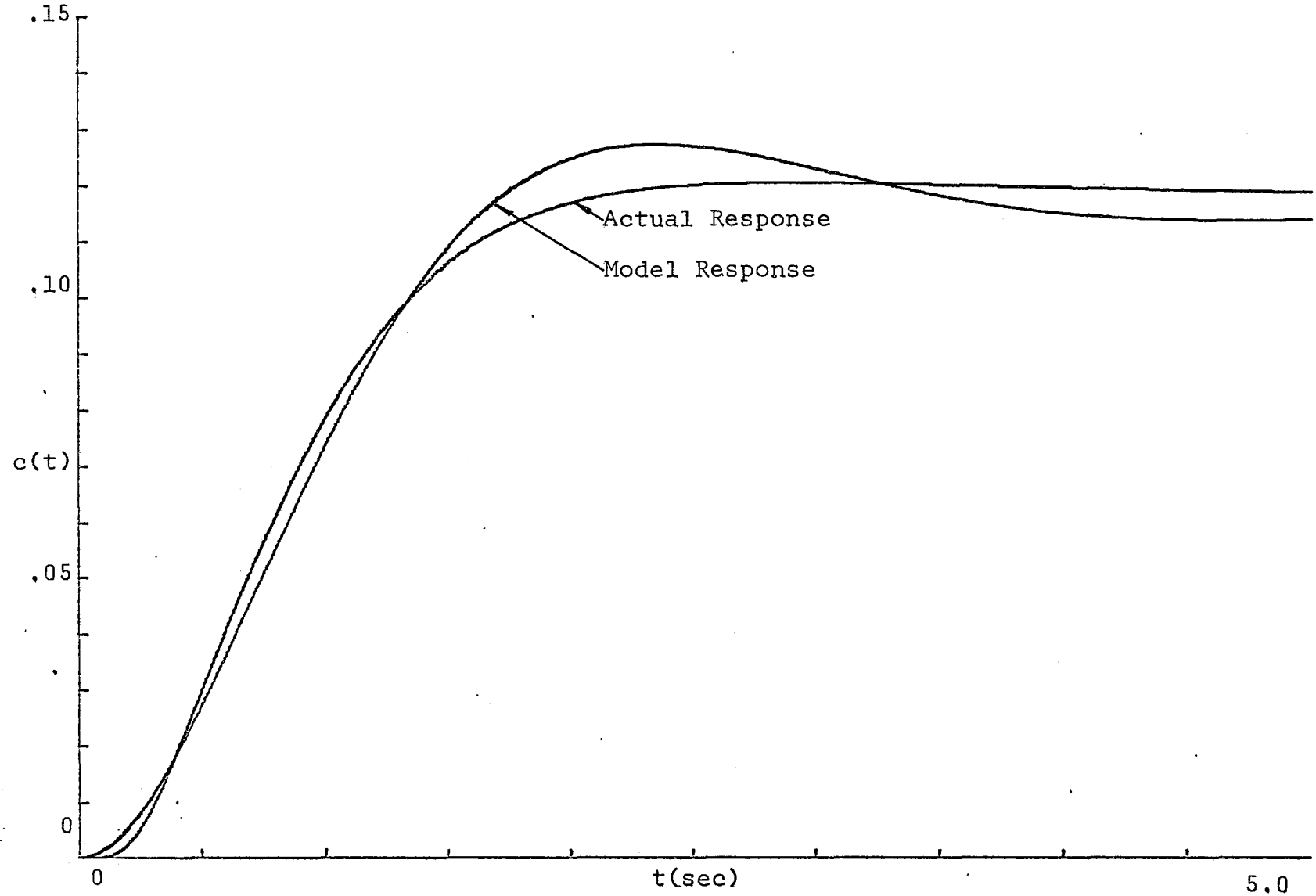


Figure 5.3 Transient Response Accuracy using Method 3.

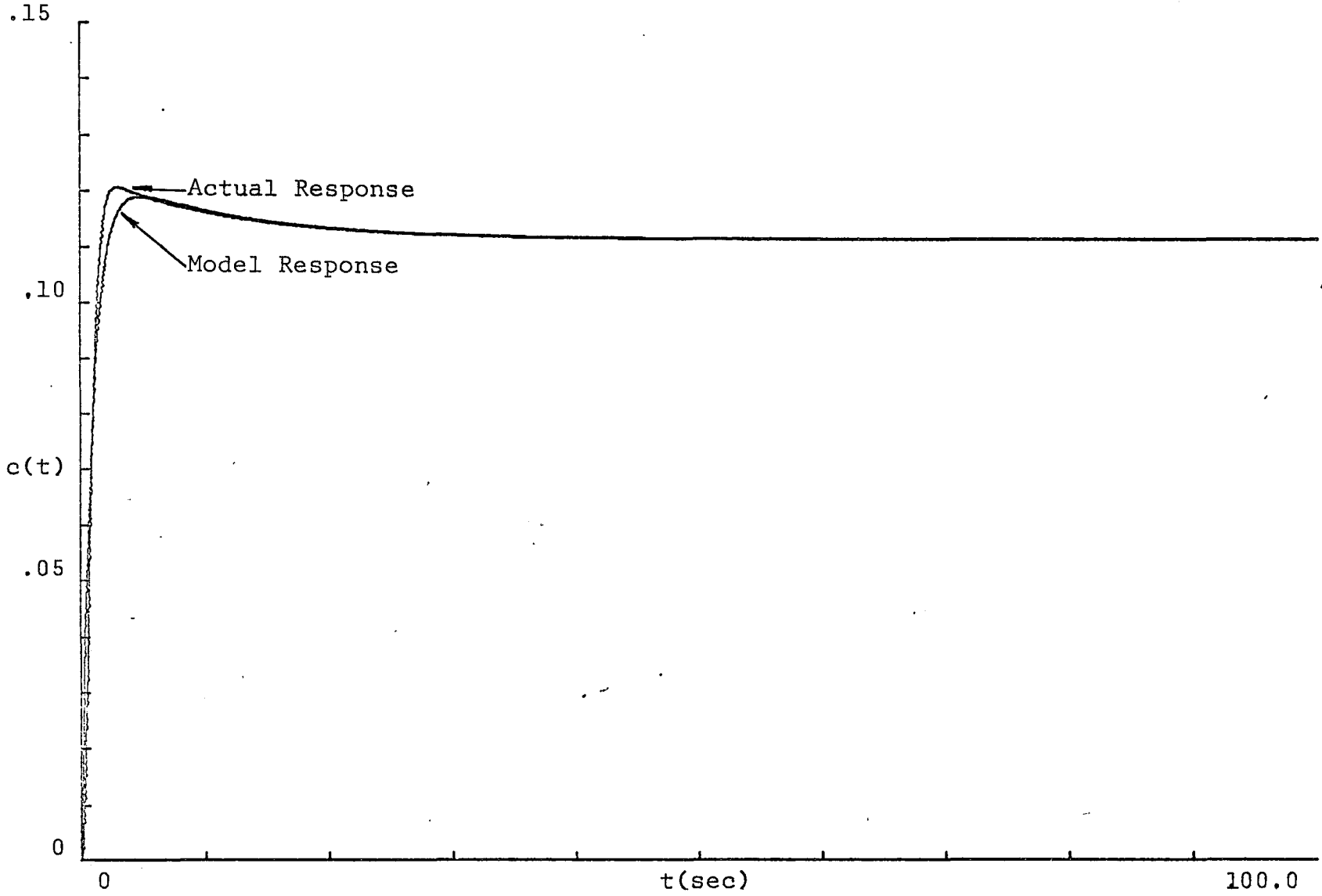


Figure 5.4 Overall Response Accuracy using Method 1.

