A Dissertation

Presented to
the Faculty of the Department of Electrical Jngineering University of Houston

In Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

by<br>Kyung Sik Yoo

Yay 1974

## ACKNOWLEDGMENT

The author expresses his hearty appreciation to Dr. E. D. Denman for guidance and valuable suggestions of the dissertation work. Sincere appreciation is also expressed to Dr. G. S. Dawkins, Prof. W. P. Schneider, Dr. L. S. Shieh, and Dr. R. D. Sinkhorn for their time and consideration. My thanks are also due to Ms. Michele Maes for her neat and fast typing.

## An Abstract of a Dissertation

Presented to the Faculty of the Department of Electrical Engineering University of Houston

In Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

by<br>Kyung Sik Yoo

May 1974

## ABSTRACT

A new uncoupling algorithm for system differential equations based on the matrix sign function is presented in this paper. Special forms of system equations arising in many classes of system optimization are investigated. The set of $n$ differential equations formed by the states and costates are uncoupled requiring only the integration of one matrix differential equations of the order $n / 2$. Among the special forms considered are the stiff state equations with constant coefficients for which a numerical algorithm is presented. The algorithm groups the system eigenvalues into separated subsets and generates completely uncoupled filter matrices. Several example solutions are developed using the new method to illustrate the outlined procedures and the numerical accuracy.
CHAPTER
PAGE
I. INTRODUCTION ..... 1
II. DEVELOPMENT OF ALGORITHMS ..... 7
2.1. Generation of the Transformation Matrix ..... 7
2.2 Block Diagonal Form of a General SquareMatrix . . . . . . . . . . . . . . 8
2.3 Continuous Iineax Systems ..... 11
2.4 Discrete Linear Systems ..... 14
2.5 Generation of a Filter Matrix ..... 22
2.5.1. Filter Matrix for Eigenvalues
with the Same Sign ..... 22
2.5.2 Filter Matrix for the Same Order of Magnitude ..... 24
2.5.3 Bandpass Filter Matrix ..... 26
2.6 Examples ..... 30
2.6.1 Continuous Systems Matrix ..... 30
2.6.2 Discrete Systems Matrix ..... 31
2.6.3 Eilter Matrix ..... 34
III. SOLUTION OF THE UNCOUPLED CONTINUOUS SYSTEM ..... 36
3.1 Linear Systems with Quadratic Criteria. ..... 36
3.1.1 Solution by a Transition Ma- trix ..... 38
3.1.2 Solution by the Riccati Equa- tion ..... 38
3.1.3 The Hamilton-Jacobi-Bellman
Equation ..... 40
3.2 A Reduced Linear Regulator ..... 41
3.3 Examples ..... 49
IV. SOLUTION OF THE USCOUPISED DISCRETE SYSTEM ..... 52
4.1 A Discrete Iinear Regulator ..... 52
4.2 The Riccati Equation for Discrete Optimal
Feedback Gain ..... 53
4.3 The Discrete Kalman Filter ..... 56
4.4 Reduced Discrete Linear Systems ..... 57
4.5 An Algebraic Solution for a Discrete Riccati Equation ..... 62
4.6 Examples ..... 65
4.6.1 Uncoupling Procedure for the Solution of a Discrete Sys- tem ..... 65
4.6.2 Solution for a Discrete Riccati Equation of Error Covariance. ..... 66
V. NUMERICAL INTEGRATION OF A STIFF SYSTEM ..... 69
5.1 The Set of Uncoupled Differential Equa- tions ..... 69
5.2 Examples ..... 72
VI. CONCLUSIONS ..... 78
BIBLIOGRAPHY ..... 81
APPENDICES ..... 84
A Matrix Sign Function ..... 84
B Asymptotic Solution of the Riccati Equation ..... 85
FIGURE PAGE
2.l Location of Eigenvaluesfor High-Pass Filter . . ..... 24
2.2 Location of Eigenvalues for Band-Pass Filter ..... 27

## LIST OF TABLES

## TABLE

PAGE

$$
\text { 3.1 Numerical Results for Example } 1 \text {. . . . . . . . } 50
$$

3.2 Numerical Results for Example 2 ..... 51
4.1 Optimal States and Controls for a Discrete
Linear Regulator ..... 66
4.2 Solution Matrix for a Discrete Riccati Equation ..... 68
5.1 Solution Vector $\mathrm{X}_{2}$ for a High-Pass Matrix . . . ..... 73
5.2 Solution Vector $X_{1}(t)$ for a Low-Pass Matrix ..... 75
5.3 Solution Vector $\mathrm{X}_{2}(\mathrm{t})$ for a Band-Pass Matrix ..... 76
5.4 Solution Vector $X_{3}(t)$ for a High-Pass Matrix ..... 76
5.5 Solution Vector $X(t)=X_{1}(t)+X_{2}(t)+X_{3}(t) \cdot \cdot$ ..... 77

## IriRRODUCTION

Considerable research has been carried out in the problem areas of optimal linear giadratic control and optimal filtering [2]. The two main theoretical approaches to the solution of a control optimization problem have been Bellman's maximum principle, which is based on the principle of optimality, and Pontryagin's minimum principle, which is an extension of the classical calculus of variations $[3,7]$. The minimum principle was originally developed for continuous-time problems and recently has been extended to discrete-time problems [32].

The application $0=$ the minimum principle of Pontryagin to the optimal control problem results in a numerical solution of the "two-point boundary value problem." The computational procedure for solving the linear two-point boundary value problems has been the subject of many recent technical reports. The most common technique for the solution of linear optimal control developed to date deals with the algebraic Riccati equation generated from a canonical set of system equations [20,23].

In addition, the appearance of Riccati equations in the optimal linear filter with Gaussian randomness has been reported in many papers. The solution matrix for the error covariance Riccati equation goverrs all the behavior and contains the necessary information for characterizing the optimal filter [9,21, 29,30,39].

The theorem of duality has proved that the error
covariance equation in optimal estimation is closely related to the Riccati equation for an optimal gain in the linear quadratic optimal control [19]. The integration of Riccati equations instead of the basic form of the state transition approach has been favored because of the improved stability properties [15,31,34, 38]. The two most widely-used techniques for solving the Riccati equation are the direct numerical integration and the automatic synthesis program (ASP) matrix integration procedure [22]. Several papers have shown the matrix Riccati equation from an algebraic point of view and determined the asymptotic or steady-state solution to the equation $[27,28,33]$. The results are applied to the spectral factorization of a class of matrices arising in filtering theory and network synthesis [1,28]. A few additional papers have been devoted to the determination of the algebraic expression for the solution of the matrix Riccati equation without iterative procedures [44,45].

During the past year, the matrix sign function has been introduced as an efficient tool in the computation of eigenvalues and eigenvectors for a square matrix and in the asymptotic solution of an algebraic Riccati equation $[4,5,6]$. A new uncoupling algorithm for system differential equations based on the matrix sign function is presented in this paper. Special forms of system equations arising in many classes of system optimization are investigated. Among the special forms considered are the stiff state equations with constant coefficients for which a numerical algorithm is presented. The purpose of this study is twofold:

1) To develop a new uncoupling algorithm to solve a set of differential or difference equations appearing in optimal control and filtering
2) To present a matriz filter algorithm for integration of system equations with constant coefficients whose eigenvalues are wicely spaced

In the development of the second purpose of this study (i.e., to present a new matrix filter algorithm for the integration of system equations with constant coefficients whose eigenvalues are widely spaced), the eigenvalues of the considered matrix are isolated into sabsets in which each eigenvalue has the same requisite characteristic, such as the same sign or the same order of magnitude. For the systems with widely spaced eigenvalues, the conventional explicit integration methods may require excessively small integrȧion steps.

Numerical integration techniques suitable for the solution of stiff state equations have appeared in the literature over the past decade $[40,43]$. One stability concept that is widely used in connection with stiff systems, called A-stability, was introduced early by Dahlquist [12]. Lawson [25] has reported on an example of the A-stable Runge Kutta method which is termed the generalized Runge-Kutta process. Rosenbrook [37] and Calahan [10] have described the implicit Runge-Kutta methods which are similar in many respects $=0$ Larson's and are implicitly analogous to the conventional process.

Certaine [11] has considered the problem of devising a predictor-corrector metho』 for stiff ordinary differential equa-
tions. Gear [17,18] devised a practical automatic integration routine based on Adam's predictor-corrector algorithm. Recently, Klopfensten, et al, [24] developed an algorithm, called PECF, consisting of a predictor-corrector associated with iterations of a pseudo Newton-Raphson method.

Outside the realm of direct consideration in this study is the work of Richards, et al, [36] who reported on an explicit method for large stiff systems associated with the qualitative picture of the solution paths (the direction of eigenvectors corresponding to dominant eigenvalues). A more systematic method, a second-order exact multistep method, for large systems has been developed by Fowler and Waten [16].

Somewhat related to the study developed in this paper is an important area often more limited by the "stiffness" phenomenon in the numerical determination of the transient response of electrical circuits [35,41,42]. Branin [8] has used the eigenvalue method to compute the frequency and time response of the circuit. More closely related to this study is the work of Richard, et al, [35] who proposed a procedure for removing the stiffness of the equations themselves during the course of the solution. The state variable corresponding to the largest pole has been discarded whenever the steady-state excitation of a linear network reach the time constant to that pole. A similar approach by Davison $[13,14]$ is based on finding the dominant poles and zeros of the system and then determining the time solution in terms of those poles and zeros. Also interested in this problem is Lee [26] who has reported a more practical pro-
cedure to integrate such linear time-invariant system equations by means of a conventional matrix filter theory. He proposed that both a low-pass and a high-pass filter be constructed, and that these filters then be used to uncouple the system equations into two sets, a set with large eiqenvalues and one with small eigenvalues. The development herein is related to his work. Numerical techniques in the optimization of a linear quadratic system and the Kalman filter are reviewed in Chapter I along with a brief historical background. In addition, certain recent theoretical developments in the area of numerical integration techniques for stiff orainary differential equations are presented.

New algorithms are developed in Chapter II to uncouple the particular forms of system matrices arising in many important classes of optimal control and estimation problems. Examples using the new algorithm in both time domains, continuous and discrete, are presented. Darticular emphasis is placed on developing a new technique that croups the system eigenvalues into separated subsets and generates completely uncoupled filter matrices.

Chapter III includes an outline of existing methods for the solution of a continuous-time linear regulator with quadratic criteria. A new computational procedure is then developed using the uncoupling algorithm. This new procedure employs a forwarding integration scheme and does not require the storage of feedback gains for a complete solution.

In Chapter IV, the uncoupling algorithm is extended to
the discrete time case, and uncoupled discrete linear system equations are studied. The application of the algorithm to solve a discrete regulator problem is shown, along with a discussion of the variance equation for the Kalman filter. The derivation of a non-recursive expression for a solution matrix of a discrete Riccati equation is included.

Chapter $V$ discusses an integration procedure for a linear stiff equation derived from the numerical determination of the transient response of electrical circuits. The merits of the band-pass filter matrix are demonstrated.

Chapter VI concludes with remarks concerning the use of the method discussed in this paper and suggestions for possible further study.

Each chapter contains example solutions using the new method to illustrate the outlined procedures and the numerical accuracy. The mathematical concept of the matrix sign function and the derivation of the steady-state solution of the matrix Riccati equation will be presented in the Appendices.

## DEVELOPMENT OF ALGORITHMS

### 2.1 Generation of the Transformation Matrix

A recently developed similarity transformation matrix
$[4,5]$ has been constructed from a matrix sign function, and, as will be shown, reduces a matrix to a block diagonal form under proper operations.

In the generation of the matrix, if $\bar{A}$ is of the order $\mathrm{n} \times \mathrm{n}$ and \& eigenvalues have positive real parts while m eigenvalues have negative real parts so that $\ell+m=N$, then the matrix sign function sign $(\bar{m})$ is obtained by

$$
\operatorname{sign}(\bar{A})=\left[\begin{array}{cc}
M_{11} & M_{12}  \tag{2.1}\\
M_{21} & M_{22}
\end{array}\right]\left[\begin{array}{ll}
I_{2 \times 2} & 0 \\
0 & -I_{m \times m}
\end{array}\right]\left[\begin{array}{cc}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1}
$$

where $M_{i j}$ are partitions of the eigenvector matrix $M$ and $I_{\ell \times \ell}$ is an $\ell \times \ell$ identity matrix. The dimensions are then compatible for matrix operations. A transformation matrix $V$ of the order $n \times n$ can now be defined as follows:

$$
\begin{equation*}
V=\operatorname{sign}(\bar{A})+K \tag{2.2}
\end{equation*}
$$

where $K$ is the matrix of $\pm I$ in the diagonal form shown in (2.l). The matrix $K$ can be formed by computing the trace of sign $\bar{A}$. Substituting (2.1) into (2.2) and using simple matrix operations,

$$
V=\left[\begin{array}{cc}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]\left[\begin{array}{cc}
I_{2 \times 2} & 0 \\
0 & -I_{m \times m}
\end{array}\right]\left[\begin{array}{cc}
{ }^{M} 11 & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1}
$$

$$
+\left[\begin{array}{ll}
I_{\ell \times \ell} & 0  \tag{2.2a}\\
0 & -I_{m \times m}
\end{array}\right]\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1}
$$

which gives

$$
\begin{align*}
V & =\left[\left[\begin{array}{ll}
M_{11} & -M_{12} \\
M_{21} & -M_{22}
\end{array}\right]+\left[\begin{array}{cc}
M_{11} & M_{12} \\
-M_{21} & -M_{22}
\end{array}\right]\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1}\right. \\
& =2\left[\begin{array}{ll}
M_{11} & 0 \\
0 & -M_{22}
\end{array}\right]\left[\begin{array}{cc}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1} \tag{2.2b}
\end{align*}
$$

Matrix $V$ in (2.2b) is the general form of the transformation matrix and may be applied to any particular form of square matrix to obtain a block diagonal form.

### 2.2 Block Diagonal Form of a General Square Matrix

A square matrix $\bar{A}$ is assumed where

$$
\bar{A}=\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]\left[\begin{array}{ll}
J_{1} & 0 \\
0 & J_{2}
\end{array}\right]\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1}=\left[\begin{array}{ll}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{array}\right] \text { (2.3) }
$$

In this development, the submatrices $J_{1}$ and $J_{2}$ are Jordan blocks or are made up of the eigenvalues of $\bar{A}$.

The matrix in (2.2b) is then used to perform a similarity transformation on $\bar{A}$ to find a new matrix $\bar{A}_{1}$ as follows:

$$
\bar{A}_{1}=V^{-1} \bar{A} V=M\left[\begin{array}{cc}
M_{1}^{-1} & 0  \tag{2.4}\\
0 & -M_{22}^{-1}
\end{array}\right]\left[\begin{array}{cc}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{array}\right]\left[\begin{array}{cc}
M_{11} & 0 \\
0 & -M_{22}
\end{array}\right] M^{-1}
$$

$$
\begin{aligned}
& \text { (2.4a) }
\end{aligned}
$$

In the above development, the submatrices of $M^{-1}, S_{i j}$ are given by

In considering the (1.2) and (2.1) submatrices of (2.4a), the (2.1) submatrix is zero, as shown below:

$$
\begin{align*}
& M_{21} M_{11}^{-1} A_{11} M_{11}\left({ }_{1} 11^{-M} 122^{M_{2}^{-1}} 2_{21}\right)^{-1}-M_{21} 1_{11}^{-1} A_{12} M_{22}\left(M_{12}-M_{11} M_{21}^{-1} M_{22}\right)^{-1} \\
& -A_{21} M_{11}\left(M_{11}-M_{12} M_{22}^{-M_{21}}\right)^{-1}+A_{22} M_{22}\left(M_{12}-M_{11} M_{21}^{-1} M_{22}\right)^{-1} \\
& =\left(M_{21} M_{11}^{-1} A_{11}+Y_{21} M_{11}^{-1} A_{12} M_{21} M_{11}^{-1}-A_{21}-A_{22} M_{21} M_{11}^{-1}\right) \\
& \text { - }\left(I-M_{12} M_{22}^{-1} M_{2 I} M_{1 I}^{-1}\right)^{-1}=0 \tag{2.5}
\end{align*}
$$

since the bracketed term is the asymptotic solution to the Riccati equation. The (1.2) submatrix is also zero since the bracketed term in (2.5a) below is an asymptotic solution to the matrix Riccati equation. (see Appendix)

$$
\begin{align*}
& A_{11} M_{11}\left(M_{21}{ }^{-M_{2}} 2_{12}^{-1} M_{11}\right)^{-1}-A_{12} M_{22}\left(M_{22}-M_{21} \dot{M}_{11}^{-1} M_{12}\right)^{-1} \\
& -M_{12}{ }^{M_{2}-1} A_{21} M_{11}\left(M_{21}-\mathrm{M}_{22} \mathrm{M}_{12}^{-1} \mathrm{M}_{11}\right)^{-1} \\
& +M_{12} M_{22^{-1}}{ }_{22^{M}} 22\left(M_{22}-M_{21} M_{11}^{-1} M_{12}\right)^{-1} \\
& =\left(\mathrm{A}_{11} \mathrm{M}_{12} 2^{\mathrm{M}_{22}^{-1}+\mathrm{A}_{12}}{ }^{-\mathrm{M}_{12}}{ }^{\mathrm{M}_{2}^{-1}} 2^{\mathrm{A}} 21^{\mathrm{M}_{1}} 12^{\mathrm{M}_{22}^{-1}-\mathrm{M}_{12} \mathrm{M}_{22}^{-1} \mathrm{~A}_{22}}\right) \\
& \text { - }\left(I-M_{21} M_{11}^{-1} M_{12} M_{22}^{-1}\right)^{-1}=0 \tag{2.5a}
\end{align*}
$$

Rearranging these terms and using the definition of $J_{1}$ and $J_{2}(2.3)$, it can be shown that the submatrices along the diagonal are as given below:
$\operatorname{block}(1.1)=\left(M_{11}-M_{12} M_{21}^{-1} M_{21}\right) J_{1}\left(M_{11}-M_{12} M_{22}^{-1} M_{21}\right)^{-1}$
$\operatorname{block}(2.2)=\left(M_{22}-M_{21} M_{11}^{-1} M_{12}\right) J_{2}\left(M_{22}-M_{21} M_{11}^{-1} M_{12}\right)^{-1}$

The matrix $A_{1}$ is then of the form

$$
\bar{A}_{1}=\left[\begin{array}{ll}
\bar{A}_{11} & 0  \tag{2.6}\\
0 & \bar{A}_{22}
\end{array}\right]=\left[\begin{array}{ll}
\left(M_{11}-M_{12} M_{22}^{-1} M_{21}\right) J_{1}\left(M_{11} M_{12} M_{22}^{-1} M_{21}\right)^{-1} & 0 \\
0 & \left(M_{22^{-M}} M_{21} M_{11}^{-1} M_{12}\right) \\
J_{2}\left(M_{22}-M_{21} M_{11}^{-1} M_{12}\right)^{-1}
\end{array}\right]
$$

which breaks $\bar{A}$ into submatrices of a block diagonal lower order so that the eigenvalues of $\bar{A}_{1}$ are the same as those of $\bar{A}$.

On the other hand, it can be shown that the similarity transformation $V \overline{A N}^{-1}$ gives a different block diagonal form.

$$
V \overline{A V}^{-1}=\left[\begin{array}{ll}
M_{11} & 0  \tag{2.7}\\
0 & -M_{22}
\end{array}\right] M^{-1} \quad \text { A } \quad M \quad\left[\begin{array}{cc}
M_{11}^{-1} & 0 \\
0 & -M_{22}^{-1}
\end{array}\right]
$$

From Equation (2.3),

$$
\begin{align*}
V \overline{A V}^{-1} & =\left[\begin{array}{ll}
M_{11} & 0 \\
0 & -\mathrm{N}_{22}
\end{array}\right]\left[\begin{array}{cc}
J_{1} & 0 \\
0 & J_{2}
\end{array}\right]\left[\begin{array}{cc}
M_{11}^{-1} & 0 \\
0 & -M_{22}^{-1}
\end{array}\right] \\
& =\left[\begin{array}{ll}
M_{11} J_{1} M_{11}^{-1} & 0 \\
0 & M_{22 J_{2} M_{22}^{-1}}
\end{array}\right] \tag{2.7a}
\end{align*}
$$

Thus, either $V^{-l} \bar{A} V$ or $V \overline{A V}{ }^{-1}$ may be used for the block diagonalization of $\bar{A}$.