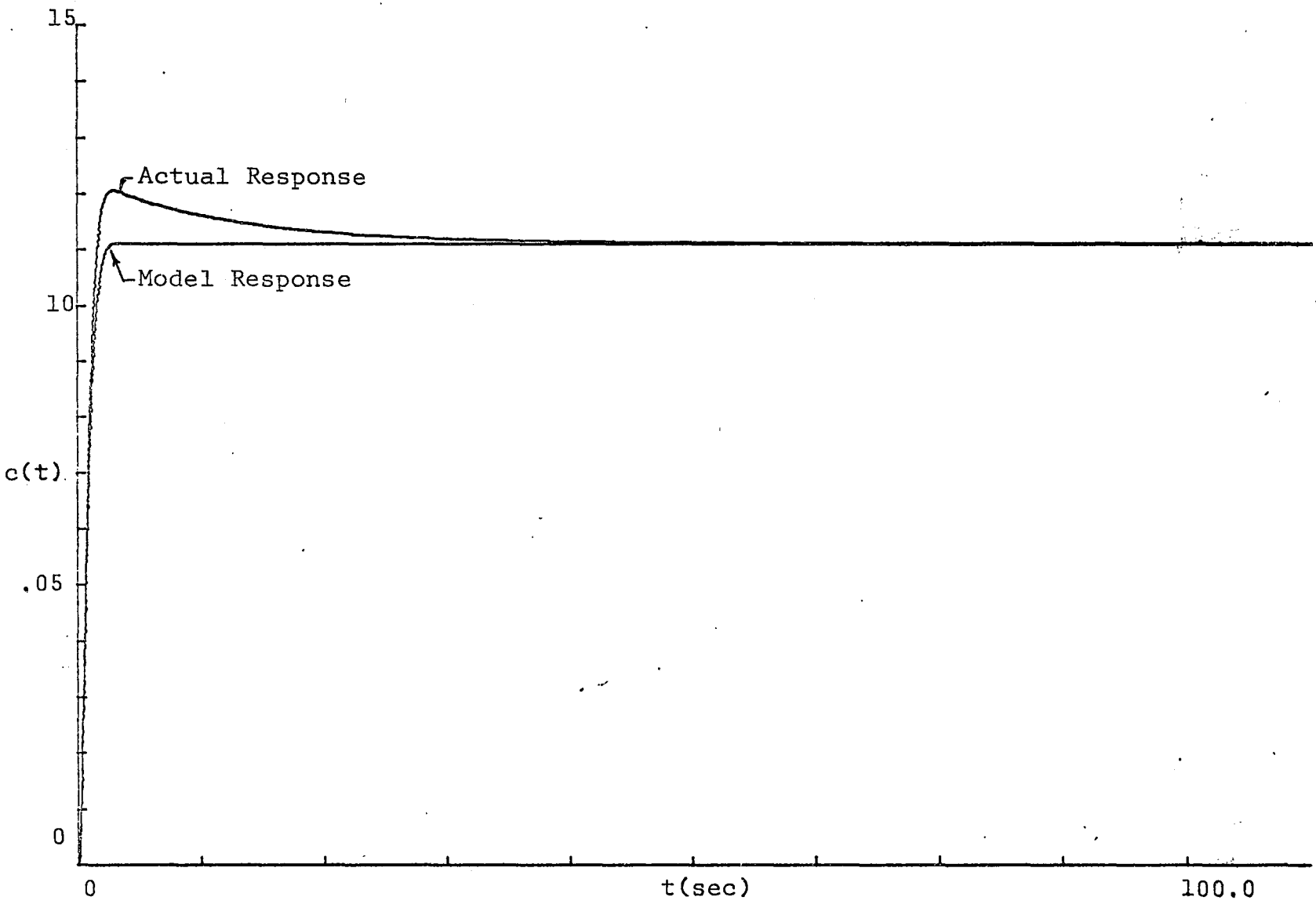


Figure 5.5 Overall Response Accuracy using Method 2

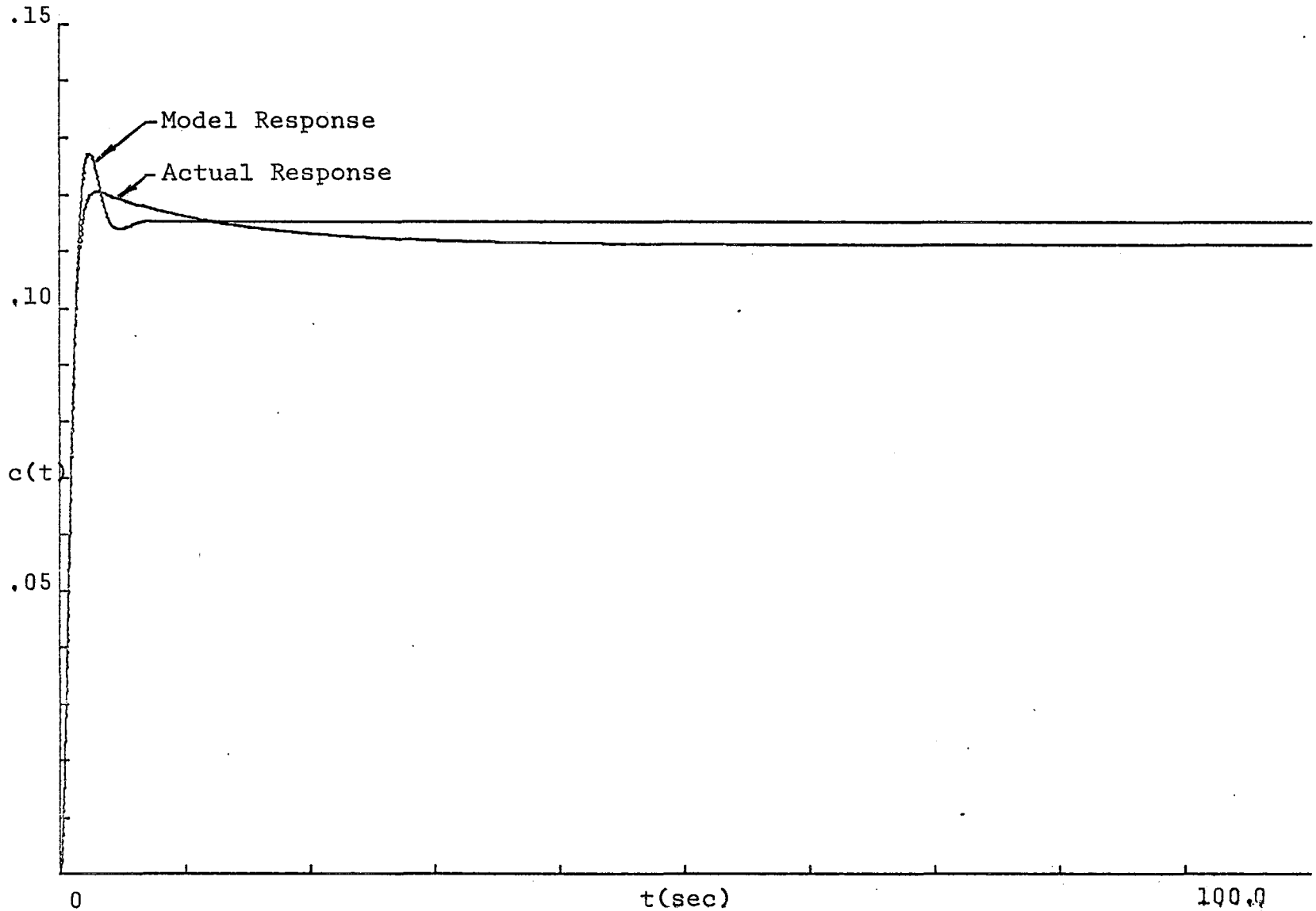


Figure 5.6 Overall Response Accuracy using Method 3.

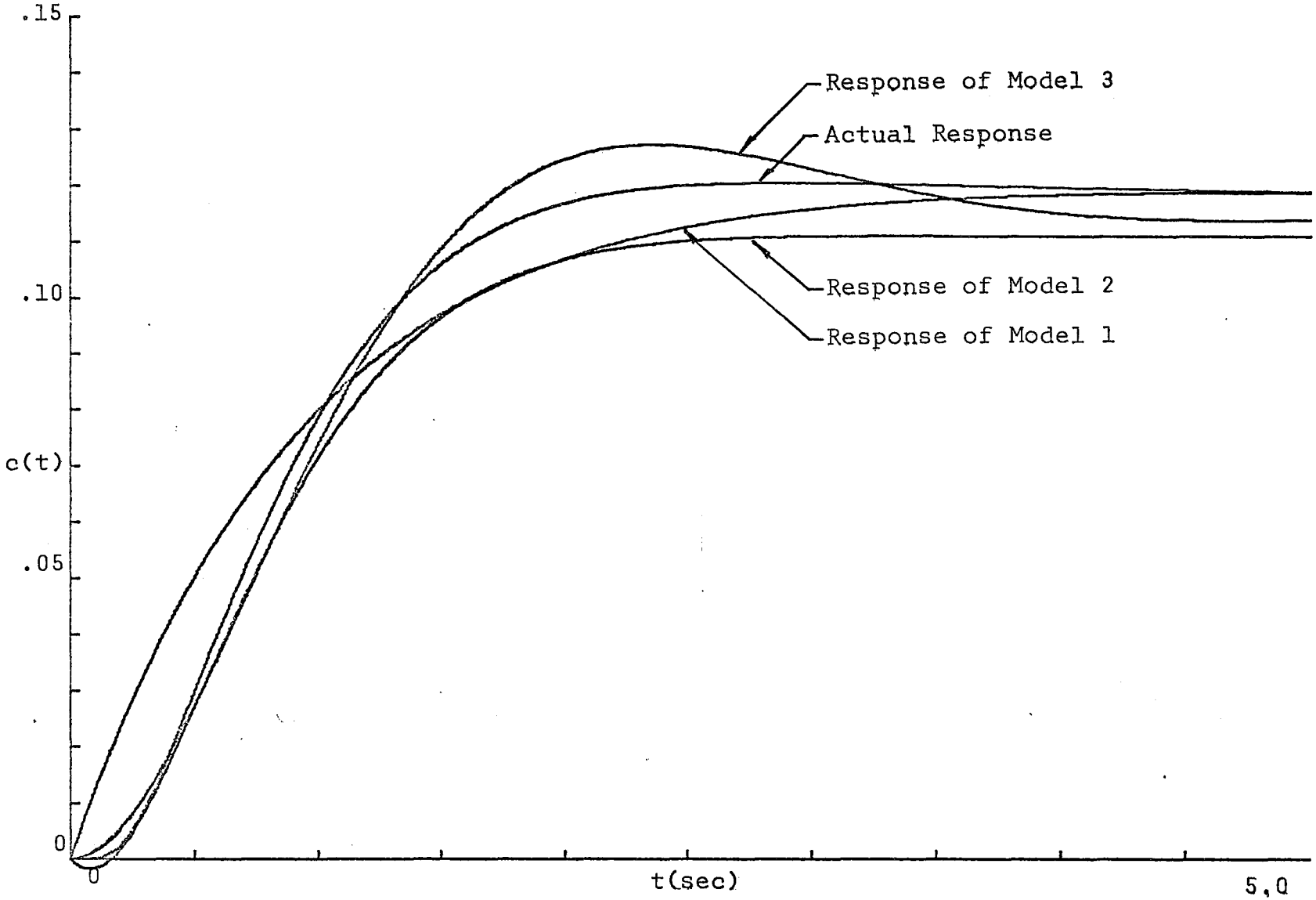


Figure 5.7 Relative Response Accuracies

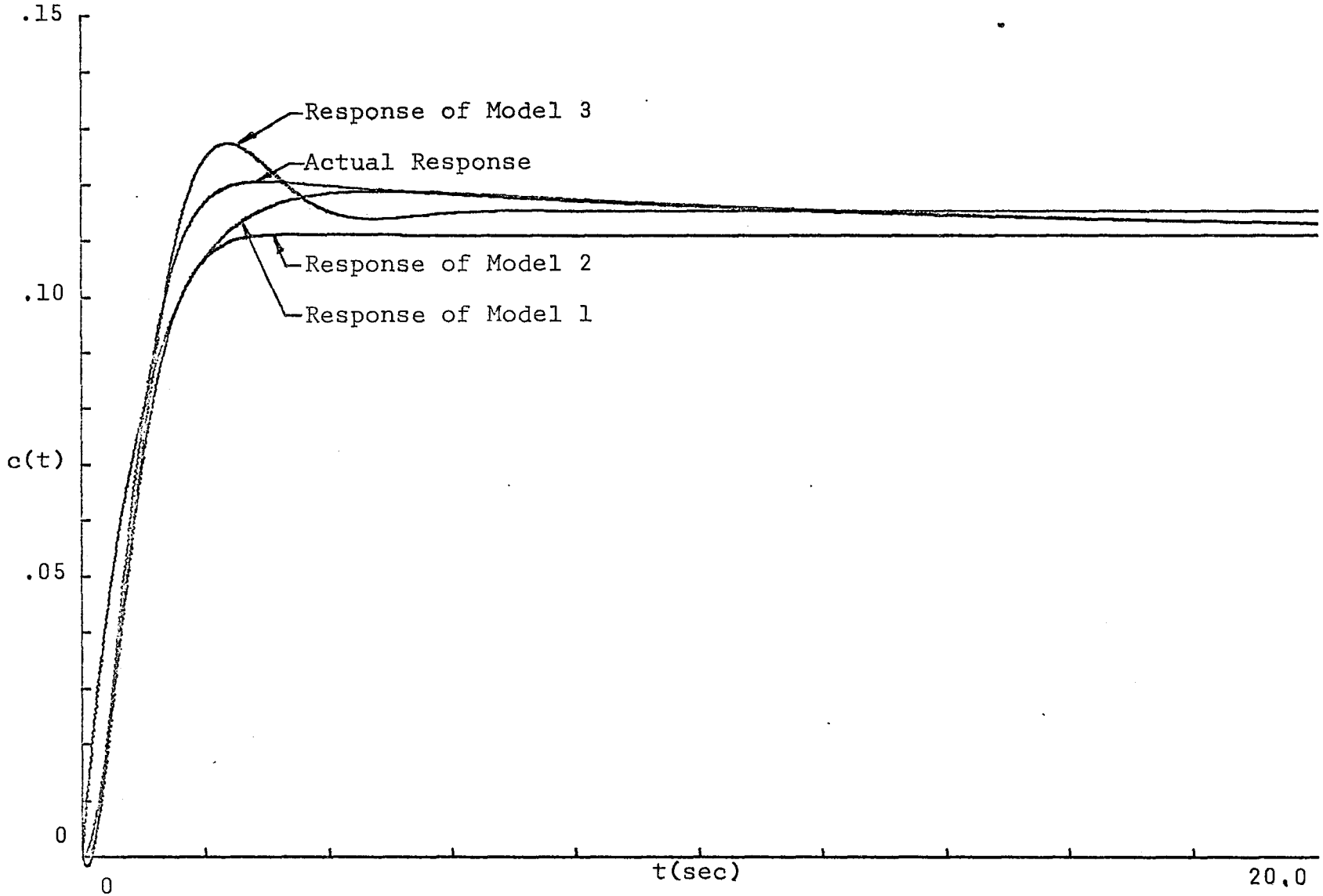


Figure 5.8 Relative Response Accuracies

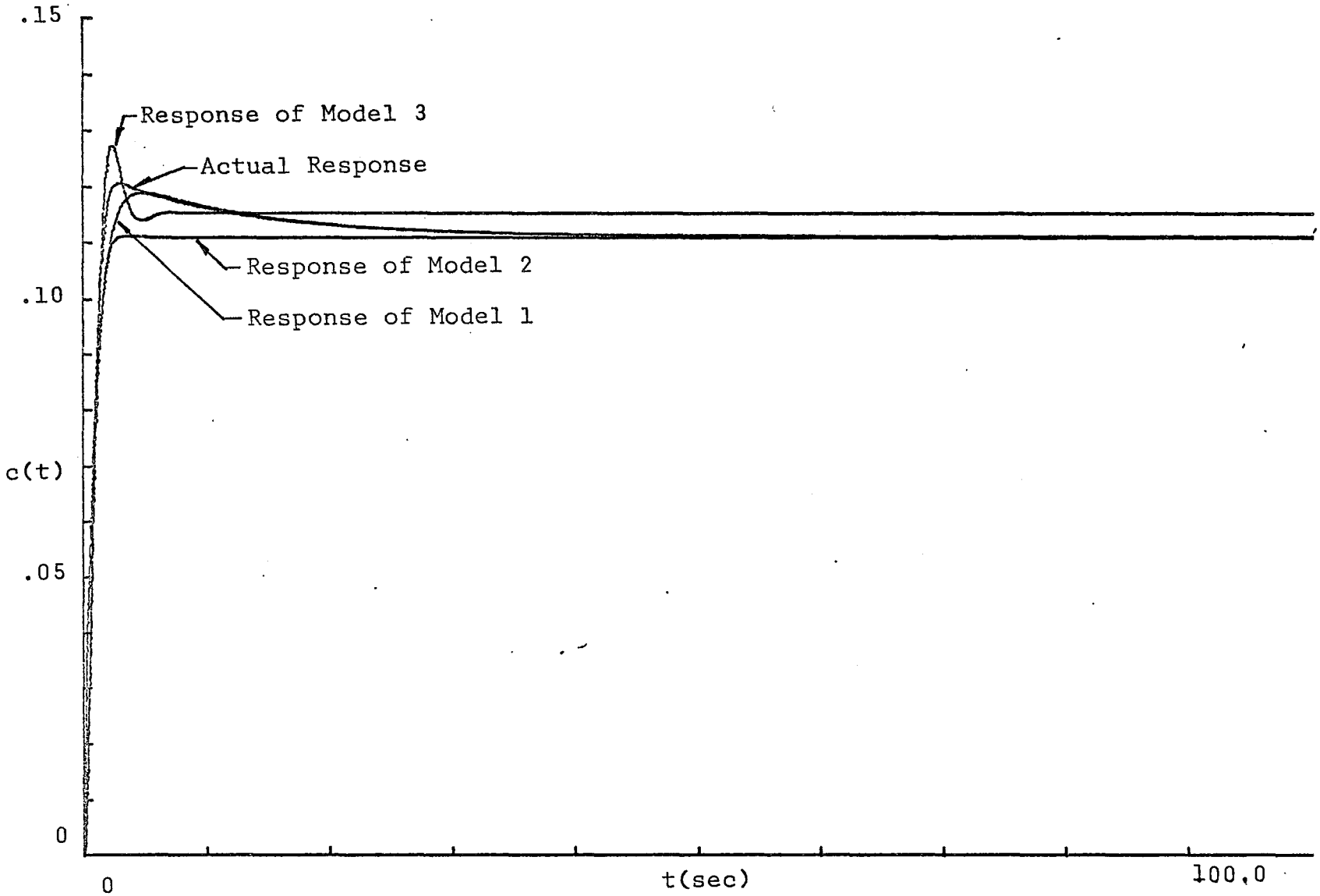


Figure 5.9 Relative Response Accuracies

small in size, is undesirable.

Figures 5.7 to 5.9 provide a composite picture of the responses for three increasing time intervals. It is interesting to note that some compromise between the systems reduced by Methods 1 and 3 could conceivably result in a system which would have a response closer to that of the actual system during the transient period with little or no error at steady state. This is very desirable, since many control systems operate mainly in the transient state.

5.2 Comparison with a View to Applications

The easiest of the three methods to both apply and conceive is certainly Method 1. The concept of the equivalent transfer function, consisting of nested paths in both novel and lucid. The continued partial fraction is easily obtained using a digital computer and with minimal effort this method can even be applied by hand computations. Once having obtained the continued fraction, the low order system can be represented by either its transfer function or state space equations quite readily. The entire operation can be accomplished by a single generalized computer program requiring only a few seconds running time. Further, Method 1 requires no knowledge of the pole locations and because of its nature, does not rely on the existence of a few dominant poles in the system to obtain satisfactory

results. The method does not, however, perform any error minimization which is commonly desirable.

The chief advantage of Method 2 is that it can be used to reduce systems with multiple inputs. This is, in itself, an attractive feature, but the application of the method is a relatively involved process.

Firstly, the eigenvalues and eigenvectors of the system are required, as well as the modal matrix and its inverse. Next, the state variables to be retained in the reduced system must be decided upon. They are required to be as different physically from each other as possible, but the physical significance of state variables may be difficult to interpret. One may, however, overcome this difficulty by deciding to retain those state variables which are measurable. Many matrix operations are then required before the reduced system is obtained. The method can be implemented by a single computer program but the program would be long and require considerable running time. In addition, the steady state value of the reduced system response is not guaranteed to be correct.

Method 3 has several important advantages. The error between the actual and reduced system responses is minimized in a least squares sense. Further, these errors can be weighted in certain regions of the response where it is desired to have a 'closer fit'.

The A matrix from the state space equations of the original system can be altered (by interchanging rows and columns) so that specific state variables (measurable ones) will be the ones retained in the reduced system. In addition, no knowledge of the poles is required and the method is particularly adaptable to computer application.

The steady state value of the reduced system response is not guaranteed to be exact, however, and it may be necessary to use a great many samples of the response from the original system in order to arrive at a satisfactory low order model.

5.3 Comparison of Limitations

The only serious limitation of Method 1 is the requirement that there be only one input to the system. This is unfortunate since many of the advanced control systems, for which reduced models are desirable, contain multiple inputs. In such cases, one may be specified a transfer function matrix, the respective elements of which may be reduced using this method.

Method 2 retains the dominant eigenvalues of the high order system in the low order model, so that in a sense, this method is limited to systems which do, in fact, have this characteristic of dominance. A high order system with poles located close together could be better approximated by a low order system if the poles of the reduced

system were shifted to compensate for the change in system order.

Method 3 has few actual limitations; the author mentions an unsuccessful first attempt to reduce a system satisfactorily,[8] but this was overcome by weighting the actual response in a certain region in order to improve the reduced system response. The method might be impractical to use on a system with poles very near to the $j\omega$ axis, since the time interval over which the samples of the response are required becomes very large. It might also be undesirable to have a reduced model which does not guarantee a close fit to the step response at steady state.

Method 3 minimizes the difference between the two curves at the various times of measurement. Since the transient part of the response has a relatively high slope, small perpendicular distances between the curves over this portion of the curve will have large response (vertical) differences, so that this region is inherently emphasized by the reduction method. This is a very desirable feature, but it could be somewhat advantageous to have some control over the range in the response curve where errors are allowed to occur.

All the methods suffer from the basic limitation that the exact system transfer function or vector differential equation must be known beforehand. A more realistic

approach would be the development of the simplified system model directly from the step-response of the system, which may often be obtained experimentally in a very simple manner.

CHAPTER 6

A NEW APPROACH TO THE REDUCTION PROBLEM

6.0 Introduction

One problem frequently encountered with control systems is that of identification. This is a very real problem in that it is seldom possible to formulate exact transfer functions or vector differential equations for real systems. Systems are therefore 'estimated', and this is commonly done on the basis of the response of the system to a step input. As mentioned at the close of Chapter 4, it would be desirable, then, to fashion the reduced model of a system directly from the step response of the actual system, rather than building the reduced system from an estimated transfer function. A new concept in system reduction which bypasses this estimation problem will be presented after the following discussion of general control system characteristics. The actual input to which an operating control system will be subjected to is usually unknown. Some standard test signals have therefore been developed, and among these, the step input is the most widely used by control engineers. The step input is a convenient signal to describe mathematically; it is easily applied to a system, and the response to this input gives

a great deal of information which has come to be regarded as the control characteristics of a system. It is commonplace then, for designers to specify certain of the step response characteristics for a given system. The characteristics most often specified include initial slope, rise time, maximum overshoot and time of maximum overshoot (which are all transient characteristics), as well as settling time and steady state response.

A close examination of Figure 5.7 shows that the response of each reduced system would satisfy only a few (if any) of the actual response characteristics, but none of the reduction methods allow the selection of specific response characteristics.

It appears that some compromise between the systems reduced by Method 1 and Method 3 could result in a more accurate overall response, and it is with this thought that a new approach to the problem has been taken. Rather than manipulating the system parameters to achieve reduction, a second order system is proposed which will have a step response consistent with some desired criteria. That is, given the step response of a high order system, a second order model may be constructed whose response will have certain of the desired characteristics of the actual response.

6.1 Second Order Model Variations

Second order models may be of four different types:

1) Two Real Poles And a Zero
2) A Pair of Complex Poles and a Zero
3) Two Real Poles without a Zero, and
4) A Pair of Complex Poles without a Zero.

In the first two cases, the step response will have a non-zero initial slope, whereas in the last two cases, the initial slope of the response will be zero. With these four models, both underdamped and overdamped systems can be approximated.

Each of the four systems are general within their range and thus have a certain number of variables, or degrees of freedom, which are accounted for by the number of response characteristics required to be met. These models, their equations and ranges of application are discussed next.

6.2 A Second Order Model with Two Real Poles

This model is used when the initial slope of the response is zero and there is no overshoot. The transfer function for the system is of the form

$$\frac{C(s)}{R(s)} = \frac{K}{(s + \alpha)(s + \beta)}$$

A typical response for the system is shown below in Figure 6.1.

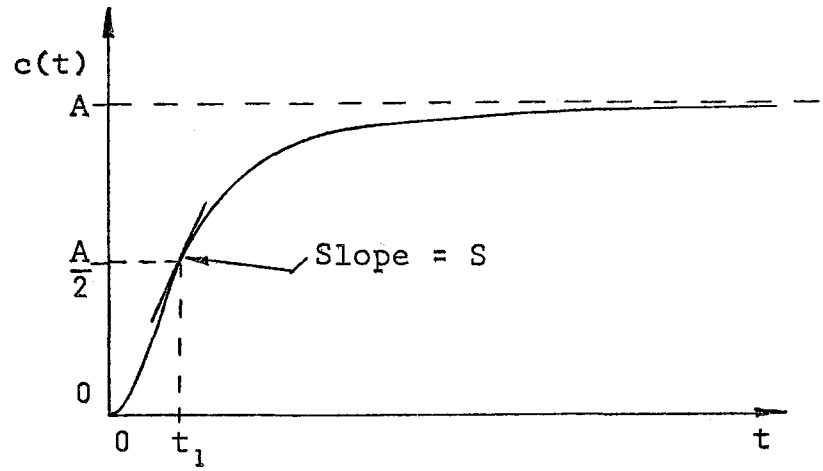


Figure 6.1: Step Response of a Second Order System with Two Red Poles.

The expressions for the step response and its slope are,

$$c(t) = \frac{K}{\alpha\beta} - \frac{K}{\alpha(\beta-\alpha)}.e^{-\alpha t} + \frac{K}{\beta(\beta-\alpha)}.e^{-\beta t}$$

$$\frac{dc}{dt} = \frac{K}{\beta-\alpha}.(e^{-\alpha t} - e^{-\beta t}) \qquad |\beta| > |\alpha|$$

These equations contain three variables which will allow the specification of three response characteristics. The steady state value (A), the time (t_1) at which $c(t) = \frac{A}{2}$, and the slope (s) at t_1 may well be selected and the following relationships are then established.

$$A = \frac{K}{\alpha\beta} \qquad \text{.....(6.2-2)}$$

$$\frac{A}{2} = \frac{K}{\alpha\beta} - \frac{K}{\alpha\beta(\beta-\alpha)}.(\beta e^{-\alpha t_1} - \alpha e^{-\beta t_1})$$

$$S = \frac{K}{\beta-\alpha}.(e^{-\alpha t_1} - e^{-\beta t_1})$$

By substitution, these equations are found:

$$\beta = \frac{1}{t_1} \ln \frac{2A\alpha}{A\alpha-2S} \qquad \text{.....(6.2-3)}$$

$$(\frac{1}{2} - e^{-\alpha t_1}).\frac{1}{t_1}.\ln \frac{2A\alpha}{A\alpha-2S} = \frac{S}{A} \qquad \text{.....(6.2-4)}$$

Equation (6.2-4) is used to find α numerically. β is then obtained from Eqn. (6.2-3), and K from Eqn. (6.2-2).