### 2.3 Continuous Linear Systams

As will be discussed in Chapter III, a solution is required for a two-point boundary-value problem. This problem is encountered in the optimization of linear control with a quadratic performance index and a nonlinear Riccati equation for the optimal estimation of a continuous Kalman filter. In both cases, a system matrix is formed in a partitioned square matrix given by (2.8).

$$
\mathrm{H}_{\mathrm{CR}}=\left[\begin{array}{ll}
\mathrm{A}_{11} & \mathrm{~A}_{12}  \tag{2.8}\\
\mathrm{~A}_{21} & \mathrm{~A}_{22}
\end{array}\right]
$$

The elements of $A_{i j}$ are the coefficient matrices of the system equations and have these properties in common: $A_{11}=-A_{22}$, $A_{12}=A_{12}^{T}$, and $A_{21}=A_{21}^{T}$. The matrix $H_{C R}$ is of the order thus, the submatrices are of the order $n / 2 \times n / 2$.

The transformation matrix $V$ may be written by the substitution of (2.4b) into (2.2b) as follows:

$$
\mathrm{V}=2\left[\begin{array}{ll}
\mathrm{S} & -\mathrm{T}_{2} \mathrm{~S}^{-1}  \tag{2.9}\\
\mathrm{~T}_{1} \mathrm{~S} & -\mathrm{S}^{\mathrm{T}}
\end{array}\right]
$$

where $T_{1}=M_{21} M_{11}^{-1}, T_{2}=M_{12} M_{22}^{-1}$ and $S=\left(I-T_{2} T_{1}\right)^{-1}$.
The first column of (2.9) is postmultiplied by $\mathrm{s}^{-1}$ to
construct a new transformation matrix. The transformation matrix $\bar{V}$ for the uncoupling of continuous systems is developed as follows:

$$
\overline{\mathrm{V}}=2\left[\begin{array}{ll}
\mathrm{I} & -\mathrm{T}_{2} \mathrm{~S}^{\mathrm{T}}  \tag{2.9a}\\
\mathrm{~T}_{1} & -\mathrm{S}^{\mathrm{T}}
\end{array}\right]
$$

where the inverse of $\overline{\mathrm{V}}$ is

$$
\overline{\mathrm{V}}^{-1}=\frac{1}{2}\left[\begin{array}{ll}
\mathrm{S} & -\mathrm{ST} \mathrm{~T}_{2}  \tag{2.9b}\\
\mathrm{~T}_{1} & -\mathrm{I}
\end{array}\right]
$$

$V_{1}$ is the desired transformation matrix for the block diagonalization process of the system matrix given by (2.8). If a similarity transformation is performed on $H_{C R}$ by $\bar{V}$, then

$$
\begin{align*}
& \bar{H}_{C R}=\bar{V}^{-1} H_{C R} \bar{V}=\left[\begin{array}{ll}
S & -S T_{2} \\
T_{1} & -I
\end{array}\right]\left[\begin{array}{ll}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{array}\right]\left[\begin{array}{ll}
I & -T_{2} S^{T} \\
T_{1} & -S^{T}
\end{array}\right] \\
& =\left[\begin{array}{l}
\mathrm{S}^{\left(\mathrm{A}_{11}\right.}{ }^{-\mathrm{T}} 2^{\mathrm{A}} 21{ }^{+\mathrm{A}} 12^{\mathrm{T}} 1^{\left.-\mathrm{T}_{2} \mathrm{~A}_{22} \mathrm{~T}_{1}\right)} \\
\mathrm{T}_{1} \mathrm{~A}_{11} 1^{-\mathrm{A}_{21}+\mathrm{T}_{1} \mathrm{~A}_{12} \mathrm{~T}_{1}{ }^{-\mathrm{A}} 22^{\mathrm{T}} 1}
\end{array}\right. \\
& \left.\begin{array}{l}
-S\left(A_{11} T_{2}-\mathrm{T}_{2} \mathrm{~A}_{21} \mathrm{~T}_{2}+\mathrm{A}_{12}-\mathrm{T}_{2} \mathrm{~A}_{22}\right) \mathrm{S}^{\mathrm{T}} \\
\left(-\mathrm{T}_{1} \mathrm{~A}_{11} 1_{2}+\mathrm{A}_{21} \mathrm{~T}_{2}-\mathrm{T}_{1} \mathrm{~A}_{12}+\mathrm{A}_{22}\right) \mathrm{S}^{\mathrm{T}}
\end{array}\right] \tag{2.10}
\end{align*}
$$

since $T_{1}$ and $T_{2}$ are the asymptotic solutions to the matrix Riccati
equations, the (1.2) and (2.1) subblocks are zero. The details on asymptotic solutions of Riccati equations will be discussed in Chapters III and IV.

$$
\overline{\mathrm{H}}_{\mathrm{CR}} \text { now becomes an uncoupled form }
$$

$$
\overline{\mathrm{H}}_{\mathrm{CR}}=\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Cl}} & 0  \tag{2.10a}\\
0 & \mathrm{H}_{\mathrm{C}}
\end{array}\right]
$$

and taking the transpose of subblock $H_{C 2}$ yields

$$
\begin{equation*}
\mathrm{H}_{\mathrm{C} 2}^{\mathrm{T}}=\mathrm{S}\left(-\mathrm{T}_{2}^{\mathrm{T}} \mathrm{~A}_{11}{ }^{\mathrm{T}} \mathrm{~T}_{1}{ }^{\mathrm{T}}+\mathrm{T}_{2}{ }^{\mathrm{T}} \mathrm{~A}_{21}{ }^{\mathrm{T}}-\mathrm{A}_{12} \mathrm{~T}_{\mathrm{T}_{1}}^{\mathrm{T}}+\mathrm{A}_{22}{ }^{\mathrm{T}}\right) \tag{2.10b}
\end{equation*}
$$

In considering that $A_{11}=-A_{22}{ }^{T}, A_{12}=A_{12}{ }^{T}, A_{21}=A_{21}{ }^{T}, T_{1}=T_{1}{ }^{T}$ and $T_{2}=T_{2}{ }^{T}$, a relation is obtained between submatrices $H_{C 1}$ and $\mathrm{H}_{\mathrm{C} 2}$ so that $\mathrm{H}_{\mathrm{Cl}}=-{ }^{-F} \mathrm{C}_{2}$. This particular form of $\mathrm{H}_{\mathrm{CR}}$ appears in the solution of various optimal controls and estimation problems.

A considerable reduction in computational efforts may be obtained by the use of the new transformation matrix given by (2.9a), since the resulting submatrices are of the order $n / 2 \times n / 2$ rather than $n \times n$. The computational procedures of an uncoupling algorithm for a continnous linear system are summarized in the following steps:
a) Obtain the state and costate differential equations and construct $\mathrm{H}_{\mathrm{CR}}$.
b) Compute the sign function matrix of $\mathrm{H}_{\mathrm{CR}}$ and construct the transformation matrices $V$ and $\bar{V}$ in (2.9) and (2.9a).
c) Apply a similarity transformation to the system matrix $H_{C R}$ by $\overline{\mathrm{V}}$.
2.4 Discrete Linear Systems

The canonical difference equations obtained for the optimization of linear discrete systems can be interpreted as either forward or backward integration schemes. As will be shown in Chapter IV, the corresponding system matrices for both cases are

$$
{ }_{\mathrm{H}}^{\mathrm{DR}} \text { (forward) }=\left[\begin{array}{lc}
\phi+\mathrm{BR}^{-1} B^{T} \phi^{-T} Q & -\mathrm{BR}^{-1_{B} \phi^{-T}}  \tag{2.11}\\
-\phi^{-\mathrm{T}_{Q}} & \phi^{-\mathrm{T}}
\end{array}\right]
$$

and

$$
H_{D R}(\text { backward })=\left[\begin{array}{ll}
\phi^{-1} & \phi^{-1} B R^{-1} B^{T}  \tag{2.11a}\\
Q \phi^{-1} & \phi^{T}+Q \phi^{-1} B_{B R}-1_{B} T
\end{array}\right]
$$

where $\mathrm{H}_{\mathrm{DR}}$ denotes the system matrix of a discrete regulator. The dual nature of optimal estimation problems and optimal regulator problems provides equivalent relations between coefficient matrices of both systems. The system matrix for the forward canonical equations for the optimal filter will have the following general form:

The sign algorithm is useful for computing the block diagonalization form of the discrete forms. Since the sign al-
gorithm separates the eigenvalues by the sign of the real parts, the discrete canonical matrix must be mapped into the continuous domain. The bilinear transformation known in sampled data theory may be considered as

$$
\begin{equation*}
z=\frac{1+S}{1-S} \tag{2.12}
\end{equation*}
$$

where $\omega$ is a complex variable; that is, $S=\sigma+j \omega$.
The transformation in (2.12) maps the interior of the unit circle in the $z$-plane onto the left half of the $\omega$-plane. Similarly, the bilinear transformation given by (2.12a) also maps the inside of the $z$-olane unit circle into the left half of the $\omega$-plane.

$$
\begin{equation*}
Z=\frac{S+1}{S^{-1}} \tag{2.12a}
\end{equation*}
$$

Either transformation will convert the original characteristic equation in $Z$ into a quotient of the polynomial in $S$ of the same order. Therefore, the domain transformation can be interpreted in terms of system eigenvalues in both planes.
Using either (2.12) or (2.12a),

$$
\begin{equation*}
\Lambda_{C}=\left(\Lambda_{D}-I\right)\left(\Lambda_{D}+I\right)^{-1} \tag{2.13}
\end{equation*}
$$

where $\Lambda_{D}$ and $\Lambda_{C}$ denote the Jordan forms of the system matrix in the discrete and complex planes, respectively. In solving (2.13) for $\Lambda_{D}$, the following is obtained:

$$
\begin{equation*}
\Lambda_{D}=\left(I-\Lambda_{C}\right)^{-1}\left(I+!_{C}\right) \tag{2.13a}
\end{equation*}
$$

It has been shown that the eigenvalues of $H_{D R}$ and $H_{D F}$ given by (2.10) through (2.10a) occur in reciprocal pairs so that for an eigenvalue $\lambda_{i}$, there exists an eigenvalue $\lambda_{j}=\frac{1}{\lambda_{i}}$ [45]. In other words, one half of the eigenvalues of the discrete system matrix lies inside the unit circle in discrete domain, while the other half lies outside. Hence, $\Lambda_{D}$ can be written in two diagonal blocks as

$$
\Lambda_{\mathrm{D}}=\left[\begin{array}{cc}
\Lambda_{\mathrm{D} 1} & 0  \tag{2.14}\\
0 & \Lambda_{\mathrm{DI}}^{-1}
\end{array}\right]
$$

Substituting (2.14) into (2.13) yields

$$
\Lambda_{\mathrm{C}}=\left[\begin{array}{ll}
\Lambda_{\mathrm{DI}}-I & 0  \tag{2.15}\\
0 & \Lambda_{\mathrm{DI}}^{-I}-I
\end{array}\right]\left[\begin{array}{ll}
\Lambda_{\mathrm{DI}}+I & 0 \\
0 & \Lambda_{\mathrm{DI}}^{-I}+I
\end{array}\right]
$$

which is

$$
\Lambda_{C}=\left[\begin{array}{ll}
\Lambda_{C 1} & 0  \tag{2.15a}\\
0 & \Lambda_{C 2}
\end{array}\right]=\left[\begin{array}{ll}
\left(\Lambda_{\mathrm{D} 1}-I\right)\left(\Lambda_{\mathrm{D} 1}+I\right)^{-1} & \\
0 & -\left(\Lambda_{\mathrm{DI}}-I\right)\left(\Lambda_{\mathrm{D} 1}+I\right)^{-1}
\end{array}\right]
$$

From (2.15a), it can be concluded that the eigenvalues of a discrete system matrix are transformed in such a manner that $\Lambda_{\mathrm{C} 1}=-\Lambda_{\mathrm{C} 2}$.

If $M$ is the eigenvector matrix of the discrete system matrix and both sides of (2.13a) are multiplied by $M$ and $M^{-1}$,

$$
\begin{equation*}
H_{D R}=M \Lambda_{D} M^{-1}=M\left(I-\Lambda_{C}\right)^{-1}\left(I+\Lambda_{C}\right) M^{-1} \tag{2.16}
\end{equation*}
$$

Then solving (2.16) for the term $H_{C R}=M \Lambda_{C} M^{-1}$ gives

$$
\begin{equation*}
H_{C R}=M \Lambda_{C} M^{-1}=\left(H_{D R}+I\right)^{-1}\left(H_{D R}-I\right) \tag{2.17}
\end{equation*}
$$

where ${ }^{H}{ }_{C R}$ denotes a new system matrix transformed into the complex plane whose eigenvalues are symmetric about the imaginary axis.

$$
\begin{align*}
& \text { The matrix sign function for } H_{C R} \text { is defined by } \\
& \text { sign } H_{C R}=\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & \because_{22}
\end{array}\right]\left[\begin{array}{cc}
I & 0 \\
0 & -I
\end{array}\right]\left[\begin{array}{ll}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array}\right]^{-1} \tag{2.18}
\end{align*}
$$

and the transformation matrix V is obtained as follows:

$$
\mathrm{V}=\operatorname{sign} \mathrm{H}_{\mathrm{CR}}+\left[\begin{array}{cc}
\mathrm{I} & 0  \tag{2.19}\\
0 & -\mathrm{I}
\end{array}\right]=2\left[\begin{array}{cc}
\mathrm{S} & -\mathrm{T}_{2} \mathrm{~S}^{\mathrm{T}} \\
\mathrm{~T}_{1} \mathrm{~S} & -\mathrm{S}^{T}
\end{array}\right]
$$

where $T_{1}=M_{21} M_{11}^{-1}, T_{2}=M_{12} 2^{M_{2}^{-1}}$ and $S=\left(I-T_{2} T_{1}\right)^{-1}$.
Again, an alterna亡ed form of matrix $\overline{\mathrm{V}}$ may be reconstructed by postmultiplying the first column of V by $\mathrm{S}^{-1}$. The desir-ed transformation matrix in the complex plane is

$$
\overline{\mathrm{V}}=2\left[\begin{array}{ll}
\mathrm{I} & -\mathrm{T}_{2} \mathrm{~S}^{\mathrm{T}}  \tag{2.19a}\\
\mathrm{~T}_{1} & -\mathrm{S}^{\mathrm{T}}
\end{array}\right]
$$

The uncoupling procedure for the new system after transformation is straightforward, using the same algorithm developed in previous sections. If ${ }^{F}$ CR is the block diagonalized form of ${ }^{H}{ }_{C R}$, and the subblocks are

$$
\overline{\mathrm{H}}_{\mathrm{CR}}=\overline{\mathrm{V}}^{-1}{ }_{\mathrm{H}}^{\mathrm{CR}} \overline{\mathrm{~V}}=\left[\begin{array}{ll}
\mathrm{H} \mathrm{Cl} & 0  \tag{2.20}\\
0 & { }_{\mathrm{H}}^{\mathrm{C} 2}
\end{array}\right]
$$

then $H_{C 1}$ and $H_{C 2}$ have the same dimensions since $I$ and $-I$ in (2.18) are in the same order. From (2.20) $\mathrm{H}_{\mathrm{CR}}$ in terms of $\bar{H}_{C R}$ is

$$
\begin{equation*}
\mathrm{H}_{\mathrm{CR}}=\overline{\mathrm{V}} \overline{\mathrm{H}}_{\mathrm{CR}} \overline{\mathrm{~V}}^{-1} \tag{2.21}
\end{equation*}
$$

Substituting (2.21) for (2.17) results in

$$
\begin{equation*}
H_{D R}=\left[I-\overline{\mathrm{V}} \overline{\mathrm{H}}_{\mathrm{CR}} \overline{\mathrm{~V}}^{-1}\right]^{-1}\left[I+\overline{\mathrm{V}} \overline{\mathrm{H}}_{\mathrm{CR}} \overline{\mathrm{~V}}^{-1}\right] \tag{2.22}
\end{equation*}
$$

from which

$$
\mathrm{H}_{\mathrm{DR}}=\overline{\mathrm{V}} \bar{H}_{\mathrm{DR}} \overline{\mathrm{~V}}^{-1}=\overline{\mathrm{V}}\left[\begin{array}{ll}
\mathrm{H}_{\mathrm{D} 1} & 0  \tag{2.22a}\\
0 & \mathrm{H}_{\mathrm{D} 2}
\end{array}\right] \overline{\mathrm{V}}^{-1}
$$

where

$$
\begin{align*}
\bar{H}_{D R} & =\left[I-\bar{H}_{C R}\right]^{-I}\left[I+\bar{H}_{C R}\right] \\
& =\left[\begin{array}{cc}
\left(I-\mathrm{H}_{C 1}\right)^{-I}\left(I+\mathrm{H}_{\mathrm{C} 1}\right) & 0 \\
0 & \left(I-H_{C 2}\right)^{-1}\left(I+\mathrm{H}_{\mathrm{C} 2}\right)
\end{array}\right] \tag{2.22b}
\end{align*}
$$

The block diagonalization of the discrete systern matrix is completed by applying a similarity transformation directly to $H_{D R}$; i.e.,

$$
\begin{equation*}
\overline{\mathrm{H}}_{\mathrm{DR}}=\overline{\mathrm{V}}^{-1} \mathrm{H}_{\mathrm{DR}} \overline{\mathrm{~V}} \tag{2.22c}
\end{equation*}
$$

The system matrix for the discrete optimal filter given by (2.11a) is chosen to investigate the relation between subblocks $H_{D 1}$ and $H_{D 2}$. From (2.11a) and (2.19a), $\bar{H}_{D R}$ becomes

$$
\begin{aligned}
& \cdot\left[\begin{array}{ll}
I & -T_{2} S^{T} \\
T_{1} & -S^{T}
\end{array}\right]
\end{aligned}
$$

$$
\begin{aligned}
& \text { (2.23) }
\end{aligned}
$$

In the (1.2) and (2.1) subblocks of (2.23), the bracketed term of the (1.2) submatrix is zero since $\mathrm{T}_{2}$ is the asymptotic solution to the Riccati equation whose system matrix can be represented by

$$
H_{D R}=\left[\begin{array}{ll}
\phi^{-1} & \phi^{-1} B R^{-1} B^{T}  \tag{2.24}\\
Q \phi^{-1} & \phi^{-1}+Q 亡^{-1} B_{B}^{-1} B^{T}
\end{array}\right]
$$

The (2.1) submatriy is also zero, and the system matrix is

$$
H_{D R}=\left[\begin{array}{ll}
\phi^{T}+Q \phi^{-1} B_{B R^{-1}} B^{T} & Q \phi^{-1}  \tag{2.24a}\\
\phi^{-1} B_{B R}^{-1} B^{T} & \phi^{-1}
\end{array}\right]
$$

The same procedures as above are applied to $H_{D R}^{-1}$ to determine the relation between the diagonal terms. The inverse matrix of $H_{D R}$ in partitioned form is given as

$$
\mathrm{H}_{\mathrm{DR}}^{-1}=\left[\begin{array}{cc}
\phi+\mathrm{BR}^{-1} \mathrm{~B}^{\mathrm{T}} \dot{C}^{-\mathrm{T}} \mathrm{Q} & -\mathrm{BR}^{-1} \mathrm{~B}^{T} \phi^{-\mathrm{T}}  \tag{2.25}\\
-\phi^{-\mathrm{T}_{Q}} & \phi^{-T}
\end{array}\right]
$$

Applying the similarity Eransformation to $H_{D R}^{-1}$ by $\bar{V}$ will result in

$$
\begin{aligned}
& \bar{V}^{-I_{H}-1} \overline{D R}=\left[\begin{array}{ll}
S & -S^{m} \\
T_{1} & -I
\end{array}\right]\left[\begin{array}{cc}
\phi+B^{-1} B_{B} \phi^{-T} Q & -B R^{-1} B^{T} \phi^{-T} \\
-\phi_{Q} & \phi^{-T}
\end{array}\right]\left[\begin{array}{ll}
I & -T_{2} S^{T} \\
T_{1} & -S^{T}
\end{array}\right]
\end{aligned}
$$

The submatrices (1.2) ana (2.1) of (2.26) also vanish for the same reason as discussed previously.