6.3 A Second Order Model with Complex Poles

This model is used when the initial slope of the response is zero and there is an overshoot. The transfer function for the system is of the form

$$\frac{C(s)}{R(s)} = \frac{K}{(s+\alpha)^2 + \beta^2} \qquad \text{.....(6.3-1)}$$

A typical step response for the system is shown in Figure 6.2

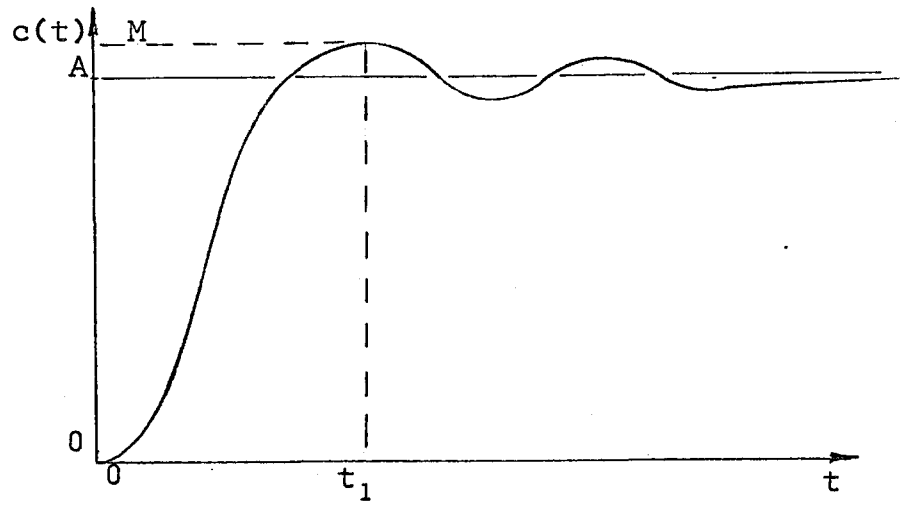


Figure 6.2 Step Response of a Second Order System with Complex Poles

The expressions for the step response and its slope are,

$$c(t) = \frac{K}{\alpha^2+\beta^2} - \frac{K}{\beta\sqrt{\alpha^2+\beta^2}}.e^{-\alpha t}.\mathrm{Sin}(\beta t + \tan^{-1}\beta)$$

$$\frac{dc}{dt} = \frac{K}{\beta\sqrt{\alpha^2+\beta^2}}.e^{-\alpha t}[\alpha\mathrm{Sin}(\beta t + \tan^{-1}\frac{\beta}{\alpha}) - \beta\mathrm{Cos}(\beta t + \tan^{-1}\frac{\beta}{\alpha})]$$

Again, these equations contain three variables and the required response characteristics here might well be the steady state value (A), the maximum response value (M), and the time of maximum response (t_1).

At the peak, $\tan(\beta t_1 + \tan^{-1}\frac{\beta}{\alpha}) = \frac{\beta}{\alpha}$

Then $t_1 = \frac{\Pi}{\beta}$ and $\beta = \frac{\Pi}{t_1}$ .....(6.3-2)

Also, $A = \frac{K}{\alpha^2 + \beta^2}$ .....(6.3-3)

$$\text{and} \quad \alpha = \frac{1}{t_1} \ln \frac{A}{M-A} \qquad \text{.....(6.3-4)}$$

By determining β from Eqn. (6.3-2) and α from Eqn. (6.3-4), K can be obtained from Eqn. (6.3-3).

6.4 A Second Order Model containing One Zero and Two Real Poles.

This type of model is used when the initial slope of the step response is not zero. The model may also accommodate a small overshoot in the response, but no undershoot.

The system will have a transfer function of the form

$$T(s) = \frac{K(s+\delta)}{(s+\alpha)(s+\beta)} \qquad \text{.....(6.4-1)}$$

Now, the expressions for the response and slope of response to a step input become

$$c(t) = \frac{K\delta}{\alpha\beta} + \frac{K(\alpha-\delta)}{\alpha(\beta-\alpha)} . e^{-\alpha t} - \frac{K(\beta-\delta)}{\beta(\beta-\alpha)} . e^{-\beta t}$$

$$\frac{dc}{dt} = -\frac{K(\alpha-\delta)}{(\beta-\alpha)} . e^{-\alpha t} + \frac{K(\beta-\delta)}{(\beta-\alpha)} . e^{-\beta t}$$

$$|\beta| > |\alpha|$$

The characteristics shown in Figure 6.2 may be used here, along with the specified initial slope of the response to give the following relationships.

At peak time (t_1),

$$(\alpha-\delta)e^{-\alpha t_1} = (\beta-\delta)e^{-\beta t_1}$$

$$M = \frac{K\delta}{\alpha\beta} + \frac{K}{\beta-\alpha}\left[\frac{(\alpha-\delta)}{\alpha}.e^{-\alpha t_1} - \frac{(\beta-\delta)}{\beta}.e^{-\beta t_1}\right]$$

At steady state,

$$A = \frac{K\delta}{\alpha\beta} \qquad .....(6.4\text{-}2)$$

at $t = 0$,

$$\left.\frac{dc}{dt}\right|_{t=0} = K \qquad .....(6.4\text{-}3)$$

Then, by substitutions, we have

$$\beta = \frac{1}{t_1}\ln\left[\frac{K-A\alpha}{\alpha(M-A)}\right] \qquad .....(6.4\text{-}4)$$

and,

$$\left[A + (M-A).e^{\alpha t_1}\right]\ln\left[\frac{K-A\alpha}{\alpha(M-A)}\right] = Kt_1 \qquad .....(6.4\text{-}5)$$

In Eqn. (6.4-3), K is evaluated as the initial slope of the response. A numerical value for α is determined from Eqn. (6.4-5); then B is obtained from Eqn. (6.4-4) and δ from Eqn. (6.4-2).

6.5 A Second Order Model containing One Zero and Complex Poles.

This system is also used when the initial slope of

the response is not zero, but the system also covers the range of responses which have large overshoots as well as some undershoots.

The transfer function for the system takes on the form

$$T(s) = \frac{K(s+\delta)}{(s+\alpha)^2 + \beta^2} \qquad .....(6.5\text{-}1)$$

The step response and its slope are expressed as

$$c(t) = \frac{K\delta}{\alpha^2+\beta^2} - \frac{K\sqrt{\beta^2+(\delta-\alpha)^2}}{\beta\sqrt{\alpha^2+\beta^2}}.e^{-\alpha t}.\mathrm{Sin}(\beta t+\phi)$$

$$\frac{dc}{dt} = \frac{K\sqrt{\beta^2 + (\delta-\alpha)^2}}{\beta\sqrt{\alpha^2 + \beta^2}}.e^{-\alpha t}\left[-\alpha\mathrm{Sin}(\beta t+\phi) + \beta\mathrm{Cos}(\beta t+\phi)\right]$$

where

$$\phi = \tan^{-1}\frac{\beta}{\alpha} - \tan^{-1}\frac{\beta}{\alpha-\delta}$$

From the characteristics shown in Figure 6.2 and the specified initial slope of the response, the following expressions are derived.

At peak time (t_1)

$$\tan(\beta t_1+\phi) = \frac{\beta}{\alpha}$$

Then,

$$\beta t_p = \tan^{-1}\frac{\beta}{\alpha-\delta} \qquad .....(6.5\text{-}1)$$

At t = 0,

$$\frac{dc}{dt} = K \qquad \text{.....(6.5-2)}$$

At steady state,

$$A = \frac{K\delta}{\alpha^2 + \beta^2} \qquad \text{.....(6.5-3)}$$

and,

$$c(t_1) = \frac{K\delta}{\alpha^2 + \beta^2} + \frac{K\sqrt{\beta^2 + (\delta-\alpha)^2}}{\alpha^2 + \beta^2}.e^{-\alpha t_1} \qquad \text{.....(6.5-4)}$$

Equations (6.5-2) and (6.5-3) may be substituted in the other two equations so that only two variables are involved in each. These equations are transcendental, however, and determination of the two variables must be done graphically. This task is relatively easy to tackle, if some forethought is given to the possible range of values for α and β. The use of a digital computer makes the problem of finding the values of α and β which satisfy Eqns. (6.5-1) and (6.5-4) less tedious, and once α and β are determined, δ is obtained from Eqn. (6.5-3).

6.6 Application of the New Reduction Approach to the Test System

The step response of the test system shown in Figure 2.3 has an initial slope of zero and an overshoot. To simulate this response, then, a second order system

with complex poles was required. From the response characteristics available, the steady state value, peak response, and peak time were selected as the criteria for a reduced system to meet. These characteristics were evaluated as follows.

Steady state value = 0.111111

Peak response value = 0.120689

Time of peak response = 2.900 seconds

From Eqn. (6.3-2) we have,

$$\beta = \frac{\pi}{2.9} = 1.083308$$

Then, by Eqn. (6.3-4),

$$\alpha = \frac{1}{2.9} \ln\left(\frac{.111111}{.120689 - .111111}\right)$$

$$= 0.845198$$

Finally, from Eqn. (6.3-3),

$$K = .111111(.714359 + 1.173556)$$

$$= 0.209768$$

The second order system transfer function is then

$$T(s) = \frac{0.209768}{s^2 + 1.690396s + 1.887915}$$

or, in state space form, the equation for the reduced

system is

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -1.887915 & -1.690396 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = \begin{bmatrix} .209768 & 0.0 \end{bmatrix} \underline{x}$$

6.7 Comparison of the Actual Response with the Reduced System Response

The actual and reduced system responses are shown in Figures 6.3, 6.4 and 6.5 for intervals of 5, 20 and 100 seconds, respectively. It is evident from these graphs that the specified characteristics have been met, as desired. It should be noted here, that some change in the specifications (such as the establishment of tolerances) could result in a response with less overall error, and the system can be easily altered according to the designer's judgment.

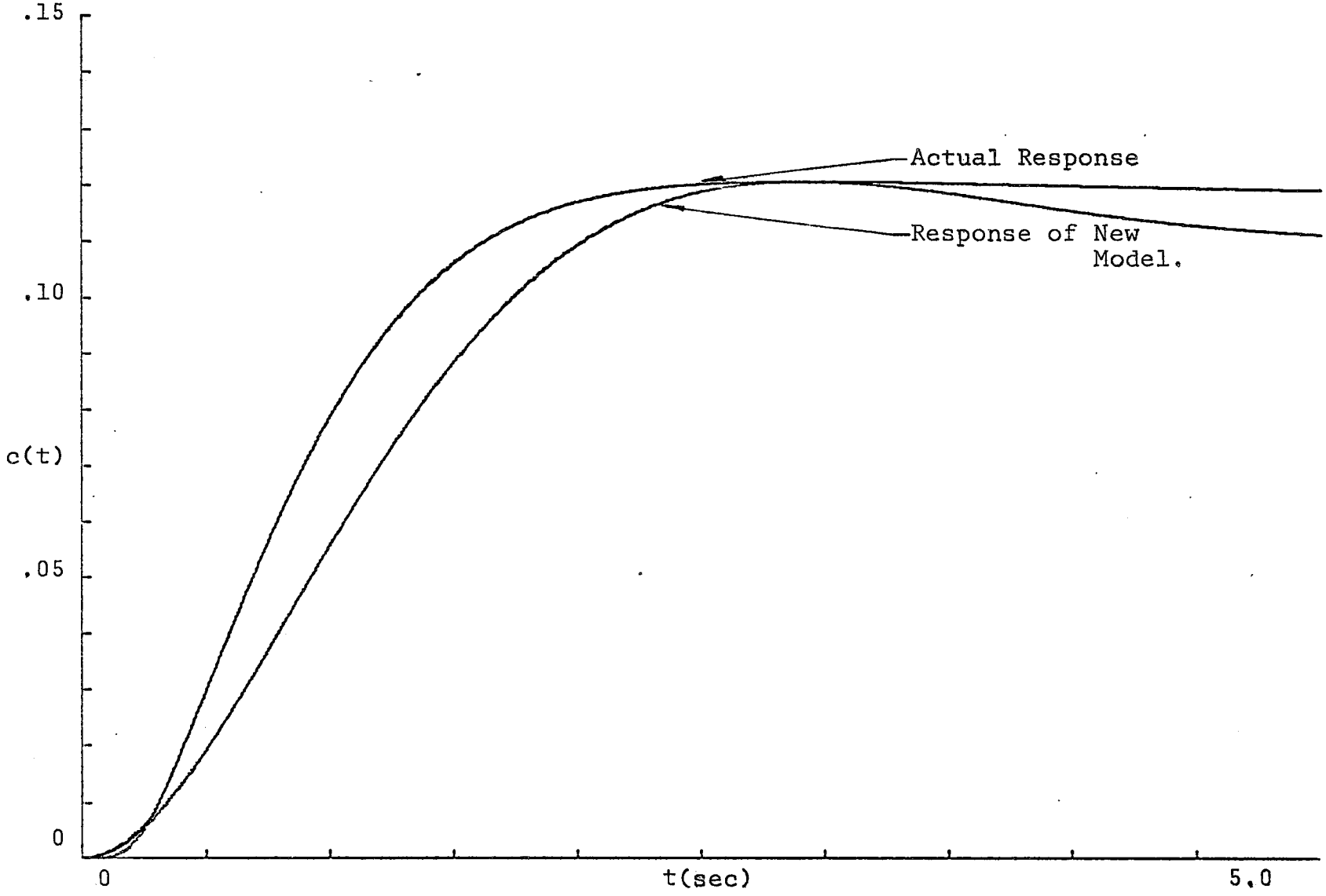


Figure 6.3 Response Comparison using the New Reduction Technique

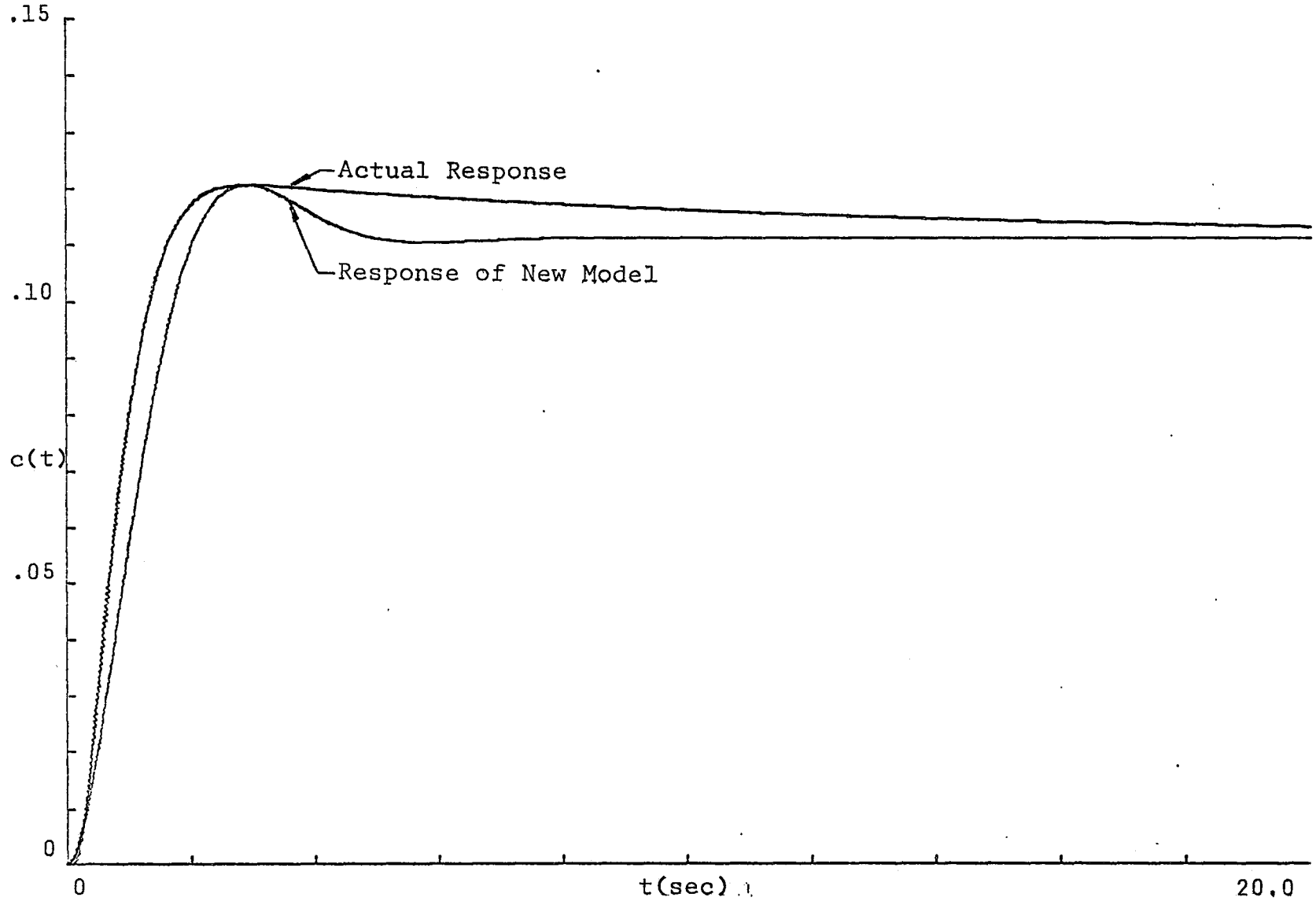


Figure 6.4 Response Comparison using the New Reduction Technique

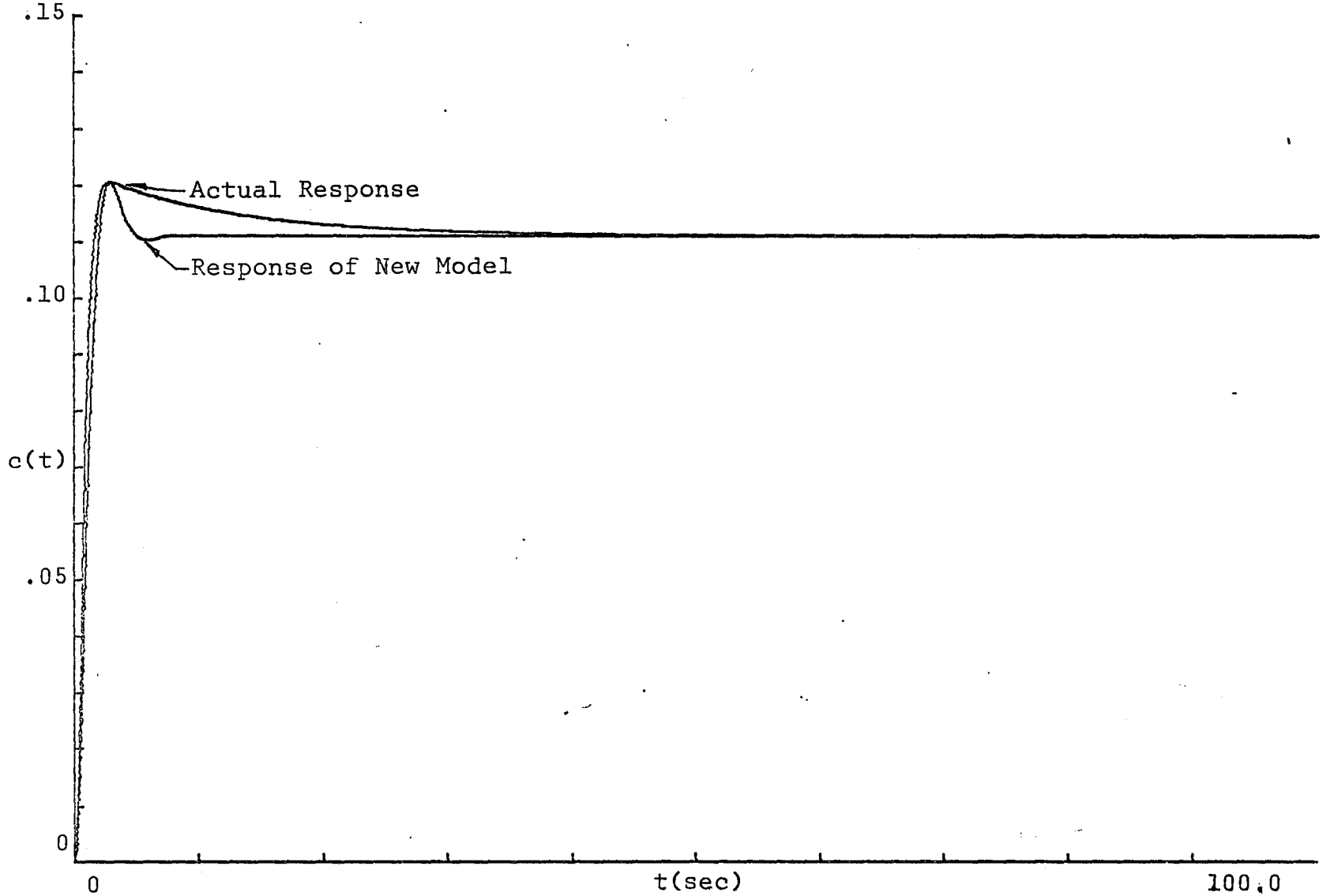


Figure 6.5 Response Comparison using the New Reduction Technique

CHAPTER 7

CONCLUSIONS

Three methods of high order system reduction have been applied to a test system and in each case a second order model was derived. These models were then tested with a step input and their responses were compared to the actual response. The state space equations and transfer functions for the three second order models appear in Table I. Each method has certain advantages and limitations. Method I is simple in concept and application, but is limited to single input systems. Method 2 can be applied to systems with multiple inputs, but the application is involved and tedious. Method 3 features a minimization of step response error and is easily applied, but does not guarantee an accurate steady state value. The third method is the most attractive one, since it involves error minimization, but there are some areas of improvement which would be desirable.

Firstly, the step response of a reduced system should ultimately approach that of the actual system. There will be some period in the response during which errors must occur because of system order reduction, but some guarantee

TABLE I

METHOD 1

State Space Equations

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -.09941 & -1.14644 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = [.01105 \quad .1299]\underline{x}$$

Transfer Function

$$\frac{C(s)}{R(s)} = \frac{0.1299s + .01105}{s^2 + 1.14644s + .09941}$$

Pole Locations

$$\lambda_1 = -1.04822$$

$$\lambda_2 = -0.09822$$

METHOD 2

State Space Equations

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -5.029956 & -4.112593 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = [.557628 \quad -.050525]\underline{x}$$

TABLE I (Continued)

Transfer Function

$$\frac{C(s)}{R(s)} = \frac{-.050525s + .557628}{s^2 + 4.112593s + 5.029956}$$

Pole Locations

$$\lambda_1 = -2.0562965 + j.8953215$$

$$\lambda_2 = -2.0562965 - j.8953215$$

METHOD 3

State Space Equations

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -2.687909 & -1.902574 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = [.30961 \quad .30961]\underline{x}$$

Transfer Function

$$\frac{C(s)}{R(s)} = \frac{.30961(s+1)}{s^2 + 1.902574s + 2.687909}$$

Pole Locations

$$\lambda_1 = -1.344 + j.316$$

$$\lambda_2 = -1.344 - j.316$$

TABLE I (Continued)

NEW METHOD

State Space Equations

$$\dot{\underline{x}} = \begin{bmatrix} 0.0 & 1.0 \\ -1.887915 & -1.690396 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$c = [.209768 \ 0.0]\underline{x}$$

Transfer Function

$$\frac{C(s)}{R(s)} = \frac{.209768}{s^2 + 1.690396s + 1.887915}$$

Pole Locations

$$\lambda_1 = -.845198 + j1.083307$$

$$\lambda_2 = -.845198 - j1.083307$$

of a very close steady state approximation should exist. Also, since there will be errors in the response curve, it would be preferable to have some choice as to where these errors occur.