In taking the irverse of both sides of (2.22a), the following is obtained:

$$
\mathrm{H}_{\mathrm{DR}}^{-1}=\overline{\mathrm{V}} \overline{\mathrm{H}}_{\mathrm{DR}}^{-1} \overline{\mathrm{~V}}^{-1}=\overline{\mathrm{V}}\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{D} 1}^{-1} & 0  \tag{2.27}\\
0 & \mathrm{H}_{\mathrm{D} 2}^{-1}
\end{array}\right] \overline{\mathrm{V}}^{-1}
$$

Equating (2.26) and (2.27) and taking the transpose of the (1.1) and (2.2) subblocks yielcs

$$
\begin{equation*}
\mathrm{H}_{\mathrm{DI}}^{-\mathrm{T}}=\left(\phi^{\mathrm{T}}+Q^{-1} \mathrm{BR}^{-1} 3^{\mathrm{T}}+\Omega_{4}^{-1} \mathrm{~T}_{2}^{2 \mathrm{~T}}-\mathrm{T}_{1}^{\mathrm{T}} \phi^{-1} \mathrm{BR}^{-1} \mathrm{~B}^{\mathrm{T}}-\mathrm{T}_{1}^{\mathrm{T}} \phi^{-1} \mathrm{~T}_{2}^{\mathrm{T}}\right) \mathrm{S}^{\mathrm{T}} \tag{2.28}
\end{equation*}
$$

and

$$
\begin{equation*}
H_{D 2}^{-T}=S\left(-T_{2}^{T} \phi^{T} T_{1}^{T}-T_{2}^{T} \phi^{-1} B_{R}^{-1} B^{T} T_{1}^{T}-T_{2}^{T} Q \phi^{-1}+\phi^{-1} B R^{-1} B_{T_{1}^{T}}^{T}+\phi^{-1}\right) \tag{2.28a}
\end{equation*}
$$

Since the asymptotic solutions to Riccati equations $T_{1}$ and $T_{2}$ are symmetric matrices, it can be determined from (2.23), (2.28), and (2.28a) that

$$
\begin{equation*}
\mathrm{H}_{\mathrm{DI}}=\mathrm{H}_{\mathrm{D} 2}^{-\mathrm{T}} \tag{2.29}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathrm{H}_{\mathrm{D} 2}=\mathrm{H}_{\mathrm{D} 1}^{-\mathrm{T}} \tag{2.29a}
\end{equation*}
$$

It has been proven that the backward system matrix for a discrete linear regulator can be reduced to subblocks which hold the relations in (2.29) and (2.29a). The same procedure, when applied to the matrices in (2.11) and (2.11b), will uncouple the systems.

The algorithm for uncoupling the canonical equations for the discrete system is summarized in the following four steps:
a) Obtain the state and costate difference equations to construct $H_{D R}$ and $H_{D F}$.
b) Use the bilinear transformation in (2.17) to transform the discrete system matrices into a complex plane.
c) Obtain the transformation matrix in (2.19) by the use of matrix sign algorithm and then construct $\overline{\mathrm{V}}$ in (2.19a).
d) Apply a similarity transformation directly to the discrete system matrices.

### 2.5 Generation of a Filter Matrix

A technique is developed here to generate the matrix filters by means of the sign of a matrix discussed previously. The conventional filter theory requires approximate values of the eigenvalues, and, in most cases, the eigenvalues are restricted to lie along one of the coordinate axes. [26].

In addition, the filters are not ideal, and even a weak coupling may seriously disrupt nunerical calculations, accumulating errors if the filter matrices are not of sufficiently high order. In the new algorithm, the eicenvalues of a matrix are isolated into subsets in which each eigenvalue has similar characteristics, such as the same sign or the same order of magnitude. The filters constructed from the sign function axe of infinite order, and there is no interaction between the eigenvalues even though closely spaced eigenvalues are separated into adjacent filters. If it is assumed that $A$ is an $n \times n$ system matrix of differential equations to be integrated and has the Jordan form $\Lambda=\operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}\right)$ with a corresponding eigenvector matrix M , the sign matrix of A is then given by

$$
\operatorname{sign} A=M K M^{-1}
$$

where

$$
K=\operatorname{diag}\left[\operatorname{sign}\left(R_{e} \lambda_{1}\right), \operatorname{sign}\left(R_{e} \lambda_{2}\right), \ldots\right]
$$

2.5.1 Filter Matrix for Eigenvalues with the Same Sign In the case where $i_{1}=\operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{S}\right)$, a block for eigenvalues with positive real parts, and $\Lambda_{2}=\operatorname{diag}\left(\lambda_{s+1}, \ldots, \lambda_{n}\right)$
for negative real parts, the sign of $A$ is then given as

$$
\operatorname{sign} A=M\left[\begin{array}{ll}
I_{s \times s} & 0  \tag{2.30}\\
0 & -I_{n-s \times n-s}
\end{array}\right] \quad M^{-1}
$$

The eigenvalues with positive real parts can now be isolated. The first step in the procedure is to add the identity matrix to sign $A$ which gives

$$
S_{p}=\operatorname{sign} A+I=2 \because\left[\begin{array}{ll}
I_{S \times s} & 0  \tag{2.31}\\
0 & 0
\end{array}\right] \mathrm{M}^{-1}
$$

The next step is to multiply $S$ by the matrix $S_{p}$ to form $A^{+}$, the desired matrix,

$$
A^{+}=\frac{1}{2} S_{p} A=\frac{1}{2} A S_{Y}=\because\left[\begin{array}{cc}
i_{1} & 0  \tag{2.32}\\
0 & 0
\end{array}\right] \quad M^{-1}
$$

Since $A^{-}$must be the complement of $A^{+}$, the eigenvalues with negative real parts are then isolated by subtracting $A^{+}$from $A$.

$$
A^{-}=A-A^{+}=S_{N N} A=A S_{Y N}=M\left[\begin{array}{ll}
0 & 0  \tag{2.32}\\
0 & \Lambda_{2}
\end{array}\right] M^{-1}
$$

where

$$
\begin{equation*}
S_{N}=I-S_{p}=\frac{1}{2}(I-\operatorname{sign} A) \tag{2.32a}
\end{equation*}
$$

Thus, two fundamental filter natrices, $S_{p}$ and $S_{N}$, have been generated which pass only positive and negative eigenvalues of $A$, respectively.
2.5.2 Filter Matrix for the Same Order of Magnitude

All eigenvalues of A may have negative real parts with $\left|R_{e} \lambda_{1}\right|<\left|R_{e} \lambda_{2}\right|<\ldots<\left|R_{e} \lambda_{n}\right|$ and be as shown in Figure 2.1. All complex eigenvalues are assumed to occur in conjugate pairs. It may also be assumed the matrix $A$ can be separated into subblocks so that the corresponding subsets of eigenvalues, $\Lambda_{1}$ and $\Lambda_{2}$, where $\Lambda_{1}=\operatorname{diag}\left(\lambda_{1}, \ldots, \lambda_{s}\right)$ and $\Lambda_{2}=\operatorname{diag}\left(\lambda_{s+1}, \ldots, \lambda_{n}\right)$ contain eigenvalues of the same order of magnitude.

Since the trace of $A$ is equal to the sum of the eiger:values, the mean eigenvalues of any matrix $A$ is given by

$$
\begin{equation*}
L=\frac{1}{n} \operatorname{trace} A=\frac{1}{n} \sum_{K=1}^{n} \lambda_{K} \tag{2.33}
\end{equation*}
$$

If $I$ is assumed to be in the range of $\lambda_{S}>L>\lambda_{S+1}$, a simple shift of the origin of the eigenvalues space can be made so that the subset $\Lambda_{1}$ lies in the positive domain and $\Lambda_{2}$ in the negative domain.


Figure 2.1

If $L$ is as shown in Figure 2.I, a shift of the origin by -II will produce a new matrix $A_{L}=A-L I$, and $A_{L}$ can be written as

$$
\begin{align*}
A_{L} & =M \Lambda M^{-1}-L I=\because[1-L I] M^{-1} \\
& =M\left[\begin{array}{llll}
\lambda_{1}-L & & \\
& \lambda_{2}-I & & \\
& & \ddots & \\
& & & \lambda_{n}-I
\end{array}\right] M^{-1} \tag{2.34}
\end{align*}
$$

The sign of $A_{L}$ will be of the form

$$
\operatorname{sign} A_{L}=M\left[\begin{array}{ll}
I_{S \times s} & 0  \tag{2.35}\\
0 & -I_{n-s \times n-s}
\end{array}\right] M^{-1}
$$

Either of the subblocks in (2.35) can be isolated by adding or subtracting the identity matrix $I$ of the order $n \times n$.

Since $\left|R_{e} \lambda_{1}\right|<\left|R_{e} ;_{2}\right|<\ldots<\left|R_{e} \lambda_{n}\right|$, the subset $\Lambda_{1}$ represents the eigenvalues with the smaller $\left|R_{e} \lambda_{1}\right|$. This set of eigenvalues should be included in the low pass filter. The low pass filter, $S_{\text {Lp }}$, is then

$$
\begin{equation*}
S_{L p}=S_{1}=\frac{1}{2}[I+\operatorname{sign}(A-I I)]=M U_{1} M^{-1} \tag{2.36}
\end{equation*}
$$

where

$$
\begin{aligned}
& U_{1}=\operatorname{diag}\left(U_{11}, \ldots, U_{n 1}\right), \text { and } \\
& U_{i 1}= \begin{cases}1 & i=1, \ldots, s \\
0 & i=s+1, \ldots, n\end{cases}
\end{aligned}
$$

The trace of $S_{1}$ will be an integer which is equal to the number of eigenvalues in the subset $\Lambda_{1}$. The product $S_{1} A$ or $A S_{1}$ isolates the subset $\Lambda_{1}$ so that

$$
A_{1}=S_{1} A=A S_{1}=M\left[\begin{array}{ll}
\Lambda_{1} & 0  \tag{2.37}\\
0 & 0
\end{array}\right] \quad M^{-1}
$$

where the trace of $A_{1}$ equals the trace of $\Lambda_{1}$
Since the high pass filter $S_{H p}$ is the complement of $S_{L p}$,

$$
\begin{align*}
S_{H p}=S_{2} & =I-S_{L p}=\frac{1}{2}[I-\operatorname{sign}(A-L I)] \\
& =M U_{2} M^{-1} \tag{2.38}
\end{align*}
$$

where $\mathrm{U}_{2}=$ diag $\left(\mathrm{U}_{21}, \ldots, \mathrm{U}_{2 \mathrm{n}}\right)$, and

$$
U_{i 2}= \begin{cases}0 & i=1, \ldots, s \\ 1 & i=s+1, \ldots, n\end{cases}
$$

then $\Lambda_{2}$ is isolated as

$$
A_{2}=S_{2} A=A S_{2}=M\left[\begin{array}{ll}
0 & 0  \tag{2.39}\\
0 & \Lambda_{2}
\end{array}\right] \quad M^{-1}
$$

### 2.5.3 Band-pass Filter Matrix

A band-pass filter matrix is now derived for the set of eigenvalues with intermediate magnitude. The case where $\Lambda_{1}=$ $\operatorname{diag}\left(\lambda_{1}, \ldots, \lambda_{r}\right), \Lambda_{2}=\operatorname{diag}\left(\lambda_{r+1}, \ldots, \lambda_{s}\right)$ and $\Lambda_{3}=\operatorname{diag}\left(\lambda_{s+1}\right.$, $\ldots, \lambda_{n}$ ) is shown in Figure 2.2 .

The generation of the high-pass matrix, $S_{h p}$ is analogous to the previous case. $S_{h p}$ is then

$$
\begin{align*}
S_{h p} & =S_{3}=\frac{1}{2}[I-\operatorname{sign}(A-I I)] \\
& =M U_{3} M^{-1} \tag{2.40}
\end{align*}
$$

where $\mathrm{U}_{3}=\operatorname{diag}\left(\mathrm{U}_{13}, \ldots, \mathrm{U}_{\mathrm{n} 3}\right)$ and

$$
U_{i 3}= \begin{cases}0 & i=1, \ldots, r \\ 0 & i=r \div 1, \ldots, s \\ 1 & i=s \div 1, \ldots, n\end{cases}
$$

The matrix with the isolated subset $\Lambda_{3}$ becomes

$$
A_{3}=S_{3} A=A S_{3}=\because\left[\begin{array}{lll}
0 & 0 & 0  \tag{2.41}\\
0 & 0 & 0 \\
0 & 0 & \Lambda_{3}
\end{array}\right] \mathrm{M}^{-1}
$$

To form the low-pass filter matrix, trace $A_{3}$ should be subtracted from trace $A$ and the remainder should be divided by the number of remaining eigenvilues for a second shift of the origin

$$
\begin{equation*}
L_{1}=\frac{\operatorname{trace} A-\operatorname{trace} A_{3}}{n-\operatorname{trace} S_{3}} \tag{2.42}
\end{equation*}
$$

which is assumed to be esual to $I_{1}$ on Figure 2.2.


Figure 2.2

The new shifted matrix is then

$$
\left.\begin{array}{rl}
A_{L 1} & =\dot{A}-L_{1} I=M \Lambda M^{-1}-L_{1} I \\
& =M\left[\begin{array}{llll}
\lambda_{1}-L_{1} & & & \\
& \lambda_{2}-I_{1} & & \\
& & & \ddots
\end{array}\right] M^{-1}  \tag{2.43}\\
& \\
& \\
& \\
& \\
& \\
&
\end{array}\right]
$$

and the sign of $A_{\text {Ll }}$ is found as shown in

$$
\operatorname{sign} A_{L I}=M\left[\begin{array}{cc}
I_{r \times r} & 0  \tag{2.44}\\
0 & -I_{n-r \times n-r}
\end{array}\right] M^{-1}
$$

Since the subset $\Lambda_{1}$ represents the low-pass filter,

$$
\begin{align*}
S_{I \rho} & =S_{1}=\frac{1}{2}\left(\operatorname{sign} A_{L I}+I\right) \\
& =M U_{1} M^{-1} \tag{2.45}
\end{align*}
$$

where

$$
\begin{aligned}
& U_{1}=\operatorname{diag}\left(U_{11}, \ldots, U_{n 1}\right), \text { and } \\
& U_{i 1}= \begin{cases}1 & i=1, \ldots, r \\
0 & i=r+1, \ldots, s \\
0 & i=s+1, \ldots, n\end{cases}
\end{aligned}
$$

Similar to $S_{H p}$, the irace of $S_{1}$ is equal to the number of eigenvalues in subset $\Lambda_{1}$. The isolation of $\Lambda_{1}$ is given below:

$$
A_{1}=S_{1} A=A S_{1}=\cdots \begin{array}{lll}
\hat{A}_{1} & 0 & 0  \tag{2.46}\\
0 & 0 & 0 \\
0 & 0 & 0
\end{array} \quad M^{-1}
$$

The band-passfilter $S_{B P}$ must be the complement of $S_{L p}$ and $\mathrm{S}_{\mathrm{Hp}}$; that is,

$$
S_{B p}=I-S_{L p}-S_{H p}
$$

Then $S_{B p}$ is expressed as

$$
\begin{equation*}
S_{B p}=M U_{2} M^{-1} \tag{2.47}
\end{equation*}
$$

where

$$
\begin{aligned}
& U_{2}=\operatorname{diag}\left(U_{12}, \ldots, U_{n 2}\right), \text { and } \\
& U_{i 2}= \begin{cases}0 & i=1, \ldots, r \\
1 & i=r+1, \ldots, s \\
0 & i=s+1, \ldots, n\end{cases}
\end{aligned}
$$

The eigenvalues in the suiset $\therefore_{2}$ can now be isolated directly by forming $\mathrm{A}_{2}$ as

$$
A_{2}=A-A_{1}-A_{3}=U\left[\begin{array}{lll}
0 & 0 & 0  \tag{2.48}\\
0 & A_{2} & 0 \\
0 & 0 & 0
\end{array}\right] \mathrm{M}^{-1}
$$

Any number of band-passfirters can be generated by the origin shifting procedure outlined for $S_{H p}$ and $S_{L p}$.

The algorithm described above does not consider the imaginary part $\omega_{i}$ of any complex eigenvalue. The imaginary parts can be examined since knowledge of $\omega_{i}$ for such eigenvalues may
be important for the integration of step sizes. This examination can be accomplished by rotating the spectrum of any $A_{i}$.

To rotate the spectrum, $A_{i}$ can be multiplied by $j$ in which case the sign algorithm becomes

$$
\begin{equation*}
(\operatorname{sign} j A)^{i+1}=\frac{1}{2}(\operatorname{sign} j A)^{i}+\frac{1}{2}\left((\operatorname{sign} j A)^{i}\right)^{-1} \tag{2.49}
\end{equation*}
$$

Since $j$ is a scalar, (2.49) can be written as

$$
\begin{equation*}
(\operatorname{sign} A)^{i+1}=\frac{1}{2}(\operatorname{sign} A)^{i}-\frac{1}{2}\left((\operatorname{sign} A)^{i}\right)^{-1} \tag{2.49a}
\end{equation*}
$$

which is the desired form for the rotated spectrum. The imaginary parts of the eigenvalues have now been mapped into the real domain and the real parts into the imaginary. The sign matrix of the rotated matrix will be the identity matrix if all imaginary parts of the eigenvalues are zero or are of the same sign.

The magnitude of $\omega_{i}$ can be determined by a slift of the origin of the rotated matrix.

### 2.6 Examples

Two example problems are designed to illustrate the numerical procedure and the properties of uncoupled systems. Examples for a matrix filter generation will be given in Chapter $v$.

### 2.6.1 Continuous System Matrix

The backward system matrix for a linear regulator was obtained using the following coefficient matrices:

$$
A=\left[\begin{array}{cc}
0.0 & 1.0 \\
2.0 & -1.0
\end{array}\right] \quad B=\left[\begin{array}{l}
0.0 \\
1.0
\end{array}\right] \quad \Omega=\left[\begin{array}{ll}
2.0 & 0.0 \\
0.0 & 1.0
\end{array}\right] \text { and } R=[1.0]
$$

which give

$$
\mathrm{H}_{\mathrm{CR}}=\left[\begin{array}{rr|rr}
0.0 & -2.0 & -2.0 & 0.0 \\
-1.0 & 1.0 & 0.0 & -1.0 \\
\hdashline 0.0 & 0.0 & 0.0 & \frac{1.0}{1.0} \\
0.0 & -2.0 & 2.0 & -1.0
\end{array}\right]
$$

A transformation matrix is constructed from the sign function of ${ }^{\mathrm{H}} \mathrm{CR}^{\text {as }}$

$$
\bar{V}=\left[\begin{array}{cc:cc}
1.0 & 0.2220-15 & -1.556 & -0.1988 \\
-0.2220-15 & 1.0 & -0.1988 & -0.4798 \\
\hdashline-0.6786 & -0.7189 & -0.8012 & 0.4798 \\
-0.7189 & -1.2005 & 1.3572 & -1.2811
\end{array}\right]
$$

The uncoupled system matrix is then formed by the similarity transforms

$$
\begin{aligned}
\overline{\mathrm{V}}^{-1} \mathrm{H}_{\mathrm{CR}} \overline{\mathrm{~V}} & =\left[\begin{array}{cc:cc}
1.3572 & -0.5622 & 0.2617-15 & -0.3451-17 \\
-0.281 \mathrm{I} & 2.2005 & 0.1388-15 & -0.1388-15 \\
\hdashline-0.1388-15 & 0.1388-16 & -1.35722 & 0.2811 \\
-0.4302-15 & -0.441-15 & 0.5622 & -2.2005
\end{array}\right] \\
& =\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Cl}} & 0 \\
0 & \mathrm{H}_{\mathrm{C} 2}
\end{array}\right]
\end{aligned}
$$

The desired relationship has been established hetween submatrices $\mathrm{H}_{\mathrm{Cl}}=-\mathrm{H}_{\mathrm{C} 2}^{\mathrm{T}}$.