The method proposed in Chapter 5, while simple in its application, offers a good deal of latitude in the fashioning of a low order model after the step response of a high order system. The problem of system identification is avoided since only thestep response of the system is required. This can normally be obtained to a sufficient degree of accuracy with relative ease. Areas of the actual response can be emphasized in the model at the discretion of the designer. For instance, in some systems it might be desirable to duplicate the transient portion of the response very closely, with little regard to settling time. In another system, the settling time could be the most important feature to approximate.

The new method not only allows an emphasis to be placed on different areas of the response, but because of its simplicity, it can be altered quickly to obtain a range of possible models which can be assessed by comparing the step responses. The transfer function for the model obtained using this method is shown in Table I.

In Chapter 5 this reduction concept is applied in order to arrive at a second order model for comparison

purposes. The concept is easily extended, however, to higher order models by specifying more of the desired response characteristics. For each increase in system order one more characteristic may be specified so that there will be the same number of equations as there are variables.

A second order model is frequently used because of its simplicity, especially in connection with optimal control theory. Too often this size of model does not contain step response characteristics which are accurate in some desired area. If this model is then used to arrive at an 'approximate' optimal control policy for the large system, the control policy may be considerably far from the optimum. By specifying certain of the model characteristics then, a 'approximate' optimal control may be derived which is consistent with the actual performance of the system.

This method of system reduction, with its simplicity, variability as to emphasis, and accuracy should make it useful in the design, analysis and optimal control areas of system engineering.

APPENDIX A

STATE SPACE EQUATIONS FOR THE TEST SYSTEM

A.1 From the block diagram of the test system in Figure 3.1, along with the chosen values for the variables, the transfer function of the system is calculated in Chapter 3.2 as $T(s) = \frac{C(s)}{R(s)}$

If we let $\frac{Y(s)}{R(s)} = \frac{1}{\text{Denominator of } T(s)}$

then $\frac{C(s)}{R(s)} = 375000sY + 31250Y$

By defining

$$x_1 = y$$

$$x_2 = \dot{x}_1 = \dot{y}$$

$$\vdots \quad \vdots$$

$$x_7 = \dot{x}_6$$

we can formulate the state space equations[2] as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -a_1 & -a_2 & -a_3 & -a_4 & -a_5 & -a_6 & -a_7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} R$$

where

$$a_1 = 281250.$$

$$a_2 = 3310875.$$

$$a_3 = 2814271.$$

$$a_4 = 853703.3$$

$$a_5 = 70341.905$$

$$a_6 = 4097.4038$$

$$a_7 = 83.635$$

and the output relationship is given by

$$c = [31250 \quad 375000] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The above equations are, then, in the standard form

$$\dot{\underline{x}} = A\underline{x} + Bu$$

$$c = H\underline{x}$$

There are various techniques[2] for reducing these equations to the simple form

$$\dot{\underline{x}} = A\underline{x} + Eu$$

where $c = x_1$. In effect, the B matrix from the original equations is changed to compensate for the derivatives of the driving function, which are present.

In order to apply one of the reduction techniques, this simplified form of the state space equation is required, so that the equations are restated as

$$\dot{\underline{x}} = A\underline{x} + Eu$$

where

$$E^T = [0\ 0\ 0\ 0\ 0\ 375000.\ -31331875.]$$

and now $c = x_1$

A.2 Evaluation of the System Eigenvalues and Eigenvectors

As mentioned in Chapter 3, it was necessary to solve for the eigenvalues of A, rather than finding the roots of the polynomial F(s), to obtain the accuracy desired.

The rounding off of numbers within the computer has a less significant effect on eigenvalues determined by matrix methods, than on roots determined by various non-matrix methods. This is increasingly evident for higher order equations.

Jacobi's method[4] for determining eigenvalues is

probably the most common of the matrix methods. It involves rotating the co-ordinate axes of the system until the off diagonal elements are reduced to zero. The remaining diagonal elements are then the eigenvalues. This is easily accomplished in one step for the second order system, but many more steps are required for more complex systems.

In higher ordered systems, the axes are rotated so as to reduce the largest off diagonal elements to zero. As this is done, however, previously zero-valued elements may take on non-zero values. The values are always reducing, so that with enough rotations the off diagonal elements may be reduced to some satisfactory tolerance. Jacobi's method has the limitation that the matrix to which the technique is applied must be symmetric. Eberlein modifies this method so that an arbitrary matrix may also be reduced to the stage where the eigenvalues appear on the main diagonal. In a combination of integrated steps the matrix undergoes two-dimensional transformations which firstly make the matrix arbitrarily close to normal, and then reduce off diagonal elements to steadily decreasing values.

The research work leading to the above method was partially sponsored by a computer company so that the method is particularly adapted to machine computation.

The eigenvalues (or roots) for the test system were obtained using Eberlein's method, which is a standard subroutine built into the computer library, and these eigenvalues are as follows.

$$\lambda_1 = -.09193240$$

$$\lambda_{2,3} = -2.024383 \pm j.9646465$$

$$\lambda_{4,5} = -7.672749 \pm j13.44463$$

$$\lambda_{6,7} = -32.07440 \pm j38.85934$$

These roots were tested in the equations and found to be correct to seven decimal places which was deemed sufficient accuracy for practical applications of the reduction methods requiring these roots.

The modal matrix, which is also required later, consists of the eigenvectors arranged in columns. Since the A matrix is in the normal form, the eigenvectors are easily determined from the eigenvalues as

$$\underline{X}_i^T = [1 \ \lambda_i \ \lambda_i^2 \ \ldots \ \lambda_i^7]$$

where $\underline{X}_i$ is the eigenvector corresponding to the ith eigenvalue (λ_i).

The eigenvectors, calculated from the eigenvalues in the same order, as shown below

$$\underline{x}_1 = \begin{bmatrix} 1.0 \\ -9.19324 \times 10^{-2} \\ 8.451567 \times 10^{-3} \\ -7.769728 \times 10^{-4} \\ 7.142898 \times 10^{-5} \\ -6.566638 \times 10^{-6} \\ 6.036868 \times 10^{-7} \end{bmatrix}$$

$$\underline{x}_{2,3} = \begin{bmatrix} 1.0 \\ -2.024383 \pm j.9646465 \\ 3.167583 \mp j3.905628 \\ -2.644851 \pm j10.96208 \\ -5.220345 \mp j24.74280 \\ 34.43603 \pm j45.05312 \\ -113.1721 \mp j57.98616 \end{bmatrix}$$

$$\underline{x}_{4,5} \begin{bmatrix} 1.0 \\ -7.672749 \mp j13.44463 \\ -1.218870 \times 10^2 \pm j2.063146 \times 10^2 \\ 3.709032 \times 10^3 \pm j5.572618 \times 10^3 \\ -2.770925 \times 10^4 \mp j5.029414 \times 10^4 \\ -4.63580 \times 10^5 \pm j7.58435 \times 10^5 \\ 1.375381 \times 10^7 \pm j4.133803 \times 10^5 \end{bmatrix}$$

$$\underline{x}_{6,7} = \begin{bmatrix} 1.0 \\ -3.207440 \times 10 \pm j3.885934 \times 10 \\ -4.812811 \times 10^{2} \mp j2.492780 \times 10^{3} \\ 1.123046 \times 10^{5} \pm j6.125217 \times 10^{4} \\ -5.982322 \times 10^{6} \pm j2.399456 \times 10^{6} \\ 9.863811 \times 10^{7} \mp j3.094302 \times 10^{8} \\ 8.860495 \times 10^{9} \pm j1.375780 \times 10^{10} \end{bmatrix}$$

APPENDIX B

SOLUTION OF THE STATE SPACE EQUATIONS USING A DIGITAL COMPUTER

B.1 Method of Solution

The solution of the equations

$$\dot{\underline{x}}(t) = A\underline{x}(t) + B\underline{u}(t)$$

is

$$\underline{x}(t) = e^{A(t - t_0)} \cdot \underline{x}(t_0) + e^{At} \int_{t_0}^{t} e^{-A\tau} B\underline{u}(\tau) d\tau$$

If t_0 is taken to be zero, which is not unduly restrictive, the solution becomes

$$\underline{x}(t) = e^{At} \left[\underline{x}(0) + \int_0^t e^{-A\tau} B\underline{u}(\tau) d\tau \right]$$

In order to use a digital computer to solve the state space equations, the solution must first be expressed in a discrete form, since the computing process is not continuous. This solution is easily obtained if the driving function, or input, is held constant between the sampling intervals, or if the input can be suitably approximated by a series of step functions which can be considered constant between the sampling intervals.

The input used here is a step function, so that the solution of the state space equations can be represented equivalently in its discrete form as

$$\underline{x}(k+1)T = e^{AT}\left[\underline{x}(kT) + \int_0^T e^{-A\tau} B\underline{u}(kT)d\tau\right]$$

.....(B.1-1)

where T = sampling interval

and k = sampling instant

Now, e^{AT} can be represented in an infinite series form as

$$e^{AT} = (I + AT + \frac{(AT)^2}{2!} + \frac{(AT)^3}{3!} + \ldots)$$

.....(B.1-2)

Also,

$$e^{-AT} = (I - AT + \frac{(AT)^2}{2!} - \frac{(AT)^3}{3!} + \ldots)$$

.....(B.1-3)

If Eqn. (B.1-1) is integrated between the limits shown and Eqn. (B.1-3) substituted in the resulting equation, the solution becomes

$$\underline{x}(k+1)T = e^{AT}\left[\underline{x}(kT) + [I - \frac{AT}{2!} + \frac{(AT)}{3!} - \ldots]B\underline{u}T\right]$$

.....(B.1-4)

With the series substitution for e^{AT} in Eqn. (B.1-4), the solution is in a form that lends itself to machine computation.

Each series in this expression must be truncated

at some point, but each series is uniformly convergent, so that assessment of the value of additional terms need only be made and a suitable criterion applied for truncation.

Increasing the value of T requires taking a greatly increasing number of terms in the series in order to obtain reasonable accuracy. However, if a very small value of T is selected, then each series converges quickly and a high degree of accuracy can be demanded. Having obtained the solution at this time T, the process can be repeated using the same accurate values for the series and merely updating the initial conditions. The final solution is, in fact, arrived at by taking many small steps and, since this iterative procedure does not require continual evaluation of the series, minimal computer time is required.

A general computer program for the solution of equations in the form $\dot{\underline{x}} = A\underline{x} + B\underline{u}$, using the technique above, is included in Appendix B.2. This program was used for obtaining the responses of the test system and all the reduced models of this system. A time interval of 0.01 seconds was found to provide close agreement (five figure accuracy) between the program output and actual values calculated for a sample system.

The following terms are applied to various parts of the solution for use in the computer program.

$$ANS \equiv (I + AT + \frac{(AT)}{2!} + \ldots)$$

$$DMA \equiv (I - \frac{AT}{2!} + \frac{(AT)}{3!} - \ldots)T$$

$$DBM \equiv [DMA][B\underline{u}]$$

$$DBX \equiv \underline{X}(0) + DBM$$

$$SOLN \equiv [ANS][DBX]$$

The values of ANS and DMA are calculated in subroutines EXPAT and DSTEP, respectively.

It should be noted here that the step function used in this example may be substituted with any linear driving function by changing the integration in Eqn. (B.1-1) and thus revising the subroutine DSTEP (which yields DMA) in the program. The convergence requirements for each series is that successive terms in the expansion be less than .0001% of the preceding term.

APPENDIX B.2

COMPUTER PROGRAM FOR SOLUTION OF STATE SPACE EQUATIONS USING A DIGITAL COMPUTER

```
C       THIS IS MAIN PROGRAM FOR STATE SPACE SOLN. OF LINEAR
C       DIFFERENTIAL EQUATIONS BY A SERIES METHOD
C       EQUATIONS ARE OF FORM  DX/DT=(A)X+(B)U,  C=(H)X
C       TRANSITION MATRIX IS SOLVED NUMERICALLY USING A SMALL TIME(DT).
C       INITIAL CONDITIONS ARE CONTINUALLY UPDATED USING OUTPUT DATA.
C       X=INITIAL CONDITIONS. T=TIME INCREMENTS(SEC). TTIM=TOTAL TIME
C       H IS OUTPUT MATRIX.
C       KOUNT IS THE DESIRED TIME INTERVALS FOR WRITTEN OUTPUT.
C       KOUNTR IS THE DESIRED TIME INTERVAL FOR PUNCHED OUTPUT.
C       NUMBR IS THE NUMBER OF RESPONSE POINTS PUNCHED OUT.
C
        DIMENSION A(25,25),B(25),X(25),DBM(25),DOX(25),H(25)
        DIMENSION ANS(25,25),DMA(25,25),PRT(25,25)
        DIMENSION Y(500),TIM(500)
C
C       READ TIME INCREMENT, TOTAL TIME, MATRIX SIZE, AND COUNTERS.
C
        READ(5,10)N,T,TTIM,KOUNT,KOUNTR,NUMBR
        DO3I=1,N
    3   READ(5,11)(A(I,J),J=1,N)
        READ(5,11)(B(I),I=1,N)
        READ(5,11)(X(I),I=1,N)
        READ(5,11)(H(I),I=1,N)
C
C       ECHO PRINT INPUT DATA.
C
        WRITE(6,22)N,T,TTIM,KOUNT,KOUNTR
        WRITE(6,24)
        DO18I=1,N
   18   WRITE(6,20)(A(I,J),J=1,N)
        WRITE(6,26)
        DO19I=1,N
   19   WRITE(6,21)I,B(I),I,X(I),I,H(I)
        WRITE(6,27)
        L=1
        M=1
        K=1
        TS=T
```

```
C
C       FIND INTEGRAL (EXP.-AT)*U
C
        CALL DSTEP(A,T,N,DMA,DE)
        IF(DE.GT.500.) GO TO 300
C
C       FIND EXP.AT
C
        CALL EXPAT(A,T,N,ANS,D)
        IF(D.GT.500.) GO TO 300
C
C       WRITE NUMBER OF ITERATIONS FOR CONVERGENCE OF ABOVE SERIES.
C
        WRITE(6,28)D,DE
C
C       WRITE INTEGRAL (EXP.-AT)*U
C
        WRITE(6,29)
        DO 101I=1,N
        WRITE(6,20)(DMA(I,J),J=1,N)
  101   CONTINUE
C
C       WRITE EXP.AT
C
        WRITE(6,201)
        DO 102I=1,N
        WRITE(6,20)(ANS(I,J),J=1,N)
  102   CONTINUE
        WRITE(6,202)
C
C       MULTIPLY DMA MATRIX BY B COLUMN
C
        DO 16I=1,N
        DBM(I)=0.0
        DO 16J=1,N
   16   DBM(I)=DBM(I)+DMA(I,J)*B(J)
C
C       ADD DBM COLUMN TO X(O) VECTOR
C
   17   DO 12I=1,N
   12   DBX(I)=DBM(I)+X(I)
C
C       MULTIPLY EXPAT BY DBX COLUMN
C
        DO 13I=1,N
        X(I)=0.0
        DO 13J=1,N
   13   X(I)=X(I)+ANS(I,J)*DBX(J)
C
C       CALCULATE RESPONSE
C
```

```
      C=0.0
      DO 5 I=1,N
      C=H(I)*X(I)+C
    5 CONTINUE
      IF(L.EQ.KOUNTR) GO TO 1
    8 IF(K.EQ.KOUNT) GO TO 2
    4 T=TS+T
      L=L+1
      K=K+1
      IF(T.GT.TTIM) GO TO 400
     ,GO TO 17
C
C     STORE DATA POINTS AT DESIRED INTERVALS.
C
    1 Y(M)=C
      TIM(M)=T
      L=L-KOUNTR
      M=M+1
      GO TO 8
C
C     WRITE OUTPUT AT DESIRED INTERVALS
C
    2 WRITE(6,23)T
      WRITE(6,204)(I,X(I),I=1,N)
      WRITE(6,203) C
      K=K-KOUNT
      GO TO 4
C
C     PUNCH DATA POINTS FOR GRAPHING.
C
  400 WRITE(7,205)(Y(M),M=1,NUMBR)
      WRITE(7,206)(TIM(M),M=1,NUMBR)
C
C     READ FORMATS
C
   10 FORMAT(I10,2F10.4,3I10)
   11 FORMAT(7F10.6)
C
C     WRITE FORMATS
C
   20 FORMAT(1X,10E13.6)
   21 FORMAT(10X,1HB,I2,2H= ,E13.6,10X,1HX,I2,5H(0)= ,E13.6,10X,1HH,I2,2
     1H= ,E13.6)
   22 FORMAT(1H1,10X,13HMATRIX DIM.= ,I2,5X,9HDELTA T= ,F6.4,4HSEC.,
     15X,12HTOTAL TIME= F9.4,4HSEC.,5X,6HKOUNT=,I4,5X,7HKOUNTR=,I4//)
   23 FORMAT(1X,3HT= ,F8.3,3HSEC)
   24 FORMAT(55X,8HA MATRIX//)
   26 FORMAT(////8X,14HCONTROL VECTOR,10X,14HINITIAL VALUES,10X,13HOUTP
     1UT MATRIX)
   27 FORMAT(1H1,15X,18HOUTPUT INFORMATION)
```

```
28  FORMAT(/10X,20HEXPAT SERIES ITER.= ,F6.0,5X,
   120HDSTEP SERIES ITER.= ,F6.0)
29  FORMAT(///50X,19HINTEGRAL(EXP.-AT)*U/)
201 FORMAT(///53X,13HEXP.AT MATRIX/)
202 FORMAT(1H1,50X,34HSOLUTION OF DIFFERENTIAL EQUATIONS///)
203 FORMAT(10X,10HRESPONSE= ,F15.6//)
204 FORMAT(1X,6(2HX(,I2,3H)= ,E13.6))
205 FORMAT(6E13.6)
206 FORMAT(11F7.3)
300 STOP
    END
```

```
      SUBROUTINE EXPAT(A,T,N,ANS,D)
C     THIS SUBROUTINE CALCULATES EXPAT BY SERIES METHOD TO A CONVERGENCE
C     OF .0001 PERCENT, WITH A LIMIT OF 500 ITERATIONS.
C     T=TIME IN SECONDS/N=DIMENSION OF SQUARE A MATRIX
C
      DIMENSION ANS(25,25),S(25,25),A(25,25),C(25,25),PRT(25,25)
C
C     BUILD IDENTITY MATRIX
C
      DO34I=1,N
      DO34J=1,N
      IF(I.EQ.J)S(I,J)=1.0
      IF(I.EQ.J) GO TO 34
      S(I,J)=0.0
   34 CONTINUE
C
C     TAKE FIRST SUM IN SERIES
C
      DO33I=1,N
      DO33J=1,N
      PRT(I,J)=S(I,J)
   33 ANS(I,J)=S(I,J)
C
C     SET UP GENERAL TERM FOR COEFFT. P
C
      D=1.0
      P=1.0
   30 P=(T/D)*P
C
C     CALCULATE NEXT TERM IN SERIES
C
      DO40I=1,N
      DO40J=1,N
      C(I,J)=0.0
      DO40K=1,N
   40 C(I,J)=C(I,J)+A(I,K)*S(K,J)
C
C     CHECK FOR CONVERGENCE
C
      DO42I=1,N
      DO42J=1,N
      TOL=.000001*PRT(I,J)
      CUTE=ABS(C(I,J)*P)
      IF(CUTE.GT.TOL) GO TO 35
   42 CONTINUE
      GO TO 100
C
C     ADD THIS TERM TO PREVIOUS ONES
C
```

```
 35 DO41I=1,N
    DO41J=1,N
    ANS(I,J)=C(I,J)*P+ANS(I,J)
    PRT(I,J)=ABS(C(I,J)*P)
 41 S(I,J)=C(I,J)
    D=D+1.0
    IF(D.GT.500.) GO TO 43
    GO TO 30
 43 WRITE(6,45)
 45 FORMAT(1X,26HA MATRIX HAS NOT CONVERGED)
100 RETURN
    END
```

```
      SUBROUTINE DSTEP(A,T,N,DMA,DE)
C     THIS SUBR. CALCULATES INTEGRAL (EXP-AT*U) BY A SERIES, TO A
C     CONVERGENCE OF .0001 PERCENT, WITH A LIMIT OF 500 ITERATIONS.
C     U=UNIT STEP I/P. N=DIMENSION OF MATRIX. T=TIME IN SEC.
C
      DIMENSION DMA(25,25),S(25,25),A(25,25),C(25,25),PRT(25,25)
C
C     BUILD I*T MAT-IX
C
      DO54I=1,N
      DO54J=1,N
      IF(I.EQ.J)S(I,J)=T
      IF(I.EQ.J) GO TO 54
      S(I,J)=0.0
   54 CONTINUE
C
C     TAKE FIRST SUM IN SERIES
C
      DO53I=1,N
      DO53J=1,N
      PRT(I,J)=S(I,J)
   53 DMA(I,J)=S(I,J)
C
C     SET UP GENERAL TERM FOR COEFFT. P
C
      DE=2.0
      P=1.0
   50 P=-1.0*P*(T/DE)
C
C     CALCULATE NEXT TERM IN SERIES
C
      DO60I=1,N
      DO 60J=1,N
      C(I,J)=0.0
      DO60K=1,N
   60 C(I,J)=C(I,J)+A(I,K)*S(K,J)
C
C     CHECK FOR CONVERGENCE
C
      DO62I=1,N
      DO62J=1,N
      TOL=.000001*PRT(I,J)
      CUTE=ABS(C(I,J)*P)
      IF(CUTE.GT.TOL) GO TO 55
   62 CONTINUE
      GO TO 200
C
C     ADD THIS TERM TO PREVIOUS ONES
C
```

```
 55 DO61I=1,N
    DO61J=1,N
    DMA(I,J)=C(I,J)*P+DMA(I,J)
    PRT(I,J)=ABS(C(I,J)*P)
    S(I,J)=C(I,J)
 61 CONTINUE
    DE=DE+1.0
    IF(DE.GT.500.) GO TO 63
    GO TO 50
 63 WRITE(6,65)
 65 FORMAT(1X,26HD MATRIX HAS NOT CONVERGED)
200 RETURN
    END
```