### 2.6.2 Discrete System Matrix

The block diagonalization procedure for a discrete system matrix is illustrated by a numerical example. The coefficient matrices considered are

$$
\begin{aligned}
& \phi=\left[\begin{array}{rrr}
-1.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 2.0 \\
0.0 & -2.0 & 0.0
\end{array}\right] G=\left[\begin{array}{r}
2.0 \\
2.0 \\
-1.0
\end{array}\right] R=[1.0] \text { and } \\
& Q=\left[\begin{array}{rrr}
1.0 & 1.0 & 0.0 \\
-1.0 & 1.0 & 0.0 \\
0.0 & 0.0 & 0.0
\end{array}\right]
\end{aligned}
$$

which give a forward system matrix in the discrete domain

$$
\mathrm{H}_{\mathrm{DR}}=\left[\begin{array}{rrr:rrr}
-6.0 & 5.0 & 0.0 & 4.0 & -1.0 & -2.0 \\
-5.0 & 5.0 & 2.0 & 4.0 & -1.0 & -2.0 \\
2.5 & -4.5 & 0.0 & -2.0 & 0.5 & 1.0 \\
\hdashline 1.0 & -1.0 & 0.0 & -1.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5 \\
-0.5 & 0.5 & 0.0 & 0.0 & -0.5 & 0.0
\end{array}\right]
$$

Now $H_{D R}$ is transformed to the complex plane by the following:

$$
\begin{aligned}
&\left(\mathrm{H}_{\mathrm{DR}}+\mathrm{I}\right)^{-1}\left(\mathrm{H}_{\mathrm{DR}}-\mathrm{I}\right)=\left[\begin{array}{lc:c:c}
1.8 & -0.4 & 0.8 & -2.0 \\
0.8 & 0.6 & 0.8 & 0.1332-15 \\
0.6 & -0.8 & 0.6 & -0.3331-16 \\
\hdashline-0.5 & 0.1249-15 & 0.1631-16 & -1.8 \\
0.4163-16 & -0.4163-16 & 0.1388-17 & -0.4 \\
-0.5551-16 & 0.4163-16 & -0.2776-17 & -0.8
\end{array}\right] \\
& \\
&\left.\begin{array}{cc}
0.1388-15 & 0.1388-15 \\
0.9853-16 & 0.8604-16 \\
-0.8119-16 & -\frac{0.7910-16}{} \\
-0.8 & -0.6
\end{array}\right] \\
&-0.8
\end{aligned}
$$

A transformation matrix is constructed as before

$$
\begin{aligned}
& \overline{\mathrm{V}}=\left[\begin{array}{ccc:cc}
1.0 & 0.1337-16 & 0.2255-16 & -1.0518 & 0.1942 \\
0.1789-17 & 1.0 & -0.1301-16 & 0.1942 & -0.1295 \\
-0.1735-17 & -0.2602-17 & 1.0 & 0.4147 & -0.9710-01 \\
\hdashline-0.1476 & 0.2289-02 & 0.5948-01 & -1.8196 & -0.3474-01 \\
0.2289-02 & -0.1356-01 & -0.1132-01 & -0.9733-02 & -1.9967 \\
0.5948-01 & -0.1132-01 & -0.5832-01 & -0.8895-01 & 0.1868-01 \\
-0.4147 \\
-0.9710-01 \\
-0.9442 \\
-0.93-01 \\
0.8424-02 \\
-1.9425
\end{array}\right]
\end{aligned}
$$

A similarity transformation can be applied directly on $H_{n R}$ by $\bar{V}$

$$
\begin{aligned}
& \bar{V}^{-1} H_{D R} \bar{V}=\left[\begin{array}{cc}
{ }^{H} D 1 & 0 \\
0 & \mathrm{~F}_{\mathrm{D} 2}
\end{array}\right] \\
& =\left[\begin{array}{ccc:}
6.7118 & -5.0453 & -0.3659 \\
5.7113 & -5.0453 & -2.3659 \\
-2.8559 & 4.5227 & 0.1829 \\
-\overline{0.1} \overline{09} \overline{1}-15 & -0 . \overline{14} \overline{40} \overline{-15} & -\frac{0.1 \overline{6} \overline{3}-16}{} \\
-0.1530-15 & 0.1278-15 & 0.1723-16 \\
-0.1847-15 & 0.1995-15 & 0.1524-17
\end{array}\right. \\
& \left.\begin{array}{ccc}
0.4230-15 & -0.6705-15 & -0.5794-15 \\
0.5783-15 & -0.7389-15 & -0.7072-15 \\
0.1656-15 & -\frac{0.4004-15}{0.1751}- & -\frac{0.7638-16}{0.3502}- \\
-0.2997 & 0.5608-02 & -0.4888 \\
-0.2243-01 & 0.4227 & -0.1547
\end{array}\right]
\end{aligned}
$$

where the inverse is

$$
\begin{aligned}
& \bar{V}^{\left.-I_{H_{D R}}^{-I} \bar{V}=\left[\begin{array}{cc}
H_{D 1}^{-1} & 0 \\
0 & H_{D 2}^{-1}
\end{array}\right], ~\right]} \\
& =\left[\left.\begin{array}{ccc}
0.2997 & -0.2243-01 & 0.3093 \\
0.1751 & 0.5608-02 & 0.4227 \\
-\frac{0.3052}{-0.4944-16} \overline{-}-\frac{0.4888}{0.1179-15} & -\frac{0.1547}{0.25} \overline{47} \overline{-15} \overline{0} \\
0.7383-16 & -0.1061-15 & -0.2536-15 \\
-0.3260-16 & -0.1537-16 & -0.6397-17
\end{array} \right\rvert\,\right. \\
& \left.\begin{array}{ccc}
-0.1462-14 & -0.1565-14 & 0.7695-15 \\
-0.7747-15 & -0.7736-15 & 0.2183-15 \\
-0.1448-14 & 0.1450-14 & -0.7729-15 \\
-\frac{-0.7118}{-1.718} & -2.8559 \\
-5.0453 & -5.0453 & 4.5227 \\
-0.3659 & -2.3659 & 0.1829
\end{array}\right]
\end{aligned}
$$

which results in the expected relation $H_{D 1}=H_{D}^{-T}$.

### 2.6.3 Filter Matrix

The A matrix to illustrate the filter matrix generation was constructed from $A=M \Lambda M^{-1}$ with

$$
\Lambda=\operatorname{diag} \begin{array}{llll}
{[-1200} & -500 & -1.5 & -0.5]
\end{array}
$$

and

$$
M=\left[\begin{array}{llll}
0 & 1 & 2 & 2 \\
1 & 1 & 2 & 3 \\
2 & 2 & 2 & 3 \\
2 & 3 & 3 & 3
\end{array}\right]
$$

Three filter matrices were computed, a low pass which contained $\lambda=-1.5$, and -0.5 , a band pass with $\lambda=-500$ and a high pass with $\lambda=-1200$. The filter matrices are given below

$$
S_{L P}=S_{1}=\left[\begin{array}{llll}
-2 & 4 & -4 & 2 \\
-0.1 E-15 & 2 & -1 & 0.1 \mathrm{E}-15 \\
-0.2 \mathrm{E}-15 & 2 & -1 & 0.2 \mathrm{E}-15 \\
-2 & 6 & -6 & 3
\end{array}\right]
$$

$$
S_{\mathrm{BP}}=\mathrm{S}_{2}=\left[\begin{array}{llll}
3 & -4 & 4 & -2 \\
3 & -4 & 4 & -2 \\
6 & -8 & 8 & -4 \\
9 & -12 & 12 & -6
\end{array}\right]
$$

and

$$
S_{H P}=S_{3}=\left[\begin{array}{lcl}
-0.18 E-15 & -0.18 \mathrm{E}-15 & -0.18 \mathrm{E}-15 \\
-3 & 3 & -3 \\
-6 & 6 & -6 \\
-6 & 6 & -6
\end{array}\right]
$$

## CHAPTER III

SOLUTION OF THE UNCOUPLED CONTINUOUS SYSTEM

### 3.1 Linear Systems with Quadratic Criteria

The theory of the linear optimal regulator problem is well known and documented [9,39]. The process is controlled by the state equation

$$
\begin{equation*}
\dot{X}(t)=A x(t)+B U(t) \tag{3.1}
\end{equation*}
$$

with the initial condition

$$
\begin{equation*}
x\left(t_{0}\right)=x_{0} \tag{3.1a}
\end{equation*}
$$

The desired system is brought from an initial state $X_{0}$ to a terminal state $X\left(t_{f}\right)$ using acceptable levels of control $U(t)$ and the state $x(t)$ on the trajectory. One way to accomplish this is to minimize a cost functional. This functional is made up of a positive definite quadratic form in the terminal state, in addition to an integral of positive definite quadratic forms in the state and the control as shown below:

$$
\begin{equation*}
J(U)=\frac{1}{2} X^{T}\left(t_{f}\right) H x\left(t_{f}\right)+\frac{1}{2} \int_{t_{0}}^{t_{f}}\left[X^{T}(t) Q X(t)+U^{T}(t) R U(t)\right] d t \tag{3.2}
\end{equation*}
$$

where the final time $t_{f}$ is fixed, $H$ and $Q$ are real symmetric positive semi-definite matrices, and $R$ is a real symmetric positive definite matrix.

It is assumed that the state and controls are not bounded and $X\left(t_{f}\right)$ is free. The problem is solved via the Pontryagin
maximum principle crthe Hamilton-Jacobi equation. Here, the former method is used.

The Hamiltonian is formed by introducing the costate vector $\lambda(t)$ as

$$
\begin{align*}
H[X(t), U(t), \lambda(t), t] & =\frac{1}{2} X^{T}(t) Q X(t)+\frac{1}{2} U^{T}(t) R U(t) \\
& +\lambda^{T}(t) A X(t)+\lambda^{T}(t) B U(t) \tag{3.3}
\end{align*}
$$

The necessary conditions For the optimality are

$$
\begin{align*}
& \dot{X}(t)=A X(t)+B(t)  \tag{3.4}\\
& \dot{\lambda}(t)=-Q X(t)-A^{T} ;(t)  \tag{3.4a}\\
& U(t)=-R^{-1} B^{T} \lambda(t) \tag{3.4b}
\end{align*}
$$

with the terminal condition

$$
\begin{equation*}
\lambda\left(t_{f}\right)=H X\left(t_{f}\right) \tag{3.4c}
\end{equation*}
$$

Substituting (3.4b) for (3.4) and repeating (3.4a), the following linear two-point boundary-value problem may be developed:

$$
\left[\begin{array}{l}
\dot{x}(t)  \tag{3.5}\\
\dot{\lambda}(t)
\end{array}\right]={ }^{H} C R\left[\begin{array}{l}
x(t) \\
\lambda(t)
\end{array}\right.
$$

where $H_{C R}$ is the system matrix associated with the continuous system with a quadratic performance index with coefficient matrices as follows:

$$
H_{C R}=\left[\begin{array}{ll}
A & -B R^{-1} B^{2}  \tag{3.5a}\\
-Q & -A^{T}
\end{array}\right]
$$

Since the problem is linear, and the differential equations are homogeneous, it is clear that both $X(t)$ and $\lambda(t)$ are proportional to $X_{0}$. All existing solutions to this boundary value problem are designed to obtain the linear relation between the state and costate vectors, denoted by the feedback optirnal gain matrix.
3.1.1 Solution by a Jransition Matrix The solutions to (3.5) may have the form

$$
\left[\begin{array}{l}
x\left(t_{f}\right)  \tag{3.6}\\
\lambda\left(t_{f}\right)
\end{array}\right]=\left[\begin{array}{ll}
\phi_{11}\left(t_{f} t\right) & s_{12}\left(t_{f}, t\right) \\
\phi_{21}\left(t_{f} t\right) & \delta_{22}\left(t_{f} t\right)
\end{array}\right]\left[\begin{array}{l}
x(t) \\
\lambda(t)
\end{array}\right]
$$

where $\phi_{i j}$ are $n \times n$ partitions of the transition matrix of (3.5). Substituting (3.6) for the terminal condition (3.Ac),

$$
H \phi_{11}\left(t_{f}, t\right) X(t)+H \phi_{12}\left(t_{\bar{i}}, t\right) \lambda(t)=\phi_{21}\left(t_{f}, t\right) X(t)+\phi_{22}\left(t_{E}, t\right) \lambda(t)
$$

The solution for $\lambda(t)$ may be written as

$$
\begin{equation*}
\lambda(t)=p(t) x(t) \tag{3.7a}
\end{equation*}
$$

where

$$
\begin{equation*}
p(t)=\left[\phi_{22}\left(t_{f^{\prime}} t\right)=H^{\prime}{ }_{12}\left(t_{f^{\prime}} t\right)\right]-1\left[H_{12}\left(t_{f} t\right)-\phi_{21}\left(t_{f^{\prime}}, t\right)\right] \tag{3.7b}
\end{equation*}
$$

3.1.2 Solution by the Riccati Equation

The idea of an alternative approach is contained in
(3.7a) and (3.7b). The value $p(t)$ is determined directly by integrating a matrix Riccati equation derived in the following procedures:

Substituting (3.7a) for (3.4a) yields

$$
\begin{equation*}
\dot{p}(t) X(t)+p(t) \dot{X}(t)=-Q X(t)-A^{T} \lambda(t) \tag{3.8}
\end{equation*}
$$

Next, substituting $\dot{X}(t)$ from (3.1) into (3.8) and using (3.7a) again,

$$
\begin{equation*}
\left[\dot{p}(t)+p(t) A(t)+A^{T}(t) p(t)-p(t) B R^{-1} B^{T} p(t)+Q\right] X(t)=0 \tag{3.8a}
\end{equation*}
$$

Since $X(t) \neq 0$, (3.8a) requires that

$$
\begin{equation*}
\dot{p}(t)=-p(t) A \quad-A^{T} p(t)+p(t) B R^{-1} B^{T} p(t)-Q \tag{3.8b}
\end{equation*}
$$

From the terminal condition in (3.4c), it is evident that

$$
\begin{equation*}
p\left(t_{f}\right)=H \tag{3.8c}
\end{equation*}
$$

Since $p(t)$ is an $n \times n$ symmetric matrix, it is necessary to solve the $n(n+1) / 2$ íirst-order differential equations. Many numerical techniques have been developed to solve the matrix Riccati equation efficienily. In general, the integration is started at $t=t_{f}$ and proceeds backward in time to $t=t_{0}$, and $p(t)$ is stored for the feedback law and optimum state. It is then possible to determine $\lambda\left(\tau_{0}\right)$ from (3.7a) as follows:

$$
\begin{equation*}
\lambda\left(t_{0}\right)=p\left(t_{0}\right) \times\left(t_{0}\right) \tag{3.9}
\end{equation*}
$$

which may be regarded as the equivalent of the terminal boundary condition in (3.4c) at an earlier time. An alternate method has been derived to obtain the solution to (3.6) by forward integration, since $\lambda\left(t_{0}\right)$ and $X\left(t_{0}\right)$ are both known $(15,31,34)$.

### 3.1.3 The Hamilton-Jacobi-Bellman Equation

The Hamilton-Jacobi-Bellman equation can be used as a means of solving the general form of the continuous linear regulator problem [23]. The Familtonian for the use of Hamilton-Jacobi-Bellman equation is formed in this case by the following:

$$
\begin{align*}
H\left[X(t), U(t), J_{X}, t\right]= & \frac{1}{2} X^{T}(t) Q X(t)+\frac{1}{2} U^{T}(t) R U(t)+J_{X}^{T}(X(t), t) \\
& {[\operatorname{[AX}(t)+B U(t)} \tag{3.10}
\end{align*}
$$

A necessary condition for $U(t)$ to minimize $H$ is that $\frac{\partial H}{\partial U}=0 ;$ thus,

$$
\begin{equation*}
U(t)=-R^{-1} B^{T} J_{X}(X(t), t) \tag{3.11}
\end{equation*}
$$

which, when substituted in (3.10), yields

$$
\begin{equation*}
I\left(X, U, J_{X}, t\right)=\frac{1}{2} X^{T} Q X-\frac{1}{2} J_{X}^{T} B^{-1} B^{T} J_{X}+J_{X}^{T} A X \tag{3.12}
\end{equation*}
$$

Now the Hamilton-Jacobi-Bellman equation is

$$
\begin{equation*}
0=J_{t}+\frac{1}{2} X^{T} Q X-\frac{1}{2} J_{X}^{T} B R^{-1} B_{X}^{T} J_{X}+J_{X}^{T} A X \tag{3.13}
\end{equation*}
$$

with the boundary condition

$$
\begin{equation*}
J\left(X\left(t_{f}\right), t_{f}\right)=\frac{1}{2} X^{T}\left(t_{f}\right) H X\left(t_{f}\right) \tag{3.13a}
\end{equation*}
$$

It is assumed that the minimum cost for the linear regulator problem is a quadratic function of the state and has the form

$$
\begin{equation*}
J[X(t), t]=\frac{1}{2} X^{T}(t) p(t) X(t) \tag{3.14}
\end{equation*}
$$

where $p(t)$ is a real symmetric positive-definite matrix to be
determined.
Substituting (3.16) in (3.13) and using the property of symmetric matrix,

$$
\begin{equation*}
\dot{p}(t)=-Q-p(t) B R^{-1} B^{T} p(t)-p(t) A-A^{T} p(t) \tag{3.15}
\end{equation*}
$$

and the boundary condition obtained from (3.13a) and (3.14) is

$$
\begin{equation*}
p\left(t_{f}\right)=H \tag{3.15a}
\end{equation*}
$$

In the derivation of the Riccati equation developed in (3.8b) and (3.8c), the equation (3.13) gives another interpretation of the meaning of $p(t)$; that is, $\frac{1}{2} X^{T}(t) p(t) X(t)$ is the minimum value of the cost function starting at time $t$ with state vector $X(t)$.

## A Reduced Linear Pequlator

The uncoupling algorithm based on the matrix block diagonalization can be a ponerful tool in solving a two-point boundary value problem. By reducirg the system order to be integrated, the determination of the rissing initial condition will be faster and more accurate. An efficient method for the solution of a linear regulator is studied in this section. This technique can be extended easily to other optimal controls and estimation problems with quadiratic criteria since their system matrices have a structure analogous to that of the linear regulator.

The vectors $y(t)$ and $:(t)$ are defined by the relation

$$
\left[\begin{array}{l}
X(t)  \tag{3.16}\\
\lambda(t)
\end{array}\right]=\left[\begin{array}{ll}
V_{11} & V_{12} \\
V_{21} & v_{22}
\end{array}\right]\left[\begin{array}{l}
y(t) \\
\omega(t)
\end{array}\right]
$$

where the elements of $V_{i j}$ are given by (2.9a).
Differentiating both sides of (3.16) and substituting (3.5) for it,

$$
\left[\begin{array}{c}
\dot{\mathrm{X}}  \tag{3.17}\\
\dot{\lambda}
\end{array}\right]=\mathrm{H}_{\mathrm{CR}}\left[\begin{array}{l}
\mathrm{X} \\
\lambda
\end{array}\right]={ }_{\mathrm{H}}^{\mathrm{CR}}\left[\begin{array}{cc}
\mathrm{V}_{11} & \mathrm{~V}_{12} \\
\mathrm{~V}_{21} & \mathrm{~V}_{22}
\end{array}\right]\left[\begin{array}{l}
\mathrm{Y} \\
\omega
\end{array}\right]=\left[\begin{array}{cc}
\mathrm{V}_{11} & \mathrm{~V}_{12} \\
\mathrm{~V}_{21} & \mathrm{~V}_{22}
\end{array}\right]\left[\begin{array}{l}
\dot{\mathrm{Y}} \\
\dot{\omega}
\end{array}\right]
$$

from which

Using the result given by (2.10), the uncoupled system equation for the lineax regulator becomes

$$
\left[\begin{array}{c}
\dot{y}(t)  \tag{3.18a}\\
\dot{\omega}(t)
\end{array}\right]=\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Cl}} & 0 \\
0 & -\mathrm{H}_{\mathrm{C} 1}^{\mathrm{T}}
\end{array}\right]\left[\begin{array}{c}
\mathrm{Y}(\mathrm{t}) \\
\\
\omega(\mathrm{t})
\end{array}\right]
$$

If $\underset{i j}{\psi}\left(t t_{0}\right)$ is the state transition matrix for the uncoupled system (5.18a), it follows that

$$
\left[\begin{array}{lll}
\dot{\psi}_{1 I}\left(t t_{0}\right) & 0  \tag{3.19}\\
0 & \dot{\psi}_{22}\left(t t_{0}\right)
\end{array}\right]=\left[\begin{array}{ll}
\mathrm{H}_{\mathrm{C} 1} & 0 \\
0 & -\mathrm{H}_{\mathrm{Cl}}^{\mathrm{T}}
\end{array}\right]\left[\begin{array}{ll}
\psi_{11}\left(\mathrm{t} \hbar_{0}\right) & 0 \\
0 & \psi_{22}\left(t \hbar_{0}\right)
\end{array}\right]
$$

where the partitions of $\psi, \psi_{11}$, and $\psi_{22}$ are the state transition matrices for the system equations $y(t)$ and $\omega(t)$, respectively.

The initial conditions for (3.19) are

$$
\begin{equation*}
\psi_{11}\left(t_{0}, t_{0}\right)=U_{22}\left(t_{0}, t_{0}\right)=I \tag{3.19a}
\end{equation*}
$$

A direct series expansion has been adopted to integrate (3.19) and (3.19a) for an increment of time $t$.

$$
\begin{align*}
\Delta \psi_{11} & =\psi_{11}\left(t+\Delta t, 亡_{0}\right)=\exp \left(\bar{H}_{\mathrm{Cl}} \Delta t\right)=I+\mathrm{H}_{\mathrm{Cl}} \Delta t \\
& +\frac{\left(\mathrm{H}_{\mathrm{Cl}}\right)^{2} \Delta t^{2}}{2!} \div \frac{\left(\mathrm{H}_{\mathrm{Cl}}\right)^{3} \Delta t^{3}}{3!}+\ldots \tag{3.19b}
\end{align*}
$$

The following recursive equations carry on the forwarding integration

$$
\begin{align*}
\mathrm{y}_{i+1} & =\Delta \psi_{11} \mathrm{y}_{i}  \tag{3.19c}\\
\omega_{i+1} & =\Delta \psi_{22} \omega_{i} \tag{3.19d}
\end{align*}
$$

and it can be seen from (3.18a) that the differential equation $\dot{\omega}=-\mathrm{H}_{\mathrm{Cl}}^{\mathrm{T}} \omega$ is the adjoint to the system equation $\dot{\mathrm{Y}}=\mathrm{H}_{\mathrm{CI}} \mathrm{Y}$.

The importance $0 \ddagger$ this fact is that there exists a relationship between the state transition matrices of the adjoint system and the original system so that

$$
\begin{equation*}
\psi_{22}\left(t \hbar_{0}\right)=\psi_{11}^{T}\left(t \hbar_{0}\right)^{-1} \tag{3.20}
\end{equation*}
$$

This relationship will reciuce the computational effort so that it is necessary to solve only one differential equation, either $\dot{y}(t)$ or $\dot{\omega}(t)$.

Another advantage of the adjoint system is that the physically realizable solution of the adjoint system represents
the physically unrealizable portion of the solution of the original system.

Now the solution vectors $y(t)$ and $\omega(t)$ are given by

$$
\begin{equation*}
y(t)=\psi_{11}\left(t, t_{0}\right) y\left(t_{0}\right) \tag{3.21}
\end{equation*}
$$

and

$$
\begin{equation*}
\omega(t)=\left(\psi_{11}^{T}\left(t, t_{0}\right)\right)^{-1} \omega\left(t_{0}\right) \tag{3.21a}
\end{equation*}
$$

The initial conditions $y\left(t_{0}\right)$ and $\omega\left(t_{0}\right)$ must be determined to complete the formulation. Using the relationship between the coupled and uncoupled solution vectors defined by (3.16), the terminal. condition in (3.4c) can be rewritten as

$$
\begin{equation*}
V_{21} Y\left(t_{f}\right)+V_{22} \omega\left(t_{f}\right)=H\left[V_{11} Y\left(t_{f}\right)+V_{12} \omega\left(t_{f}\right)\right] \tag{3.22}
\end{equation*}
$$

from which a new terminal condition is defined as

$$
\begin{equation*}
Y\left(t_{f}\right)=\bar{H} \omega\left(t_{f}\right) \tag{3.22a}
\end{equation*}
$$

where

$$
\begin{equation*}
\bar{H}=-\left(V_{21}-H V_{11}\right)^{-1}\left(V_{22}-H V_{12}\right) \tag{3.22b}
\end{equation*}
$$

H represents a terminal weighting matrix for the uncoupled system.

Substituting $Y\left(t_{f}\right)$ and $\omega\left(t_{f}\right)$ in (3.21) and (3.2la) for
(3.22a) and solving for $y\left(t_{0}\right)$ results in

$$
\begin{equation*}
y\left(t_{0}\right)=s_{0} \omega\left(t_{0}\right) \tag{3.23}
\end{equation*}
$$

where

$$
\begin{equation*}
S_{0}=\left[\psi_{11}\left(t_{f} t_{0}\right)\right]^{-1} \bar{H}\left[\stackrel{U}{11}_{T}\left(t_{f}, t_{0}\right)\right]^{-1} \tag{3.23a}
\end{equation*}
$$

The initial conditions for the uncoupled system are determined from the given initial condition $X_{0}$. Using (3.16) again,

$$
\begin{equation*}
\omega\left(t_{0}\right)=\left(v_{11} S_{0}+v_{12}\right)^{-1} x_{0} \tag{3.24}
\end{equation*}
$$

and

$$
\begin{equation*}
y\left(t_{0}\right)=s_{0} \omega\left(t_{0}\right)=\left(v_{11}+V_{12} s_{0}^{-1}\right)^{-1} x_{0} \tag{3.24a}
\end{equation*}
$$

The desired solution for the original system equations can then be reconstructed from the solution of the uncoupled system

$$
\left[\begin{array}{l}
x(t)  \tag{3.25}\\
\lambda(t)
\end{array}\right]=\left[\begin{array}{ll}
v_{11} & v_{12} \\
v_{21} & v_{22}
\end{array}\right]\left[\begin{array}{ll}
i_{11}\left(t, t_{0}\right) & 0 \\
0 & {\left[\psi_{11}^{T}\left(t t_{0}\right)-1\right.}
\end{array}\right]\left[\begin{array}{l}
y\left(t_{0}\right) \\
\omega\left(t_{0}\right)
\end{array}\right]
$$

The technique described offers several advantages over conventional methods. The main advantage is that there is no need for any calculation of the feedback gain matrix, thus eliminating costly storage problems. Since the uncoupled system equations are adjoint to each other, the $2 n \times 2 n$ original coupled system is transformed to an $n \times n$ subsystem without changing any system characteristics.

The missing initial condition $\lambda_{0}$ can be obtained from (3.24) and (3.24a).

$$
\begin{equation*}
\lambda_{0}=P_{0} X_{0} \tag{3.26}
\end{equation*}
$$

where

$$
\begin{equation*}
p_{0}=\left(v_{21} s_{0}+v_{22}\right)\left(v_{11} s_{0}+v_{12}\right)^{-1} \tag{3.26a}
\end{equation*}
$$

It can be seen that $P_{0}$ is the initial condition in Riccati equations for the continuous optimal feedback gain of linear regulators (3.8b) and is equivalent to the kickback equation in the matrix increment coefficient (NIC) algorithm developed earlier [31, 34].

Recalling that the solution matrix $P(t)$ of the Riccati equation is defined by $\lambda(t)=p(t) x(t)$, and using the relation in (3.16), $P(t)$,

$$
\begin{equation*}
P(t)=\left[V_{21} Y(t)+V_{\left.22^{\prime \nu}(t)\right]\left[V_{11} Y(t)+V_{12}\right.}(t)\right]^{-1} \tag{3.27}
\end{equation*}
$$

Substituting (3.21), (3.2la), and (3.23) for (3.27), $p(t)$ becomes

$$
\begin{align*}
P(t)= & {\left[V_{21} \psi_{11}\left(t, t_{0}\right) S_{0}+V_{22} \psi_{11}^{-T}\left(t, t_{0}\right)\right] } \\
& {\left[V_{11}{ }_{11}\left(t, t_{0}\right) S_{0}+V_{12} \psi_{11}^{-T}\left(t, t_{0}\right)\right]^{-1} } \tag{3.27a}
\end{align*}
$$

At $t=t_{0}$, the value $P\left(t_{0}\right)$ is obtained. This value is equivalent to $P_{0}$ in (3.26a). Equation (3.27a) is an algebraic expression for the optimal feedback gain in the linear regulator. This procedure can be applied to the estimation problem, and an algebraic solution for the covariance matrix can be obtained.

Since the reduced subblock $\mathrm{H}_{\mathrm{Cl}}$ is assumed to contain positive eigenvalues only, the inverse of its transition matrix
is numerically stable and will vanish when the terminal time becomes infinite. The situation in which the process is to be controlled for an interval of infinite duration attracts special attention in the performance evaluation of many types of linear filtering, prediction, and optimum controls. In general, the terminal weighting matrix $H$ is assumed to be zero.

When the terminal time is infinite, the uncoupled systems have simpler solutions as follows:

$$
\begin{align*}
& \text { Since } S_{0}=0 \text { as } t_{f}^{f+\infty} \\
& \lim _{f} \rightarrow \infty  \tag{3.28}\\
& \\
& \lim _{0}\left(t_{0}\right)=V_{12}^{-1} X_{0}
\end{align*}
$$

and

$$
\begin{equation*}
\lim _{t_{f}^{\rightarrow \infty}} Y\left(t_{0}\right)=0 \tag{3.28a}
\end{equation*}
$$

Therefore, from (3.21) and (3.21a),

$$
\begin{equation*}
y(t)=0 \tag{3.29}
\end{equation*}
$$

and

$$
\begin{equation*}
\omega(t)=\left[\psi_{11}^{T}\left(t t_{0}\right)\right]^{-1} v_{12}^{-1} x_{0} \tag{3.29a}
\end{equation*}
$$

The steady state solutions for the original system become

$$
\begin{align*}
& X(t)=V_{12}\left[\psi_{11}^{T}\left(t \hbar_{0}\right)\right]^{-1} V_{12}^{-1} x_{0}  \tag{3.30}\\
& \lambda(t)=V_{22}\left[\psi_{11}^{T}\left(t \hbar_{0}\right)\right]^{-1} V_{12}^{-1} x_{0} \tag{3.30a}
\end{align*}
$$

The steady state solution for $p(t)$ can be obtained either from (3.27a) or (3.30) and (3.30a). Both cases result in

$$
\begin{align*}
\lim _{\mathrm{t}_{\mathrm{f}} \rightarrow \infty} \mathrm{p}(\mathrm{t}) & =\mathrm{V}_{22^{\mathrm{V}_{1}}-\frac{1}{2}=\mathrm{s}^{\mathrm{T}}\left(\mathrm{~T}_{2} \mathrm{~S}^{\mathrm{T}}\right)^{-1}} \\
& =\mathrm{T}_{2}^{-1}=\operatorname{M}_{2} 2^{\mathrm{H}^{-1}} 12 \tag{3.31}
\end{align*}
$$

which is equivalent to the conventional positive definite solution of a matrix quadratic equation $[28,33]$. In many practical problems, the steady sta亡e solutions have been implemented without significant performance degradation to compare system performance using the time varying optimal gain.

The algorithm for solving the constant coefficient linear regulator problem can be summarized in the six steps below:
a) Obtain the state and costate equations to construct $\mathrm{H}_{\mathrm{CR}}$.
b) Use the sign algorithm to construct the transformation matrices $V$ and $\bar{V}$.

d) Obtain a rew 亡orminal weighting matrix $H$ as given in (3.22b).
e) Compute $S_{0}$ fron (3.23a).
f) Compute new initial condition vectors as specified in (3.24) and (3.24a).
g) Reconstruct solution vectors for a coupled system by (3.25).

## 3.3 <br> Examples

The algorithm has been used to solve numerous linear regulator problems. There is excellent agreement when results obtained by using the algorithm are compared with results obtained by using other computational procedures. Two examples will be discussed. The numerical values given were obtained by using the algorithm and by using a procedure for the EISPAK subroutine package.

The state equations for the first example are as follows: $\dot{X}(t)=A x(t)+B u(t)$
$A=\left[\begin{array}{cccccc}-0.21053 & -0.10526 & -0.0007378 & 0.0 & 0.0706 & 0.0 \\ 1.0 & -0.03537 & -0.0001180 & 0.0 & 0.0004 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -605.16 & -4.92 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0 & 0.0 & 1.0 \\ 0.0 & 0.0 & 0.0 & 0 & -3906.25 & -12.5\end{array}\right]$
$B^{T}=\left[\begin{array}{llllll}-7.211 & -0.05232 & 0.0 & 1.0 & 0.0 & -1.0\end{array}\right]$
$Q=[I]_{6 \times 6}$
$R=1.0$
$H=[0] 6 \times 6$
$X_{0}^{T}=\left[\begin{array}{llllll}1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0\end{array}\right]$

The computational results are given below where $\mathrm{X}_{1}, \mathrm{X}_{2}$, and $u$ are given. The values from the EISPAK program are given in the first column; the values from the algorithm are in the second column.

Table 3.1
Numerical Results for Example 1

|  |  |  | $\mathrm{X}_{1}$ |  | $\mathrm{X}_{2}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $t$ | EISPAK | algorithm | EISPAK | algorithm | EISPAK | algorith |
| 0.2 | 0.84040 | 0.84040 | 0.099001 | 0.099001 | -1.3221 | -1.3221 |
| 2.0 | 0.022247 | 0.022247 | 0.4430 | 0.4430 | -0.39998 | -0.39998 |
| 4.0 | -0.16908 | -0.16908 | 0.32909 | 0.32909 | -0.021178 | -0.02117 |
| 6.0 | -0.12993 | -0.12993 | 0.16242 | 0.16242 | 0.056753 | 0.05675 |
| 8.0 | -0.069071 | -0.069072 | 0.058280 | 0.058280 | 0.038792 | 0.03884 |
| 10.0 | -0.040360 | -0.041177 | 0.0043283 | 0.0043727 | $0.6605-02$ | $-0.357-0$ |

A system in which A is $9 \times 9$ was selected as the second example.


The first two states and the control are given in Table 3.2.

Table 3.2
Numerical Results for Example 2

| t |  | $\mathrm{X}_{2}$ |  | algorithm | u |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | EISPAK | algorith | EISPAK |  | EISPAK | algorithm |
| 0.2 | -4.4914 | -4.4914 | $-1.9499$ | -1.9499 | -4.6531 | -4.6531 |
| 2.0 | $-3.7688$ | -3.7688 | 1.1625 | 1.1625 | -0.18305 | -0.18305 |
| 4.0 | -2.4121 | -2.4121 | 0.19929 | 0.19929 | 0.40611 | 0.4051 .1 |
| 6.0 | -2.3989 | -2.3989 | -0.14663 | -0.14663 | 0.29003 | 0.29003 |
| 8.0 | $-2.6751$ | -2.6751 | -0.12266 | -0.12266 | 0.10517 | 0.10517 |
| 10.0 | $-2.7900$ | -2.7900 | 0.064273 | 0.064273 | 0.22-11 | 0.17-09 |

Both problems were run on an IBM 360 Model 44 in double precision. The first problem required 77.1 seconds of execution time by the new algorithm versus 88.8 seconds with EISPAK. The program for the algorithm was not optimized.

SOLUTION OF THE URCOUPIED DISCRETE SYSTEMS
4.1 A Discrete Linear Regulator

This section considers a linear constant coefficient discrete system represented by

$$
\begin{equation*}
\mathrm{X}_{\mathrm{K}+1}=\phi \mathrm{X}_{\mathrm{K}}+\mathrm{BU}_{\mathrm{K}} \tag{4.1}
\end{equation*}
$$

with the initial condiさion vector

$$
\begin{equation*}
X(0)=X_{0}, \quad Y=0,1,2, \ldots, N \tag{4.la}
\end{equation*}
$$

where $\phi$ is the state transition matrix and $N$ is assumed to be a fixed integer. The quadrȧic eerformance critexion is
where it is assumed that the weighting matrices $Q$ and $R$ are independent to stage $K$.

The Hamiltonian is formed by

$$
\begin{equation*}
\left.H=\frac{1}{2} X_{K}^{T} Q X_{K}+\frac{1}{2} T_{K}^{T} U_{K} \div\right)_{K+1}^{T}\left[\phi X_{K}+B U_{K}\right] \tag{4.3}
\end{equation*}
$$

When the discrete raximan principle is applied, the optimal control is given by

$$
\begin{equation*}
\mathrm{U}_{\mathrm{K}}=-\mathrm{R}^{-1} \mathrm{~B}^{\mathrm{T}} \lambda_{\mathrm{K}+\mathrm{I}} \tag{4.4}
\end{equation*}
$$

where $\lambda_{K+1}$ is the costate vector which obeys the equation

$$
\begin{equation*}
\lambda_{\mathrm{K}}=Q \mathrm{X}_{\mathrm{K}}+\phi^{T} \lambda_{\mathrm{K}+1} \tag{4.5}
\end{equation*}
$$

The boundary condition for the unspecified terminal states is given as

$$
\begin{equation*}
\lambda_{N}=H X_{N} \tag{4.5}
\end{equation*}
$$

Equations (4.1), (4.4), and (4.5) represent a set of linear difference equations for an open-loop control.

The system matrix of the discrete Riccati equation has been introduced previously (Section 2.2). The matrix is formulated from the canonical forward equations instead of the conventional backward approach, but it will be shown in a later part of this chapter that both structures of the system matrix produce equivalent feedback gain matrices at each stage.

The equation can be solved for $\lambda_{K+1}$ in terms of $\lambda_{K}$ since the state transition matrix, $\dot{=}$, always has an inverse

$$
\begin{equation*}
\left.\lambda_{\mathrm{K}+1}=-\phi^{-\mathrm{T}} \mathrm{QX}_{\mathrm{K}}+\phi^{-\mathrm{T}}\right\rangle_{\mathrm{K}} \tag{4.7}
\end{equation*}
$$

Substituting (4.7) into (4.1), with the expression of optimal control in (4.4), yields this equation for the state vector.

$$
\begin{equation*}
X_{K+1}=\left(\phi+B R^{-1} B_{B^{T}}{ }^{-T} Q\right) X_{K}-B R^{-1} B_{\phi^{T}}^{-T} \lambda_{K} \tag{4.7a}
\end{equation*}
$$

Equations (4.7) and (4.7a) are the desired canonical state and costate difference equations for the discrete linear regulator. This set of coupled equations may be expressed in the following compact form:

$$
\left[\begin{array}{l}
\mathrm{X}  \tag{4.8}\\
\lambda
\end{array}\right]_{K+1}=H_{D R}\left[\begin{array}{c}
\mathrm{X} \\
\lambda
\end{array}\right.
$$

where $H_{D R}$ represents the forward system matrix associated with the coefficients of the ciscrete linear regulator so that
4.2 The Riccati Equation for Discrete Optimal Feedback Gain Several alternatire, but equivalent, recursive relations have been derived to solve the set of discrete boundary value problems in Section 4.l. In general, a closed loop solution $\mathrm{p}_{\mathrm{K}}$ is assumed as

$$
\begin{equation*}
\lambda_{K}=P_{K} X_{K} \tag{4.9}
\end{equation*}
$$

Substituting (4.4) and (4.9) into (4.1) and (4.5) to eliminate $\lambda_{\mathrm{K}}$ yiclds

$$
\begin{equation*}
X_{K+1}=\phi X_{K}-B R^{-1} S^{T} P_{K+1} X_{K+1} \tag{4.10}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathrm{P}_{\mathrm{K}} \mathrm{X}_{\mathrm{K}}=Q \mathrm{X}_{\mathrm{K}}+\dot{G}^{\mathrm{T}} \mathrm{P}_{\mathrm{K}+1} \mathrm{X}_{\mathrm{K}+1} \tag{4.10a}
\end{equation*}
$$

By solving for $X_{K+1}$ and eliminating it, the following is obtained:

$$
\begin{equation*}
P_{K} X_{K}=Q X_{K}+\phi^{T} P_{K+1}\left[I+B R^{-1} B^{T} P_{K+1}\right]^{-1} \phi X_{K} \tag{4.11}
\end{equation*}
$$

from which

$$
\begin{equation*}
P_{K}=Q+\phi^{T} P_{K+1}\left[I+B R^{-1} B_{B+1}\right]^{-1} \phi \tag{4.11a}
\end{equation*}
$$

or equivalently

$$
\begin{equation*}
P_{K}=Q+\phi^{T}\left[P_{K+1}^{-1}+B R^{-1} B^{T}\right]^{-1} \phi \tag{4.11b}
\end{equation*}
$$

The terminal condition for $P_{K}$ is provided by (4.6) and (4.9) so that

$$
\begin{equation*}
\mathrm{P}_{\mathrm{N}}=\mathrm{II} \tag{4.12}
\end{equation*}
$$

Equations (4.1la) and (4.11b) are developed in the well known discrete matrix Riccati equation associated with discrete optimal controls.