APPENDIX C

APPLICATION OF METHOD 1 USING A DIGITAL COMPUTER

```
C     THIS PROGRAM FORMS THE CONTINUED FRACTION EXPANSION OF A TRANSFER
C     FUNCTION. IF NUMERATOR IS OF SMALLER ORDER THAN DENOMINATOR, FILL
C     OUT TO SAME ORDER WITH ZERO COEFFTS.. THE NUMERATOR COEFFTS. ARE
C     READ IN ASCENDING ORDER TO FORM TN VECTOR. THE DENOMINATOR
C     COEFFTS. ARE READ IN ASCENDING ORDER TO FORM TD VECTOR.
C     N=NO. OF COEFFTS. IN DENOMINATOR. M= ORDER OF REDUCED SYSTEM
C
      DIMENSION TN(25),TD(25),RUDIV(25),REM(25)
C
C     READ ORDER OF PRESENT SYSTEM AND ORDER OF DESIRED SYSTEM.
C
      READ(5,10) N,M
      READ(5,11)(TN(I),I=1,N)
      READ(5,11)(TD(I),I=1,N)
C
C     WRITE ORDER OF SYSTEM AND DESIRED ORDER.
C
      WRITE(6,20) N,M
C
C     WRITE COEFFTS. OF NUMERATOR AND DENOMINATOR.
C
      WRITE(6,23)
      DO 7I=1,N
      K=I-1
    7 WRITE(6,21)K,TN(I),K,TD(I)
      WRITE(6,24)
C
C     INITIALIZE COUNTERS.
C     L=RUNNING ORDER OF TRANSFER FUNCTION.
C     KOUNT=RUNNING ORDER OF REDUCED SYSTEM.
C
      KOUNT=0
      L=N
C
C     THIS IS DIVISION WITH SAME POWERS OF S.
C
    6 H1=TD(1)/TN(1)
      DO 1I=1,L
      RUDIV(I)=H1*TN(I)
      REM(I)=TD(I)-RUDIV(I)
```

```
C
C       FORM NEW FRACTION
C
    1   TD(I)=TN(I)
C       SHIFT FOR LOST TERM AND PUT IN NUMERATOR.
C
        J=L-1
        DO 2I=1,J
    2   TN(I)=REM(I+1)
C
C       THIS IS DIVISION WITH UNLIKE POWERS OF S.
C
        H2=TD(1)/TN(1)
        DO 3I=1,J
    3   RUDIV(I)=H2*TN(I)
        RUDIV(L)=0.0
        DO 4I=1,L
    4   REM(I)=TD(I)-RUDIV(I)
C
C       FORM NEW FRACTION
C
        DO 5I=1,J
        TD(I)=TN(I)
C
C       SHIFT AND PUT IN NUMERATOR.
C
    5   TN(I)=REM(I+1)
        L=L-1
        WRITE(6,22)H1,H2
        KOUNT=KOUNT+1
        IF(KOUNT.EQ.M) GO TO 100
        GO TO 6
C
C       READ FORMATS
C
   10 FORMAT(2I5)
   11 FORMAT(5F15.5)
C
C       WRITE FORMATS
C
   20 FORMAT(1H1,10X,25HORDER OF EXISTING SYSTEM=,I2,10X,24HORDER OF DES
      1IRED SYSTEM=,I2///)
   21 FORMAT(2(28X,2HS(,I2,3H) =,F15.5))
   22 FORMAT(28X,F15.5,28X,F15.5)
   23 FORMAT(20X,22HNUMERATOR COEFFICIENTS,20X,23HDENOMINATOR COEFFICIEN
      1TS/)
   24 FORMAT(1H1,22X,14HEVEN QUOTIENTS,22X,16HQUOTIENTS OVER S/)
  100 STOP
        END
```

APPENDIX D

THE APPLICATION OF METHOD 2

D.1 Formulation of the Reduced Matrices

The various matrices which were evaluated during the application of this method are given below. These matrices are in the author's notation and follow a sequence which leads from the reduced system back to the test system.

$$DA^{*}D^{-1} = \begin{bmatrix} -3.907092 & 1.29 \times 10^{-6} \\ -3463089. & -.1416739 \end{bmatrix}$$

$$DB^{*} = \begin{bmatrix} -.05052537 \\ 437819.7343 \end{bmatrix}$$

$$D = \begin{bmatrix} 0.9002 & 0.0 \\ 0.0 & -11597.88912 \end{bmatrix}$$

$$A^{*-1}B^{*} = \begin{bmatrix} -.1233025 \\ 32.33347 \end{bmatrix}$$

$$A^{-1}B = \begin{bmatrix} -.111014 \\ 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \\ 375000. \end{bmatrix}$$

$$A^* = \begin{bmatrix} -3.907092 & -.01663581 \\ 269.0063 & -.1416739 \end{bmatrix}$$

$$U_1 = \begin{bmatrix} 1.0 & +j0.0 & 1.0 & +j0.0 \\ -113.1721 & -j57.98616 & -113.1721 & +j57.98616 \end{bmatrix}$$

$$\lambda_1 = \begin{bmatrix} -2.024383 + j.9646465 & 0 \\ 0 & -2.024383 - j.9646465 \end{bmatrix}$$

$$B^* = \begin{bmatrix} -.0561393 \\ -37.74995 \end{bmatrix}$$

Matrices U, A, and B are recorded in Appendices A.1 and A.2.

D.2 Inversion of a Complex Matrix with Singular Real and Imaginary Parts

If $(A + jB)$ is non singular, there exists a unique matrix $(C + jD)$ such that

$$(A + jB)^{-1} = C + jD$$

or, $(A + jB)(C + jD) = I$, i.e., $AC - BD = I$

$$BC + AD = 0$$

and $(C + jD)(A + jB) = I$, i.e., $CA - DB = I$

$$CB + DA = 0$$

Now, if we let $E = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$ and $F = \begin{bmatrix} C & D \\ -D & C \end{bmatrix}$

$$\text{Then, } EF = \begin{bmatrix} AC-BD & AD+BC \\ -BC-AD & -BD+AC \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = I$$

$$\text{and} \quad FE = \begin{bmatrix} CA-DB & CB+DA \\ -DA-CB & -DB+CA \end{bmatrix} = I$$

$$\text{Thus} \quad F = E^{-1}$$

Since the existence and uniqueness of C and D is guaranteed if (A + jB) is non-singular, the existence of F is also guaranteed.

Hence, E possesses a unique inverse F, and is therefore, non-singular.

APPENDIX E

THE APPLICATION OF METHOD 3

E.1 Formulation of the Reduced Matrices

Anderson applies his method to a continuous system model which can also be represented by discrete-time equations.[9] The test system used here can also be described by discrete time equations with an input which does not vary between sampling instances. Then, given $\underline{x}(0)$ and the sequence u(kT), k = 0,1,---ℓ, the state of the test system at the sampling instances can be computed for all t > 0. T was chosen to be .01 seconds for convenience, and ℓ = 500 to satisfy the total time interval requirements.

With the reduced system described by the equations

$$\dot{\underline{x}}_m = A_m \underline{x}_m + B_m u$$

the required matrices for the second order case are defined as $[A_m B_m] = B_1^T M (M^T M)^{-1}$

where

$$B_1 = \begin{bmatrix} \dot{x}_1(0) & \dot{x}_2(0) \\ \dot{x}_1(T) & \dot{x}_2(T) \\ \vdots & \\ \dot{x}_1(kT) & \dot{x}_2(kT) \end{bmatrix}$$

and

$$M = \begin{bmatrix} x_1(0) & x_2(0) & u(0) \\ x_1(T) & x_2(T) & u(T) \\ \vdots & \vdots & \vdots \\ x_1(kT) & x_2(kT) & u(kT) \end{bmatrix}$$

Using the computer program in Appendix E.2, the required matrices are determined as

$$A_m = \begin{bmatrix} 0.0 & 1.0 \\ -2.687909 & -1.902574 \end{bmatrix}$$

and $$B_m = \begin{bmatrix} 0.0 \\ 0.309610 \end{bmatrix}$$

APPENDIX E.2

DIGITAL COMPUTER PROGRAM FOR APPLICATION OF METHOD 3

```
C     THIS PROGRAM APPLIES METHOD NO.3 TO AN ARBITRARY SYSTEM.
C     N= NO. OF RESPONSE POINTS EVALUATED. T= TIME INTERVAL BETWEEN
C     RESPONSE POINTS. L= DESIRED ORDER OF SYSTEM. K= L + NO. OF INPUTS.
C
      DIMENSION AM(500,3),AMT(3,500),BT(2,500),A(10,10),N1(10,10)
      DIMENSION PROD(3,3),AMBM(3,3)
      READ(5,10)N,T,L,K
      WRITE(6,20) N,T,L,K
      DO 1I=1,N
      READ(5,11)(AM(I,J),J=1,K)
    1 CONTINUE
      DO 2I=1,N
      BT(1,I)=AM(I,2)
      BT(2,I)=AM(I,3)
      AM(I,3)=1.
      AMT(1,I)=AM(I,1)
      AMT(2,I)=AM(I,2)
      AMT(3,I)=AM(I,3)
    2 CONTINUE
      WRITE(6,26)
      DO 3I=1,K
      DO 4J=1,K
      A(I,J)=0.0
      DO 4M=1,N
      A(I,J)=AMT(I,M)*AM(M,J)+A(I,J)
    4 CONTINUE
      WRITE(6,21)(A(I,J),J=1,K)
    3 CONTINUE
      NNN=10
      NN=K
      ZERO=1.E-10
      CALL INVMAT(A,NNN,NN,ZERO,IERR,N1)
      IF(IERR.EQ.0) GO TO 5
      WRITE(6,22)IERR
      GO TO 100
    5 WRITE(6,23)
      DO 6I=1,NN
      WRITE(6,21)(A(I,J),J=1,NN)
    6 CONTINUE
      WRITE(6,24)
```

```
      DO 7I=1,L
      DO 8J=1,K
      PROD(I,J)=0.0
      DO 8M=1,N
      PROD(I,J)=BT(+,M)*AM(M,J)+PROD(I,J)
    8 CONTINUE
      WRITE(6,21)(PROD(I,J),J=1,K)
    7 CONTINUE
      WRITE(6,25)
      DO 9I=1,L
      DO 30J=1,K
      AMBM(I,J)=0.0
      DO 30M=1,K
      AMBM(I,J)=PROD(I,M)*A(M,J)+AMBM(I,J)
   30 CONTINUE
      WRITE(6,21)(AMBM(I,J),J=1,K)
    9 CONTINUE
C
C     READ FORMATS
C
   10 FORMAT(I10,F10.4,2I10)
   11 FORMAT(3E13.6)
C
C     WRITE FORMATS
C
   20 FORMAT(1H1,20X,7H(K+1)T=,I3,5X,2HT=,F5.3,5X,2HR=,I3,5X,4HR+M=,I3/)
   21 FORMAT((7(5X,E13.6))///)
   22 FORMAT(//5X,22HINVERSION FAILS, IERR=,I2)
   23 FORMAT(1H1,10X,13H(MTM) INVERSE//)
   24 FORMAT(//10X,13H(BTM) PRODUCT//)
   25 FORMAT(//10X,18HAMBM OUTPUT MATRIX//)
   26 FORMAT(//10X,13H(MTM) PRODUCT//)
  100 STOP
      END
```

REFERENCES

1. RICHARD C. DORF: "Modern Control Systems", Addison-Wesley Inc., pp.167-168,1967.

2. KATSUHIKO OGATA: "State Space Analysis of Control Systems", Prentice-Hall Inc., 1967.

3. P.J. EBERLEIN: J. Soc. Ind. Appl. Math. Vol. 10, No. 1, Page 74.

4. SHAN S. KUO: "Numerical Methods and Computers", Addison-Wesley Inc., 1965.

5. C.F. CHEN; L.S. SHIEH: "A Novel Approach to Linear Model Simplification" Preprints of the Joint Automatic Control Conference, pp.454-460, June, 1968.

6. E.J. DAVISON: "A Method for Simplifying Linear Dynamic Systems", IEEE Trans. Automatic Control, vol. AC-11, pp.93-101, January 1966.

7. E.J. DAVISON: "A New Method for Simplifying Large Linear Dynamic Systems" IEEE Trans. Automatic Control, vol. AC-13, pp.214-215, April, 1968.

8. J.H. ANDERSON: "Geometrical approach to reduction of dynamical systems", Proc. IEE (London), vol. 114, pp.1014-1018, July, 1967.

References Cont'd.

9. J.H. ANDERSON; H.W. KWAN: "The Systematic Reduction of Complex Process Models; Two Case Studies": I.E.E. Conference on Computer aided Design, April 1969, Southampton, England.

10. B.Z. VULIKH: "Introduction to Functional Analysis", Addison-Wesley Inc., 1963.

11. P.KOKOTOVIC; P. SANNUTI: "Singular Perturbation Method for Reducing the Model Order in Optimal Control Design", Preprints of the Joint Automatic Control Conference, pp.468-476, June,1968.

BIBLIOGRAPHY

1. DeRUSSO; ROY; CLOSE: "State Variables for Engineers", Wiley, 1967.

2. KATSUHIKO OGATA: "State Space Analysis of Control Systems", Prentice-Hall Inc., 1967.

3. ANDREW P. SAGE: "Optimum Systems Control", Prentice-Hall Inc., 1968.

4. RICHARD C. DORF: "Modern Control Systems", Addison Wesley Inc., 1967.

5. SHAN S. KUO: "Numerical Methods and Computers", Addison-Wesley Inc., 1965.