Usually, the matrix Riccati equation is solved recursively backward in time, starting $K=N$ to $K=0$, and the feedback gain matrix $P_{K}$ is stored at each stage. From (4.10), the recursive equation for the closed-loop state vector is expressed in terms of the prestored gain.

$$
\begin{equation*}
X_{K+1}=\left(I+B R^{-1} B^{T} P_{K+1}\right)^{-1} X_{K} \tag{4.13}
\end{equation*}
$$

The optimal control $U_{K}$ is then obtained as

$$
\begin{equation*}
U_{K}=-R^{-1} B^{T} \phi^{-T}\left[P_{K}-Q\right] X_{K} \tag{4.13a}
\end{equation*}
$$

or in terms of $\mathrm{P}_{\mathrm{K}+1}$

$$
\begin{equation*}
\mathrm{U}_{\mathrm{K}}=-\mathrm{R}^{-1} \mathrm{~B}^{\mathrm{T}}\left[\mathrm{P}_{\mathrm{K}+1}^{-1}+\mathrm{BR}^{-1} \mathrm{~B}^{\mathrm{T}}\right]^{-1} \phi \mathrm{X}_{\mathrm{K}} \tag{4.13b}
\end{equation*}
$$

### 4.3 The Discrete Kalman Filter

The development of the discrete filter model may be given by the following set of first-order difference equations:

$$
\begin{equation*}
\mathrm{X}_{\mathrm{K}+1}=\phi \mathrm{X}_{\mathrm{K}}+\Gamma \omega_{\mathrm{K}} \quad \mathrm{~K}=0, \ldots, \mathrm{~N}-1 \tag{4.14}
\end{equation*}
$$

where

$$
\begin{align*}
& E\left(X_{0}\right)=\bar{X}_{0}  \tag{4.14a}\\
& E\left(\omega_{K}\right)=0  \tag{4.14b}\\
& E\left(X_{0}-\bar{X}_{0}\right)\left(X_{0}-\bar{X}_{0}\right)^{T}=P_{0}  \tag{4.14c}\\
& E\left(\omega_{i}, \omega_{j}^{T}\right)=Q_{i j}  \tag{4.14d}\\
& E\left(\omega_{K}\right)\left(X_{0}-\bar{X}_{0}\right)^{T}=0 \tag{4.14e}
\end{align*}
$$

Measurements $Z_{K}$ are made at stage $K$, and are related to the state $X_{K}$ by

$$
\begin{equation*}
\mathrm{Z}_{\mathrm{K}}=\mathrm{HX}_{\mathrm{K}}+\mathrm{V}_{\mathrm{K}} \quad \mathrm{~K}=0, \ldots, \mathrm{~N} \tag{4.15}
\end{equation*}
$$

where

$$
\begin{align*}
& E\left(V_{K}\right)=0  \tag{4.15a}\\
& E\left(V_{i} V_{j}^{T}\right)=R_{i j}  \tag{4.15b}\\
& E\left(\omega_{i} V_{j}^{T}\right)=0  \tag{4.15c}\\
& E\left(X_{0}-\bar{X}_{0}\right) V_{K}^{T}=0 \tag{4.15d}
\end{align*}
$$

It is assumed that the state transition matrix $\phi$, input noise $\omega$, system noise $Q$, and observation noise $R$ are independent of stage K .

Numerous studies are available concerning least-square curve fitting, variational techniques, and maximum likelihood estimations to deal with the filiering problems. In cases where an additive input noise $\ddot{H}_{\underline{n}}$, as well as the output measurement error $\mathrm{V}_{\mathrm{K}}$ exist, variable approaches have resulted in the following sets of estimation equations from either deterministic or statistical considerations.

The optimum state estimate $\mathrm{X}_{\mathrm{K}+1}$ satisfies

$$
\begin{equation*}
\mathrm{X}_{\mathrm{K}+1}=\phi \mathrm{X}_{\mathrm{K}}+\mathrm{P}_{\mathrm{K}+1} \mathrm{~S}^{-1}\left(\mathrm{Z}_{\mathrm{K}+1}-\mathrm{H} \phi \mathrm{X}_{\mathrm{K}}\right) \tag{4.16}
\end{equation*}
$$

and the equation for the corariance of the estimation error $P_{K+1}$ is

$$
\begin{equation*}
P_{K+1}=M_{K+1}-H_{K+1} E^{T}\left[H r_{K+1} H^{T}+R\right]^{-1} H_{K+1} \tag{4.16a}
\end{equation*}
$$

where $M_{K+l}$ is defined by

$$
\begin{equation*}
M_{K+1}=\phi \mathrm{P}_{\mathrm{K}} \phi^{T}+\Gamma Q^{-T} \tag{4.16b}
\end{equation*}
$$

The initial conditions are

$$
\begin{equation*}
x\left(K_{0}\right)=x_{0} \tag{4.16c}
\end{equation*}
$$

and

$$
\begin{equation*}
P\left(K_{0}\right)=P_{0} \tag{4.16d}
\end{equation*}
$$

As shown in (4.16a) and (4.16b), $\mathrm{P}_{\mathrm{K}}$ is independent of the measurement $Z_{1}$; thus, the covariance matrix can be precom-
puted and stored to carry out the computation of the updated estimate in (4.16) with the current measurement.

In the case where $\therefore, P, R, Q$, and $H$ are all constant matrices, the filtering process may reach a steady state in the sense that $M_{K}$ and $P_{K}$ become constant matrices, $M$ and $P$, as $K \rightarrow \infty$. The two matrix equations for determining $\because$ and $P$ are

$$
\begin{align*}
& P^{-1}=M^{-1}+H^{T} R^{-1}  \tag{4.17}\\
& \text { or } \quad P=M-M H^{T}\left(R+M^{-1}\right)^{-1} F 1  \tag{4.17a}\\
& \text { and } \quad M=\phi P \phi^{T}+\Gamma Q \Gamma^{T}
\end{align*}
$$

An algebraic nonzecursive expression for $P_{K}$ is dexived in Section 4.4. The new technigue based on the uncoupling algorithm developed in Section 2.2 provides the determination of the transient solution $a 亡$ any particular stage without accumulating iteration from a giren starting value. Furthermore, the steady-state solution is immediate and does not require the information about the eigenralue and eigenvector of the system.

### 4.4 Reduced Discrete Zinear Systems

In Section 3.3, a linear regulator problem was solved when the set of continuous system equations was coupled. An analogous technique to the continuous case is developed for a solution to a discrete linear regulator. An uncoupling algorithm for canonical system mȧrices associated with the discrete optimal controls and the Kalman filters has been previously discussed in detail. An algebraic expression for the solution matrix to a Riccati difference equation is presented as a subsidiary
of the procedure.
The coupled set of difference equations for the optimization of the discrete linear regulator was derived in section 4.2

$$
\left[\begin{array}{l}
\mathrm{X}  \tag{4.18}\\
\lambda
\end{array}\right]_{\mathrm{K}+1}=\mathrm{H}_{\mathrm{DR}}\left[\begin{array}{l}
\mathrm{X} \\
\lambda
\end{array}\right]_{\mathrm{K}}
$$

The uncoupling of (4.18) may be preceded by introducing a set of new vectors $Y_{K}$ and $\omega_{K}$ by the relation

$$
\left[\begin{array}{l}
\mathrm{X}  \tag{4.19}\\
\lambda
\end{array}\right]_{\mathrm{K}}=\overline{\mathrm{V}}\left[\begin{array}{l}
\mathrm{Y} \\
\omega
\end{array}\right]_{\mathrm{K}}=\left[\begin{array}{ll}
\mathrm{v}_{11} & \mathrm{v}_{12} \\
\mathrm{v}_{21} & \mathrm{v}_{22}
\end{array}\right]\left[\begin{array}{l}
\mathrm{Y} \\
\omega
\end{array}\right]_{\mathrm{K}}
$$

where the transformation matrix $\bar{V}$ has the structure given by (2.19a). Substituting (4.19) for (4.18) and premultiplying both sides by $V^{-1}$ gives

Using the result of the uncoupling algorithm developed in Section 2.4 , it follows that the uncoupled system equations for the discrete linear regulator will be

$$
\left[\begin{array}{l}
\mathrm{Y}  \tag{4.20a}\\
\omega
\end{array}\right]_{\mathrm{K}+1}=\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Dl}} & 0 \\
0 & \mathrm{H}_{\mathrm{Dl}}^{-\mathrm{T}}
\end{array}\right]\left[\begin{array}{l}
\mathrm{Y} \\
\mathrm{~W}
\end{array}\right]_{\mathrm{K}}
$$

The solution vectors for (4.20a) are then obtained as

$$
\begin{equation*}
Y_{K}=\phi_{11}\left(K_{0}, K\right) Y_{0} \tag{4.21}
\end{equation*}
$$

and

$$
\begin{equation*}
\omega_{\mathrm{K}}=\phi_{22}\left(\mathrm{~K}_{0}, \mathrm{~K}\right) \omega_{0} \tag{4.2la}
\end{equation*}
$$

The uncoupled system equaぇions in (4.21) and (4.2la) adjoin each other and have a set of properties which are directly analogous to the properties listed For a continuous system in Section 3.3. The transition matrices ${ }^{\prime} 11$ and $\phi_{22}$ have the relationship

$$
\begin{equation*}
\phi_{11}\left(K_{0}, K\right)=\left[\phi_{22}^{-1}\left(E_{0}, K\right)\right]^{T} \tag{4.22}
\end{equation*}
$$

where

$$
\begin{equation*}
\phi_{11}\left(K_{0}, K\right)=H_{D 1}^{K} \tag{4.22a}
\end{equation*}
$$

and

$$
\begin{equation*}
\phi_{22}\left(\mathrm{~K}_{0}, \mathrm{~K}\right)=\left(\mathrm{H}_{\mathrm{DI}}^{-\mathrm{T}}\right){ }^{Z} \tag{4.22b}
\end{equation*}
$$

In a development similar to the continuous case, the unknown initial conditions $\Psi_{0}$ and $\ddot{O}_{0}$ can be determined by the use of (4.19) and the terminal condition given by (4.6).

$$
\begin{equation*}
V_{21} Y_{N}+V_{22}^{(1)} N=F\left[V_{11} Y_{Y}+V_{12} \omega_{N}\right] \tag{4.23}
\end{equation*}
$$

Solving for $Y_{N}$ yields

$$
\begin{equation*}
Y_{N}=\bar{H} \omega_{N} \tag{4.23a}
\end{equation*}
$$

where $H$ is a final-value weighting matrix for the uncoupled system and is provided by

$$
\begin{equation*}
\overline{\mathrm{H}}=\left[\mathrm{V}_{21}-\mathrm{HV} 1\right]^{-1}\left[\mathrm{Hv} 12^{-\mathrm{V}} 22^{]} .\right. \tag{4.23b}
\end{equation*}
$$

Considering the solution vectors $Y_{K}$ and $\omega_{K}$ at final stage N, a constant factor betreen uncoupled initial conditions is determined from (4.23a).

$$
\begin{equation*}
Y_{0}=s_{0} \omega_{0} \tag{4.24}
\end{equation*}
$$

where

$$
\begin{equation*}
S_{0}=\bar{H}_{D I}^{-N} \bar{H}\left(H_{D I}^{-N}\right)^{T} \tag{4.24a}
\end{equation*}
$$

Then $Y_{0}$ and $\omega_{0}$ may be 0 bたニinea in terms of the given value $X_{D}$

$$
\begin{equation*}
u_{0}=\left(v_{11} S_{0}+v_{12}\right)^{-1} x_{0} \tag{4.25}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathrm{x}_{0}=\mathrm{S}_{0}^{(1)} 0 \tag{4.25a}
\end{equation*}
$$

Since $S_{0}$ can be calculated so that the $N$ powers of $H_{D l}^{-1}$ can be built up by squaring $H_{D l}^{-1}$ and squaring the results, etc., the new initial conditions do not cause computational difficulties.

At each stage, the desired solution vectors for the coupled system can be reconstructed using the relation in (4.19).

$$
\begin{align*}
{\left[\begin{array}{l}
\mathrm{X} \\
\lambda
\end{array}\right]_{\mathrm{K}} } & =\left[\begin{array}{ll}
\mathrm{V}_{11} & \mathrm{~V}_{12} \\
\mathrm{~V}_{21} & \mathrm{~V}_{22}
\end{array}\right]\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Dl}}^{\mathrm{K}} & 0 \\
0 & \left(\mathrm{H}_{\mathrm{Dl}} \mathrm{~T}\right) \mathrm{K}
\end{array}\right]\left[\begin{array}{l}
\mathrm{Y}_{0} \\
\omega_{0}
\end{array}\right] \\
& =\left[\begin{array}{ll}
\mathrm{V}_{11} & \mathrm{~V}_{12} \\
\mathrm{~V}_{21} & \mathrm{~V}_{22}
\end{array}\right]\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{Dl}}^{-(\mathrm{N}-\mathrm{K})} \overline{\mathrm{H}\left(\mathrm{H}_{\mathrm{DI}}^{-N}\right)} \mathrm{T} & 0 \\
& 0
\end{array}\right]\left[\begin{array}{l}
\omega_{0} \\
\\
\\
\left.\omega_{0}^{-\mathrm{T}}\right)^{\mathrm{K}}
\end{array}\right]\left[\begin{array}{l}
\omega_{0}
\end{array}\right] \tag{4.26}
\end{align*}
$$

Since the eigenvalues of the submatrix $H_{D 1}$ are located outside the unit circle and the elements of (4.26) have only negative powers of $\mathrm{H}_{\mathrm{Dl}}$, the expression in (4.26) will be stable numerically throughout the whole stages.

For convenient reference, the necessary computational procedures, similar to those previously outlined in section 3.2, are summarized in the following steps:
a) Obtain the set of difference equations in (4.1) and (4.5) and construct the system matrix $H_{D R}$.
b) Use the sign algorithm for a discrete system outlined in Section 2.4 to obtain the transformation matrix $V$ and $\bar{V}$.
c) Compute Nth power of subblock $H_{D 1}$ where $N$ is the terminal time stage.
d) Obtain $\bar{H}$ given by (4.25b).
e) Compute $\mathrm{S}_{0}$ using (4.24a).
f) Compute initial conditions for uncoupled system.
g) Reconstruct the desired solution vectors $X_{K}$ and $\lambda_{K}$ by (4.26).

## 4.5 <br> An Algebraic Solution for a Discrete Riccati Equation

It has been reported that an algebraic solution to the discrete Riccati equation can be obtained by use of the eigenvalues and eigenvectors of the system matrix [45]. A non-recursive expression for a solution matrix $P_{K}$ made it possible to determine the steady-state solution of the Riccati equation as well as the transient solution at any particular time stage without iteration. The method reported, although conceptually simple, is not suitable for this study because of the complexity of finding the associated eigenvalues and eigenvectors.

In this section, a new procedure is presented which does not require either the iterational process or the burden of finding eigenvalues and eigenvectors. An algebraic expression for a feedback gain matrix of a discrete linear regulator is developed from the results described in section 4.3. In (4.9) and (4.19),

$$
\begin{equation*}
V_{21} Y_{K}+V_{22} \omega_{K}=P_{K}\left(V_{11} Y_{K}+V_{12} \omega_{K}\right) \tag{4.27}
\end{equation*}
$$

Using the solution vectors in (4.21) and (4.2la), a non-recursive expression is derived for $P_{K}$ as follows:

$$
\begin{equation*}
\mathrm{P}_{\mathrm{K}}=\left(\mathrm{V}_{21} \phi_{11} \mathrm{~S}_{0}+\mathrm{V}_{22}{ }_{22}\right)\left(\mathrm{V}_{11} \phi_{11} \mathrm{~S}_{0}+\mathrm{V}_{12} \phi_{22}\right)^{-1} \tag{4.28}
\end{equation*}
$$

From (2.28), it can be shown that

$$
\begin{equation*}
\lim _{\substack{\mathrm{N}, \mathrm{~K} \rightarrow \infty \\ \mathrm{~K} \leq \mathrm{N}}} \mathrm{P}_{\mathrm{K}}=\mathrm{V}_{22} \mathrm{~V}_{12}^{-1}=\mathrm{T}_{2}^{-1}=\mathrm{M}_{22^{\mathrm{M}_{1}}}^{-12} \tag{4.29}
\end{equation*}
$$

since

$$
\begin{align*}
& \lim _{\mathrm{N}, \mathrm{~K} \rightarrow \infty} \mathrm{~S}_{0}=0 \\
& \mathrm{~K}<\mathrm{N} \tag{4.29a}
\end{align*}
$$

The result of (4.29) is a steady-state solution consisting of eigenvectors of the system matrix corresponding to these eigenvalues which lie inside the unit circle.

A similar approach provides a nonrecursive solution for an error covariance matrix in the discrete Kalman filter. Since the initially known error covariance matrix in (4.16d) is dual to the terminal condition of the linear regulator given by (3.8c), the relation between the uncoupled initial vector is also dual in integration time.

From (4.9) and (4.19),

$$
\begin{equation*}
V_{21} Y_{0}+V_{22} \omega_{0}=P_{0}\left[V_{11} Y_{0}+V_{12} \omega_{0}\right] \tag{4.30}
\end{equation*}
$$

Assuming that

$$
\begin{equation*}
\omega_{0}=S_{0} Y_{0} \tag{4.30a}
\end{equation*}
$$

then

$$
\begin{equation*}
S_{0}=-\left[\mathrm{V}_{22}-\mathrm{P}_{0} \mathrm{~V}_{12}\right]^{-1}\left[\mathrm{~V}_{21}-\mathrm{P}_{0} \mathrm{~V}_{11}\right] \tag{4.30b}
\end{equation*}
$$

The linear relation between solution vectors $Y_{K}$ and $\omega_{K}$ can be written

$$
\begin{equation*}
\omega_{K}=S_{K} Y_{K} \tag{4.31}
\end{equation*}
$$

where

$$
\begin{equation*}
S_{K}=\left(\mathrm{H}_{\mathrm{Dl}}^{-\mathrm{T}}\right)^{\mathrm{K}} \mathrm{~S}_{0^{\mathrm{H}}}-\mathrm{Kl} \tag{4.31a}
\end{equation*}
$$

Substituting (4.19) and (4.31) for (4.9) yields

$$
\begin{equation*}
\mathrm{P}_{\mathrm{K}}=\left[\mathrm{V}_{21}+\mathrm{V}_{22} \mathrm{~S}_{\mathrm{K}}\right]\left[\mathrm{v}_{11}+\mathrm{V}_{12} \mathrm{~S}_{\mathrm{K}}\right]^{-\mathrm{I}} \tag{4.32}
\end{equation*}
$$

Equations (4.30b), (4.31a), and (4.32) constitute the desired nonrecursive algebraic solution for the error covariance matrix $P_{K}$. At any particular $K, P_{K}$ can be obtained only by computing $K$ powers of $H_{D T}^{-1}$.

The steady-state solution $P$ is obtained by letting $K \rightarrow \infty$.

$$
\begin{equation*}
P_{\infty}=\lim _{K \rightarrow \infty} P_{K}=V_{21}{ }_{I I}^{-I}=T=M_{21} M_{l l}^{-1} \tag{4.33}
\end{equation*}
$$

since $\mathrm{Fl}_{\mathrm{Dl}}$ corresponds to the eigenvalues outside the unit circle and $S_{K}$ contains only neçaive powers of $H_{D l}$,

$$
\begin{equation*}
S_{\infty}=\lim _{K \rightarrow \infty} S_{K}=\frac{1 i-}{K \rightarrow \infty}\left(H_{D I}^{-T}\right)^{K} S_{0} H_{D I}^{-K}=0 \tag{4.33a}
\end{equation*}
$$

The steady-state solution derived in (4.33) is analogous to that. of the continuous Riccati equation reported earlier [27,33].
4.6 Examples
4.6.1 Uncoupling Procecare for the Solution of a Discrete System

The uncoupled discrete obtained from Section 2.6.2 is solved completely by the outlined procedure in Section 4.4. Initial conditions and the terminal weighting matrix of the original problem are given as

$$
\begin{aligned}
& x_{0}=\left[\begin{array}{lll}
-2.0 & -1.0 & 2.0
\end{array}\right]^{\mathrm{T}}, \\
& H=\left[\begin{array}{lll}
0.5 & 0.1 & 0.3 \\
0.1 & 0.2 & 0.6 \\
0.3 & 0.6 & 0.8
\end{array}\right] \text { and } \mathrm{N}=20 .
\end{aligned}
$$

The initial vectors for the reduced system are calculated as
$Y_{0}=[-0.3722-17$
$0.4568-17$
$0.3241-17]^{T}$
and

$$
W_{0}=[3.1803
$$

15.508
$-4.0186]^{T}$

The desired optimal states and optimal controls for the coupled system are tabulated in Table 4.1 .

Table 4.1
Optimal States and Controls for a Discrete Linear Regulator

| K | $\mathrm{X}_{1 \mathrm{~K}}$ | $\mathrm{X}_{2 \mathrm{~K}}$ | $\mathrm{X}_{3 \mathrm{~K}}$ | $\mathrm{U}_{\mathrm{K}}$ |
| :--- | :--- | :--- | :--- | :--- |
| 1 | 1.3904 | -0.6096 | -3.6952 | 1.6952 |
| 2 | -4.7513 | 1.2487 | 1.8517 | -3.0709 |
| 3 | -0.8559 | 0.1920 | 0.5497 | 1.9477 |
| 4 | 0.1830 | $-0.6054-01$ | -0.1354 | 0.5195 |
| 5 | $-0.6955-01$ | $0.1827-01$ | $0.5196-02$ | -0.1263 |
| 10 | $-0.1939-03$ | $0.4172-04$ | $0.1495-03$ | $0.6260-03$ |
| 15 | $0.3686-06$ | $-0.1050-06$ | $-0.4741-06$ | $-0.6190-06$ |
| 20 | $-0.1410-06$ | $-0.5022-07$ | $-0.1043-06$ | $-0.3052-07$ |

4.6.2 Solution for a Discrete Riccati Eruation of Error Co-

## variance

An initial navigation problem from Mehra [30] is adopted as a numerical example. The discrete Kalman filter is constructed from the system whose coefficient matrices are

$$
\phi=\left[\begin{array}{ccclc}
0.75 & -1.75 & -0.3 & 0 & -0.15 \\
0.09 & 0.91 & -0.0015 & 0 & -0.008 \\
0 & 0 & 0.95 & 0 & 0 \\
0 & 0 & 0 & 0.55 & 0 \\
0 & 0 & 0 & 0 & 0.905
\end{array}\right]
$$

$$
\begin{aligned}
& \Gamma=\left[\begin{array}{lll}
0 & 0 & 0 \\
0 & 0 & 0 \\
24.64 & 0 & 0 \\
0 & 0.835 & 0 \\
0 & 0 & 1.83
\end{array}\right] \quad H=\left[\begin{array}{lllll}
1 & 0 & 0 & 0 & 1 \\
& & & & \\
0 & 1 & 0 & 1 & 0
\end{array}\right] \\
& \Omega=\left[\begin{array}{lll}
0.25 & 0 & 0 \\
0 & 0.5 & 0 \\
0 & 0 & 0.75
\end{array}\right] \quad R=\left[\begin{array}{lll}
0.4 & 0 & \\
0 & & 0.6
\end{array}\right]
\end{aligned}
$$

and the initial error covariance matrix is given

$$
P_{0}=\left[\begin{array}{lllll}
0.1 & 0.2 & 0.3 & 0.4 & 0.5 \\
0.2 & 0.1 & 0.2 & 0.3 & 0.4 \\
0.3 & 0.2 & 0.1 & 0.2 & 0.3 \\
0.4 & 0.3 & 0.2 & 0.1 & 0.2 \\
0.5 & 0.4 & 0.3 & 0.2 & 0.1
\end{array}\right]
$$

Two programs were used to calculate the error covariance matrix, one based on the new non-recursive technique in Section 4.5 and the other based on the conventional recursive solution of the associated discrete Riccati equation. The results from both methods are compared to show the maximum error in the 12 th and l3th significant digits in double precision on the IBM 360-44. The diagonal elements of the computed $P_{K}$ are presented in Table 4.2.

Table 4.2
Solution Matrix of a Discrete Riccati Equation

| $K$ | $P_{11}(K)$ | $P_{22}(K)$ | $P_{33}(K)$ | $P_{44}(K)$ | $P_{55}(K)$ |
| :--- | ---: | ---: | ---: | ---: | ---: |
| 1 | -0.4949 | -0.11372 | 151.67 | 0.30929 | 2.3724 |
| 5 | 24.038 | 0.21626 | 310.28 | 0.42450 | 6.8506 |
| 10 | 24.741 | 0.66461 | 320.13 | 0.46808 | 8.1641 |
| 15 | 24.810 | 0.68089 | 321.32 | 0.47168 | 8.2999 |
| 20 | 24.821 | 0.68212 | 321.33 | 0.47174 | 8.3214 |
| 25 | 24.822 | 0.68231 | 321.34 | 0.47176 | 8.3219 |
| 30 | 24.822 | 0.68232 | 321.34 | 0.47176 | 8.3220 |

Table 4.2 shows the convergence of $P_{K}$ to the steady-state solution with increasing K . The steady-state solution for the Riccati equation is obtained from the partitions of the transformation matrix used; i.e.r

$$
\mathrm{P}_{\infty}=\mathrm{V}_{21} \mathrm{~V}_{11}^{-1}=\left[\begin{array}{ccccc}
24.822 & 1.6052 & -59.609 & -0.21033 & -6.9568 \\
1.6052 & 0.68232 & -3.2971 & -0.16927 & -1.5205 \\
-59.609 & -3.2971 & 321.34 & 0.67997 & 9.122 \\
-0.21033 & -0.16927 & 0.67997 & 0.47176 & 0.26422 \\
-6.9568 & -1.5205 & 9.122 & 0.26422 & 8.322
\end{array}\right]
$$

NUMERICAL INTEGRATION OF A STIFF SYSTEM

### 5.1 The Set of Uncoupled Differential Equations

A new algorithm is developed for the numerical integration of linear stiff systems with constant coefficients having eigenvalues which differ greatly in magnitude. The intermediate size eigenvalues which may still be troublesome in either of the two sets are brought into consideration. When the new algorithm of the matrix filter generation in Section 2.5 is applied, the resulting state equations are completely uncoupled, and standard integration techniques can be utilized.

The differential equations for the eigenvalue subset matrices can now be generated by considering the first-order differential equations

$$
\begin{equation*}
\dot{X}(t)=A X(t)+B E(t) \tag{5.1}
\end{equation*}
$$

with initial conditions

$$
\begin{equation*}
x\left(t_{0}\right)=x_{0} \tag{5.la}
\end{equation*}
$$

It is assumed that the eigenvalues in the same order of magnitude are isolated into subsets and $X_{1}, \ldots, X_{l}$ are the corresponding uncoupled state vectors

$$
\begin{equation*}
x(t)=x_{1}(t)+x_{2}(t)+\ldots+x_{\ell}(t) \tag{5.2}
\end{equation*}
$$

where $\ell$ is the number of filters or subsets. In addition, $S_{i}$ denotes a matrix filter of the order $n \times n$
generated from the algorithm in Section 2.5.

$$
S_{i}=M\left[\begin{array}{cccccc}
0 & 0 & . & . & . & 0  \tag{5.3}\\
0 & 0 & . & . & \cdot & 0 \\
. & . & I & . & . & 0 \\
\cdot & . & 0 & . & . & . \\
. & . & 0 & . & . & 0
\end{array}\right] M^{-1}
$$

where the partitioned identity matrix is located at the ith column and row and where

$$
\begin{equation*}
\sum_{i=1}^{\ell} s_{i}=I \tag{5.4}
\end{equation*}
$$

The vectors $X_{i}(t)$ in (5.2) then become

$$
\begin{equation*}
x_{i}(t)=s_{i} x(t) \tag{5.5}
\end{equation*}
$$

so that

$$
\begin{equation*}
x(t)=\sum_{i=1}^{\ell} x_{i}(t)=\sum_{i=1}^{\ell} s_{i} x(t) \tag{5.6}
\end{equation*}
$$

To derive the desired differential equations in uncoupled form, the first differentiated equation (5.5) with respect to $t$ is

$$
\begin{equation*}
\dot{x}_{i}(t)=S_{i} \dot{X}(t)=S_{i}[A X(t)+B f(t)] \tag{5.7}
\end{equation*}
$$

The matrix $S_{i}$ will commute with $A$, thus

$$
\begin{equation*}
\dot{x}_{i}(t)=A S_{i} X(t) \div S_{i} B £(t) \tag{5.7a}
\end{equation*}
$$

which finally gives

$$
\begin{equation*}
\dot{x}_{i}(t)=A X_{i}(t)+S_{i} B f(t) \tag{5.7b}
\end{equation*}
$$

The terms of ( 5.7 b ) are multiplied through by the identity matrix, and (5.4) is used along with (5.5) to simplify as shown below:

$$
\begin{align*}
\dot{x}_{i}(t) & =A\left(\sum_{i=1}^{\ell} S_{i}\right) X_{i}(t)+S_{i} B f(t) \\
& =A S_{i} x_{i}(t) \div A\left(\sum_{\substack{j=1 \\
j \neq i}}^{\ell} S_{j}\right) x_{i}(t) X(t)+S_{i} B f(t) \\
& =A S_{i} x_{i}(t) \div A\left(\sum_{\substack{j=1 \\
j \neq i}}^{\ell} S_{j}\right) S_{i}(t)+S_{i} B f(t) \tag{5.8}
\end{align*}
$$

Since $S_{j}(t) S_{i}(t)=0$ for $a l l j \neq i$, then

$$
\begin{align*}
\dot{x}_{i}(t) & =s_{i} A x_{i}(t) \div s_{i} B f(t)  \tag{5.8a}\\
& =A_{i} x_{i}(t) \div s_{i} B f(t)
\end{align*}
$$

Equation (5.8a) is the desired uncoupled differential equation, and the eigenvalues are completely isolated into the subsystem $\operatorname{matrix} A_{j}$.

The initial conditions for (5.8a) are developed from (5.5)

$$
\begin{equation*}
x_{i}(0)=s_{i} x(0) \tag{5.8b}
\end{equation*}
$$

The integration step size for each uncoupled state equation is determined in the usual manner.

Since $\left|R_{e} \lambda_{1}\right|<\left|R_{e}\right\rangle_{2}\left|<\ldots<\left|R_{e} \lambda_{n}\right|\right.$, the step size can be changed accordingly for each subset. The step size can be selected by taking the average trace for each subset matrix $A_{i}$.

### 5.2 Examples

For the first example, a simple network analysis problem given by Lee was selected and compared [26]. The state equations to be integrated are

$$
\left[\begin{array}{l}
\dot{x}_{1}(t) \\
\dot{x}_{2}(t)
\end{array}\right]=\left[\begin{array}{rr}
-1000 & 0 \\
2000 & -1
\end{array}\right]\left[\begin{array}{l}
x_{1}(t) \\
x_{2}(t)
\end{array}\right]+\left[\begin{array}{l}
1 \\
0
\end{array}\right] f(t)
$$

and the initial conditions are given as

$$
\left[\begin{array}{l}
x_{1}(0) \\
x_{2}(0)
\end{array}\right]=\left[\begin{array}{l}
0 \\
0
\end{array}\right]
$$

where

$$
X=\left[\begin{array}{l}
x_{1} \\
\\
x_{2}
\end{array}\right] A=\left[\begin{array}{rr}
-1000 & 0 \\
2000 & -1
\end{array}\right] B=\left[\begin{array}{l}
1 \\
0
\end{array}\right] f(t)=I(t)
$$

The eigenvalues of $A$ are $\lambda_{1}=-1000$ and $\lambda_{2}=-1$, thus widely spaced. The step size was taken to be $h_{1}=\frac{1}{5 \lambda_{1}}=0.0002$ during the transient time period due to a large eigenvalue, then changed to $h_{2}=\frac{1}{5 \lambda_{2}}=0.2$ for the subset of small eigenvalues.

The integration interval is taken to be $0<t<5$, and 75 integration steps were used. The high-and low-pass matrices generated by the algorithm developed in Section 2.5 are

$$
S_{\mathrm{HP}}=\left[\begin{array}{cc}
1.000000000 & -0.617958115-17 \\
-2.002002002 & -0.5551115123-14
\end{array}\right]
$$

and

$$
S_{I د P}=\left[\begin{array}{cc}
-0.1110223023-15 & 0.6179585115-17 \\
2.002002002 & 1.000000000
\end{array}\right]
$$

For the purpose of comparison, the matrix filters constructed from the Butterworth function are presented.

$$
\mathrm{S}_{\mathrm{HP}}=\left[\begin{array}{cc}
1 & 0 \\
-2.00200203 & 1.0-8
\end{array}\right]
$$

and

$$
S_{L P}=\left[\begin{array}{ll}
-1.00000002-16 & 0 \\
2.00200203 & 1.00000001
\end{array}\right]
$$

The solution vector $X_{2}$ for a high-pass matrix are compared in Table 5.1

Table 5.1
Solution Vector $\mathrm{X}_{2}$ for a High-Pass Matrix

| Integ <br> Step |  | t | Filter Theory | New Algorithm |
| :--- | :--- | :---: | :---: | :---: |
| 0 | 0.0 | 0.0 | Exact $\mathrm{X}_{2}$ |  |
| 10 | 0.002 | $2.26895019 \mathrm{E}-03$ | $2.268942174-03$ | $2.69894021 \mathrm{E}-03$ |
| 20 | 0.004 | $6.02667963 \mathrm{E}-03$ | $6.026679269-03$ | $6.02670013 \mathrm{E}-03$ |
| 30 | 0.006 | $9.97900485 \mathrm{E}-03$ | $9.979008405-03$ | $9.97900962 \mathrm{E}-03$ |
| 40 | 0.008 | $1.39507856 \mathrm{E}-02$ | $1.395079204-02$ | $1.39507949 \mathrm{~F}-02$ |
| 50 | 0.01 | $1.79183299 \mathrm{E}-02$ | $1.791834164-02$ | $1.79183483 \mathrm{E}-02$ |
| $* 55$ | 1.01 | 1.27082125 E 00 | 1.270821384 E 00 | 1.27083286 E 00 |
| 60 | 2.01 | 1.73174577 E 00 | 1.731745951 F 00 | 1.73175437 E 00 |
| 65 | 3.01 | 1.90131311 E 00 | 1.901313295 E 00 | 1.90131793 E 00 |
| 70 | 4.01 | 1.96369442 E 00 | 1.963694618 E 00 | 1.96369389 E 00 |
| 75 | 5.01 | 1.98664358 F 00 | 1.986643786 E 00 | 1.98664483 E 00 |

For the second example, a $4 \times 4$ matrix with widely spaced eigenvalues was selected to illustrate the procedure of the bandpass filter algorithm. The state equation was assumed to be of the form $\dot{X}(t)=A X(t)$ where $A=M \Lambda M^{-1}$ with

$$
\Lambda=\operatorname{diag}\left[\begin{array}{llll}
-1200 & -500 & -1.5 & -0.5]
\end{array}\right.
$$

and

$$
M=\left[\begin{array}{llll}
0 & 1 & 2 & 3 \\
1 & 1 & 2 & 3 \\
2 & 2 & 2 & 3 \\
2 & 3 & 3 & 3
\end{array}\right]
$$

which gives

$$
A=\left[\begin{array}{cccc}
-1.493+03 & 1.99+03 & -1.988+03 & 9.93+02 \\
2.106+03 & -1.609+03 & 1.6105+03 & -1.406+03 \\
4.206+03 & -3.209+03 & 3.2105+03 & -2.806+03 \\
2.710+03 & -1.215+03 & 1.218+03 & -1.8105+03
\end{array}\right]
$$

The initial conditions are given as

$$
\mathrm{x}^{\mathrm{T}}(0)=\left[\begin{array}{llll}
2 & 1 & -1 & -2
\end{array}\right]
$$

The eigenvalues were isolated so that $\lambda=-1200$ was contained in the high-pass filter, $\lambda=-500$ in the band-pass filter, and the remaining two in a low-pass filter. The submatrices for the uncoupled state equations are given below

$$
A_{1}=\left[\begin{array}{rrrr}
7.0 & -10.0 & 12.0 & -7.0 \\
6.0 & -9.0 & 10.5 & -6.0 \\
6.0 & -9.0 & 10.5 & -6.0 \\
10.5 & -15.0 & 18.0 & -10.5
\end{array}\right]
$$

$$
\mathrm{A}_{2}=\left[\begin{array}{llll}
-.15 \mathrm{E}+04 & .20 \mathrm{E}+04 & -.20 \mathrm{~F}+04 & .10 \mathrm{E}+04 \\
-.15 \mathrm{E}+04 & .20 \mathrm{E}+04 & -.20 \mathrm{E}+04 & .10 \mathrm{E}+04 \\
-.30 \mathrm{E}+04 & .40 \mathrm{E}+04 & -.40 \mathrm{E}+04 & .20 \mathrm{E}+04 \\
-.45 \mathrm{E}+04 & .60 \mathrm{E}+04 & -.60 \mathrm{E}+04 & .30 \mathrm{E}+04
\end{array}\right]
$$

and

$$
A_{3}=\left[\begin{array}{llll}
.0 & .0 & .0 & .0 \\
.36 \mathrm{E}+04 & -.36 \mathrm{E}+04 & .36+04 & -.24 \mathrm{E}+04 \\
.72 \mathrm{E}+04 & -.72 \mathrm{E}+04 & .72+04 & -.48 \mathrm{E}+04 \\
.72 \mathrm{E}+04 & -.72 \mathrm{E}+04 & .72+04 & -.48 \mathrm{E}+04
\end{array}\right]
$$

Eighteen iterations were programmed in the sign algorithm.
All calculations were in double precision on the UNIVAC ll08, and the filter matrices are correct to 16 digits which were presented in Chapter II. The trace of the matrices $A_{1}, A_{2}$, and $A_{3}$ were $-2.0,-500$, and -1200 , respectively. Integration step sizes of $0.05,0.005,0.0005$ were used in a fourth-order Runge-Kutta algorithm to calculate $X_{1}(t), X_{2}(t)$, and $X_{3}(t)$.

The calculated values for $X_{1}(t), X_{2}(t)$ and $X_{3}(t)$ are given in Tables 5.2 through 5.4 where $X_{i}^{T}(t)=\left[X_{I i}, X_{2 i}, X_{3 i}, X_{4 i}\right]$. Table 5.2

Solution Vector $X_{1}(t)$ for a Low-Pass Matrix

| $t$ | $x_{11}(t)$ | $x_{21}(t)$ | $x_{31}(t)$ | $x_{41}(t)$ |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 0 | 0 | 0 | 0 |
| 1 | 2.30040 | 4.11999 | 4.11999 | 3.45060 |
| 2 | 1.90855 | 3.01219 | 3.01219 | 2.86283 |
| 3 | 1.27212 | 1.94151 | 1.94151 | 1.90819 |
| 4 | 0.79714 | 1.20314 | 1.20314 | 1.19570 |
| 5 | 0.48919 | 0.73545 | 0.73545 | 0.73379 |

Table 5.3
Solution Vector $X_{2}(t)$ for a Band-Pass Matrix

| $t$ | $X_{12}(t)$ | $x_{22}(t)$ | $x_{32}(t)$ | $x_{42}(t)$ |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 2 | 2 | 4 | 6 |
| 0.01 | $0.55373 \mathrm{E}-02$ | $0.55373 \mathrm{E}-02$ | $0.11074 \mathrm{E}-01$ | $0.16611 \mathrm{E}-01$ |
| 0.02 | $0.15331 \mathrm{E}-04$ | $0.15331 \mathrm{E}-04$ | $0.30662 \mathrm{E}-04$ | $0.45993 \mathrm{E}-04$ |
| 0.03 | $0.42446 \mathrm{E}-0$ | $0.42446 \mathrm{E}-07$ | $0.84892 \mathrm{E}-07$ | $0.12733 \mathrm{E}-06$ |
| 0.04 | $0.11751 \mathrm{E}-09$ | $0.11751 \mathrm{E}-09$ | $0.23503 \mathrm{E}-09$ | $0.35254 \mathrm{E}-09$ |
| 0.05 | $0.32082 \mathrm{E}-12$ | $0.32054 \mathrm{E}-12$ | $0.64165 \mathrm{E}-12$ | $0.96076 \mathrm{E}-12$ |
| - | - | - | - | - |
| 5 | $-0.455 \mathrm{E}-14$ | $0.482 \mathrm{E}-14$ | $-0.908 \mathrm{E}-14$ | $-0.153 \mathrm{E}-13$ |

Table 5.4
Solution Vector $X_{3}(t)$ for a High-Pass Matrix

| $t$ | $X_{13}(t)$ | $x_{23}(t)$ | $X_{33}(t)$ | $X_{43}(t)$ |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 0 | -4 | -8 | -8 |
| 0.0025 | $-0.11 \mathrm{E}-16$ | -0.19914 | -0.39829 | -0.39829 |
| 0.0050 | $0.45 \mathrm{E}-18$ | $-0.99149 \mathrm{E}-02$ | $-0.19829 \mathrm{E}-02$ | $-0.19829 \mathrm{E}-02$ |
| 0.0075 | $0.10 \mathrm{E}-17$ | $-0.49363 \mathrm{E}-03$ | $-0.98726 \mathrm{E}-03$ | $-0.98726 \mathrm{E}-03$ |
| 0.0100 | $0.10 \mathrm{E}-17$ | $-0.24576 \mathrm{E}-04$ | $-0.49153 \mathrm{E}-04$ | $-0.49153 \mathrm{E}-04$ |
| 0.0125 | $0.10 \mathrm{E}-17$ | $-0.12235 \mathrm{E}-07$ | $-0.24471 \mathrm{E}-05$ | $-0.24471 \mathrm{E}-05$ |
| 0.0150 | $0.10 \mathrm{E}-17$ | $-0.60918 \mathrm{E}-07$ | $-0.12183 \mathrm{E}-06$ | $-0.12183 \mathrm{E}-06$ |
| 0.0175 | $0.10 \mathrm{E}-17$ | $-0.30329 \mathrm{E}-08$ | $-0.60658 \mathrm{E}-08$ | $-0.60658 \mathrm{E}-08$ |
| 0.0200 | $0.10 \mathrm{E}-17$ | $-0.15100 \mathrm{E}-09$ | $-0.30200 \mathrm{E}-09$ | $-0.30200 \mathrm{E}-09$ |

The solution vector $X(t)=X_{1}(t)+X_{2}(t)+X_{3}(t)$ is
given in Table 5.4 and agrees with the exact solution to 10 digits for whole intervals $0 \leq t \leq 5$.

The exact solution is obtained from known eigenvalues and eigenvectors.

Fable 5.5
Solution Vector for $X(t)=X_{1}(t)+X_{2}(t)+X_{3}(t)$

| $t$ | $x_{1}$ | $x_{2}$ | $x_{3}$ | $x_{4}$ |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 2 | 1 | -1 | -2 |
| 0.0025 | 0.58797 | 3.38504 | 3.75894 | 1.34318 |
| 0.0050 | 0.19403 | 3.17662 | 3.33087 | 0.51746 |
| 0.0075 | 0.09170 | 3.07998 | 3.12652 | 0.20712 |
| 0.0100 | 0.07288 | 3.05789 | 3.07134 | 0.12948 |
| 0.0125 | 0.07793 | 3.05924 | 3.06310 | 0.12268 |
| 0.0150 | 0.08977 | 3.06735 | 3.06692 | 0.13631 |
| 0.0175 | 0.10350 | 3.07736 | 3.07768 | 0.15572 |
| 0.0200 | 0.11772 | 3.08787 | 3.08796 | 0.17671 |
| 0.03 | 0.17469 | 3.13002 | 3.13002 | 0.26203 |
| 0.04 | 0.23061 | 3.17120 | 3.17120 | 0.34591 |
| 0.05 | 0.28540 | 3.21133 | 3.21133 | 0.42810 |
| - | - | - | - | - |
| 1.0 | 2.30040 | 4.11999 | 4.11999 | 3.45060 |
| - | - | - | - | - |
| 5.0 | 0.48919 | 0.73545 | 0.73545 | 0.73379 |

It should be noted that the vectors $X_{2}(t)$ and $X_{3}(t)$ reach limiting values set by the elements in $S_{2}$ and $S_{3}$ which should be zero. This problem can be alleviated by setting all elements in $S_{i}$ to zero when they are at machine limit.

## CHAPTER VI

## CONCLUSIONS

In this study, new algorithms for the integration of the coupled differential equations from the optimal control with quadratic criteria and optimal filtering have been developed. The matrix sign function made it possible to uncouple the set of differential equations so that the resulting computations for the state transition matrix are of the order $n / 2$ when the original system matrix $A$ is of the order $n \times n$. The algorithm required two subblocks of state transition matrices to solve the general forms of system equations. However, certain important classes of system optimization problems require only one halfsized subblock due to the particular form of system coefficient matrices. This advantage saves considerable computational efforts particularly when the system order is high.

The new method permitted the state and costate equations from a quadratic linear regulator to be integrated simultaneously forward in time, thereby; the storage problem was overcome in calculating the feedback control law of the more conventional method.

Test problems have been made in several cases and results show a reduction in both the computation time and the size of the program required. The results of the new method show satisfactory agreement with solutions from the more common methods.

This study also developed a numerical integration technique suitable for the solution of stiff state equations by isolating the eigenvalues with large magnitude. Instead of dealing with
the integration method itself, the new technique gradually reduced the stiffness during the integration process. This was accomplished by discarding the uncoupled solution set corresponding to the eigenvalues having the largest negative real parts as their contribution becomes negligible. The effect of the troublesome eigenvalues was therefore removed.

The sign of the complex eigenvalue space was shifted by adding $\delta I$ to the system matrix $A$, and the matrix filters were generated. The amount of the shifting factor $\delta$ was selected to isolate the troublesome eigenvalues. The presence of eigenvalues with large negative real parts is readily recognized by observing the main diagonal of the matrix $A$, since the trace of $A$ is the sum of the real part of the eigenvalues.

Filter matrices cenerated by the new algorithm uncoupled the solution sets to the accuracy attainable from the machine. The advantage of the band-pass filter was demonstrated in an example problem.

In this paper the uncoupling procedure has been applied only to the linear time-invariant case. The same approach could be extended to the time-varying and nonlinear stiff equations with an increased number of calculations. The generation of the filter matrix will be more complicated by periodically repeating the computation of a matrix sign function of a time-varying system matrix or the Jacobian matrix and the uncertainty of the selection of a shifting factor $\delta$. This process should be undertaken whenever there is an abrupt change in the coefficients of the state equations.

This work suggests further studies in the following areas:
a) Extension of the sign algorithm to the time-varying matrix
b) Application of the uncoupling algorithm to the system optimization with time-varying system coefficients
c) Development of integration techniques to solve the time varying stiff equations using the matrix filter algorithm
d) Improvement of the computation of the matrix sign function
[1] Anderson, B. D. O., "An Algebraic Solution to the Spectral Factorization Problem," IEEE Trans. on Auto. Contr., Vol. AC-12, No. 4, August 1967.
[2] Athans, M., "The Status of Optimal Control Theory and Applications for Deterministic Systems," IEEE Trans. on Auto. Contr., 580-9, July 1966.
[3] Athans, M. and Falb, P. L., "Optimal Control," McGrawHill Book Co., New York, N. Y., 1966 .
[4] Beavers, A. N., Jr. and Denman, E. D., "A New Algorithm for Eigenvalues and Eigenvectors," Seventh Asilomar Circuits and Syst. Conf., Pacific Grove, Calif., 1973.
[5] Beavers, A. N., Jr. and Denman, E. D., "A New Solution Method for Quadratic Equations," to appear in Math. Biosci.
[6] Beavers, A. N., Jr. and Denman, E. D., "Asymptotic Solution to the Matrix Riccati Equation," to appear in Math. Biosci.
[7] Bellman, R. E. and Kalaba, R. E., "Dynamic Programming and Modern Control Theory," Academic Press, New York, 1965.
[8] Branin, F. H., Jr., "Computer Methods of Network Analysis," Proc. IFEE, Vol. 55, NO. 11, 1787-801, November 1967.
[9] Bryson, A. E., Jr. and Ho, Y. C., "Applied Optimal Control," Blaisdell Pub. Co., London, Eng., $\overline{1969 .}$
[10] Calahan, D. A., "A Stable, Accurate Method of Numerical Integration for Nonlinear Systems," (letter), Proc. IEEE, April, 1968, Vol. 56, 744, 1968.
[11] Certaine, T., "The Solution of Ordinary Differential Equations with Large Time Constants," Math. Methods for Digital Computers," Chap. 11, p. 128, Wiley, New York, 1960.
[12] Dahlquist, G., "Stability Questions for Some Numerical Methods for Ordinary Differential Equations," Proc. of Symp. App. Math., Vol. 15, 1963.
[13] Davison, E. J., "The Numerical Solution of Large Systems of Linear Differential Fquations," AICHE J., Vol. 14 , No. 1, Jan. 1968.
[14] Davison, E. J. and Maki, li. C., "The Numerical Solution of the Matrix Differential Equation," IFEE Trans. Auto. Cont., Vol. AC-18, 1973.
[15] Dukikar, K. L. and Denman, E. D., "Existence of Correct Initial Value Using Convergence Properties of Riccati Equation," Fourth Asilomar Conference on Circuits and Systems, Pacific Grove, Calif., Nov. 1970.
[16] Fowlex, M. E. and Warten, R. M., "A Numerical Integration Technique for Ordinary Differential Equations with Widely Separated Eigenvalues," IBM Jour., 537-43, Sept. 1967.
[17] Gear, C. W., "The Automatic Integration of Ordinary Differential Equations," Comm. of the ACM, Vol. 14, 176-9, 1971.
[18] Gear, C. W., "The Automatic Integration of Stiff Ordinary Differential Equations," Proc. IFIF Congress, Vol. I, 187-93, 1968.
[19] Kalman, R. E., "Contributions to the Theory of Optimal Control," Bol. Soc. Mat. Mex., 102-119, 1960.
[20] Kalman, R. E., "The Theory of Optimal Control and the Calculus of Variations," Mathematical Optimization Techniques, Bellman, R. E., ed., The Rand Co., Santa Monica, Calif., 1963.
[21] Kalman, R. E. and Busy, R. S., "New Results in Linear Filtering and Prediction Theory," Trans. ASME, J. Basic Engrg., Vol. 33, 95-107, March 1961.
[22] Kalman, R. F. and Engler, T. S., "A Users Manual for the Automatic Synthesis Programs," NASA Report CR-475, June 1966.
[23] Kirk, D. E., "Optimal Control Theory, An Introduction," Prentice-hall Inc., Englewood Cliffs, New Jersey, 1970.
[24] Klopfenstein, R. W. and Davis, C. B., "PECE Algorithm for the Solution of Stiff Systems of Ordinary Differential Equations," Math. Comp., Vol. 25, No. 115, 457-73, July 1971.
[25] Lawson, J. D., "Generalized Runge-Kutta for Stable Systems with Large Lipschitz Constants," SIAM J. Numer. Anal., Vol. 4, No. 3, 1967.
[26] Lee, H. B., "Matrix Filtering as an Aid to Numerical Integration," Proc. IEEE, Vol. 55, No. 11, 1826-31, Nov. 1967.
[27] MacFarlane, A. G. J., "An Eigenvector Solution of the Optimal Linear Regulator Problem," J. Electronics and Control, Vol. 14, June 1963.
[28] Martensson, K., "On the Matrix Riccati Equation," Information Services, Vol. 3, 17-49, 1971.
[29] Meditch, J. S., "Stochastic Optimal Linear Estimation and Control," McGraw-Hill Book Co., New York, N. Y., 1969.
[30] Mehra, R. K, "On the Identification of Variance and Adaptive Kalman Filtering," IEEE Trans. Auto. Cont., Vol. AC-15, No. 2, 175-84, April 1970.
[31] Musial, W., "A Numerical Investigation of the p-Equations Relating the Matrix Riccati Equation to a Linear State Equation," M.S. Thesis, U. of Houston, 1972.
[32] Pearson, J. B., Jr. and Sridhar, R., "A Discrete Optimal Control Problem," IEEE Trans. Auto. Cont., Vol. AC-ll, 171-4, April 1966.
[33] Potter, J. E., "Matrix Quadratic Solution," SIAM J. nppl. Math., Vol. 14, No. 3, May 1966.
[34] Rao, H. S. and Denman, E. D., "A Forward Integration Mlgorithm for the Linear Regulator Problem," SWIEECO Conf., Dallas, Texas, 1970.
[35] Richard S., Peschard, J., Thibault, R., and Aubé, G., "OnIine Reduction for Computer Analysis of High-Order Electrical Circuits with Widely Separated Time Constants," IEEE Trans. Circuit Theory, Vol. CT-19, No. 6, Nov. 1972.
[36] Richards, P. I., Lanning, W. D., and Torrey, M. D., "Numerical Integration of Large, Highly Damped, Nonlinear Systems," SIAM Rev., Vol. 7, 376-80, 1965.
[37] Rosenbrock, H. H., "Some General Implicit Processes for the Numerical Solution of Differential Equations," Computer J.5, 3020-30, 1963.
[38] Rybichi, G. G. and Usher, P. D., "The Generalized Riccati Transformation as a Simple Alternative to Invariant Imbedding," The Astrophysical Journal, Vol. 146, No. 3, Dec. 1966.
[39] Sage, A. P., "Optimal Systems Controls," Prentice-Hall, Inc., Englewood Cliffs, N. J. 1968 .
[40] Seinfeld, J. H., Lapidus, L. and Hwang, M., "Review of Numerical Integration Techniques for Stiff Ordinary Differential Equations," Ind. Fng. Chem. Fundam., Vol. 9, No. 2, 1970 .
[41] Shichman, II., "Integration System of a Nonlinear Network Analysis Program," IEEE Trans. Circuit Theory, Vol. CT17, No. 3, Aug. 1970.
[42] Sineman, R. W., "Digital Time Domain Analysis of Systems with Widely Separated Doles," Assoc. Comp. Mach. J., Vol. 12, 286-93, April 1965.
[43] Treanor, C. E., "A Method for the Numerical Integration of Coupled First-Order Differential Fquations with Greatly Different Time Constants," Math. Comp., Vol. 20, 39-45, 1966.
[44] Vaughan, D. R., "A Negative Exponential Solution for the Matrix Riccati Equation," IEEE Trans. Auto. Cont., Vol. AC-14, 72-5, Feb. 1969.
[45] Vaughan, D. R., "A Non-Recursive Algebraic Solution for the Discrete Riccati Equation," IEEE Trans. Auto. Cont., Vol. AC-15, Oct. 1970.

## APPENDIX A

## MATRIX SIGN FUNCTION

Consider a square matrix $A$ of the order $N$ and

$$
\begin{equation*}
A=M \Lambda M^{-1} \tag{A.1}
\end{equation*}
$$

where $M$ is the eigenvector matrix of $A$. Suppose that $p$ of the N eigenvalues are positive and n are negative so that $\mathrm{p}+\mathrm{n}=\mathrm{N}$. The matrix function sign $A$ is defined by

$$
\operatorname{sign} A=M\left[\begin{array}{cc}
I_{p \times p} & 0  \tag{A.2}\\
0 & -I_{n \times n}
\end{array}\right] \quad M^{-1}
$$

The sign $A$ is calculated by the algorithm

$$
\begin{equation*}
(\operatorname{sign} A)^{i+1}=\frac{1}{2}(\operatorname{sign} A)^{i}+\frac{1}{2}\left((\operatorname{sign} A)^{i}\right)^{-1} \tag{A.3}
\end{equation*}
$$

with $(\operatorname{sign} A)^{0}=A$.
The iterative process is terminated when trace[(sign $A)^{i+1}$ ]
$=\operatorname{trace}\left[(\operatorname{sign} A)^{\mathbf{i}}\right]$ to the desired accuracy.

## Asymptotic Solution of Riccati Equation

Consider the Riccȧi equation

$$
\begin{equation*}
A+B P+P C+P D P=0 \tag{B.1}
\end{equation*}
$$

The matrix of coefficients takes the form

$$
\bar{A}=\left[\begin{array}{rr}
B & A  \tag{B.2}\\
-D & -C
\end{array}\right]
$$

Let the Jordan form of $\vec{A}$ be

$$
\left[\begin{array}{cc}
M_{11} & M_{12}  \tag{B.3}\\
M_{21} & M_{22}
\end{array}\right]\left[\begin{array}{ccc}
\Lambda_{1} & 0 & \vdots \\
0 & \Lambda_{2} & -D \\
\hline & -C
\end{array}\right]\left[\begin{array}{cc}
B & A \\
M_{21} & M_{22}
\end{array}\right]
$$

Equating terms after the block multiplication gives

$$
\begin{array}{ll}
M_{11} \Lambda_{1}=\mathrm{BM}_{11}+\mathrm{AH}_{21} & \mathrm{M}_{12} \therefore_{2}=\mathrm{BM}_{12}+\mathrm{AM}_{22} \\
M_{21} \Lambda_{1}=-\mathrm{DM}_{11}-\mathrm{CM}_{21} & \mathrm{M}_{22} \mathrm{~N}_{2}=-\mathrm{DM}_{12}-\mathrm{CM}_{22} \tag{B.4}
\end{array}
$$

Performing multiplications on both sides, assuming the $M_{2 l}^{-1}$ and $M_{22}^{-1}$ exist,

$$
\begin{align*}
& M_{11} \Lambda_{1} M_{2 l}^{-1}=B_{11} M_{21}^{-1}+A \\
& M_{11} \Lambda_{1} M_{2 l}^{-1}=-M_{11} M_{21}^{-1} \ddots_{11} 1_{21}^{-1}-M_{11} M_{21}^{-1} C \tag{B.5a}
\end{align*}
$$

and

$$
M_{12} \Lambda_{2} M_{22}^{-1}=B_{12}^{M} M_{22}^{-1}+A
$$

$$
M_{12} \Lambda_{2} M_{22}^{-1}=-M_{12} M_{22}^{-1} D_{12} M_{22}^{-1}-M_{12} M_{22}^{-1} C
$$

Subtracting the lower equation from the upper in (B.5a) and (B.5b) gives

$$
\begin{align*}
& \mathrm{A}+\mathrm{BM}_{11} \mathrm{M}_{21}^{-1}+\mathrm{M}_{11} \mathrm{M}_{21}^{-1} \mathrm{C}+\mathrm{M}_{11}{ }^{M_{21}^{-1}} \mathrm{DM}_{11} \mathrm{M}_{21}^{-1}=0 \\
& \mathrm{~A}+\mathrm{BM}_{12} \mathrm{M}_{22}^{-1}+\mathrm{M}_{12} 2_{22}^{-1} \mathrm{C}+\mathrm{M}_{1} 2^{M_{2}^{-1}} 2^{-1} M_{12} \mathrm{M}_{22}^{-1}=0 \tag{B.6}
\end{align*}
$$

which are the Riccati equations with $P=M_{12} M_{21}^{-1}$ and $P=M_{12} M_{22}^{-1}$.

