Springer Series in Information Sciences 18

Editor: Thomas S. Huang

Springer Series in Information Sciences

Editors: Thomas S. Huang Teuvo Kohonen Manfred R. Schroeder
Managing Editor: H. K. V. Lotsch

Volume 1 **Content-Addressable Memories** By T. Kohonen 2nd Edition

Volume 2 **Fast Fourier Transform and Convolution Algorithms**
By H. J. Nussbaumer 2nd Edition

Volume 3 **Pitch Determination of Speech Signals** Algorithms and Devices
By W. Hess

Volume 4 **Pattern Analysis** By H. Niemann

Volume 5 **Image Sequence Analysis** Editor: T. S. Huang

Volume 6 **Picture Engineering** Editors: King-sun Fu and T. L. Kunii

Volume 7 **Number Theory in Science and Communication**
With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity By M. R. Schroeder 2nd Edition

Volume 8 **Self-Organization and Associative Memory** By T. Kohonen
2nd Edition

Volume 9 **Digital Picture Processing** An Introduction By L. P. Yaroslavsky

Volume 10 **Probability, Statistical Optics and Data Testing**
A Problem Solving Approach By B. R. Frieden

Volume 11 **Physical and Biological Processing of Images**
Editors: O. J. Braddick and A. C. Sleigh

Volume 12 **Multiresolution Image Processing and Analysis**
Editor: A. Rosenfeld

Volume 13 **VLSI for Pattern Recognition and Image Processing**
Editor: King-sun Fu

Volume 14 **Mathematics of Kalman-Bucy Filtering**
By P. A. Ruymgaart and T. T. Soong 2nd Edition

Volume 15 **Fundamentals of Electronic Imaging Systems**
Some Aspects of Image Processing By W. F. Schreiber

Volume 16 **Radon and Projection Transform-Based Computer Vision**
Algorithms, A Pipeline Architecture, and Industrial Applications
By J. L. C. Sanz, E. B. Hinkle, and A. K. Jain

Volume 17 **Kalman Filtering** with Real-Time Applications
By C. K. Chui and G. Chen

Volume 18 **Linear Systems and Optimal Control**
By C. K. Chui and G. Chen

C. K. Chui G. Chen

Linear Systems and Optimal Control

With 4 Figures

Springer-Verlag Berlin Heidelberg New York
London Paris Tokyo

Professor Dr. Charles K. Chui

Department of Mathematics and Department of Electrical Engineering,
Texas A & M University, College Station, TX 77843, USA

Dr. Guanrong Chen

Department of Electrical and Computer Engineering,
Rice University, Houston, TX 77251, USA

Series Editors:

Professor Thomas S. Huang

Department of Electrical Engineering and Coordinated Science Laboratory,
University of Illinois, Urbana, IL 61801, USA

Professor Teuvo Kohonen

Department of Technical Physics, Helsinki University of Technology,
SF-02150 Espoo 15, Finland

Professor Dr. Manfred R. Schroeder

Drittes Physikalisches Institut, Universität Göttingen, Bürgerstrasse 42–44,
D-3400 Göttingen, Fed. Rep. of Germany

Managing Editor: Helmut K.V. Lotsch

Springer-Verlag, Tiergartenstrasse 17,
D-6900 Heidelberg, Fed. Rep. of Germany

ISBN-13: 978-3-642-64787-1 e-ISBN-13: 978-3-642-61312-8
DOI: 10.1007/978-3-642-61312-8

Library of Congress Cataloging-in-Publication Data. Chui, C.K. Linear systems and optimal control. (Springer series in information sciences ; 18) Bibliography: p. Includes index. 1. Control theory. 2. Optimal control. I. Chen, G. (Guanrong) II. Title. III. Series. QA402.3.C5566 1988 629.8'312 88-2012

Softcover reprint of the hardcover 1989

Typesetting: Macmillan India Ltd., India

2154/3150-543210 – Printed on acid-free paper

Preface

A knowledge of linear systems provides a firm foundation for the study of optimal control theory and many areas of system theory and signal processing. State-space techniques developed since the early sixties have been proved to be very effective. The main objective of this book is to present a brief and somewhat complete investigation on the theory of linear systems, with emphasis on these techniques, in both continuous-time and discrete-time settings, and to demonstrate an application to the study of elementary (linear and nonlinear) optimal control theory.

An essential feature of the state-space approach is that both time-varying and time-invariant systems are treated systematically. When time-varying systems are considered, another important subject that depends very much on the state-space formulation is perhaps real-time filtering, prediction, and smoothing via the Kalman filter. This subject is treated in our monograph entitled "Kalman Filtering with Real-Time Applications" published in this Springer Series in Information Sciences (Volume 17). For time-invariant systems, the recent frequency domain approaches using the techniques of Adamjan, Arov, and Krein (also known as AAK), balanced realization, and H^∞ theory via Nevanlinna-Pick interpolation seem very promising, and this will be studied in our forthcoming monograph entitled "Mathematical Approach to Signal Processing and System Theory". The present elementary treatise on linear system theory should provide enough engineering and mathematics background and motivation for study of these two subjects.

Although the style of writing in this book is intended to be informal, the mathematical argument throughout is rigorous. In addition, this book is self-contained, elementary, and easily readable by anyone, student or professional, with a minimal knowledge of linear algebra and ordinary differential equations. Most of the fundamental topics in linear systems and optimal control theory are treated carefully, first in continuous-time and then in discrete-time settings. Other related topics are briefly discussed in the chapter entitled "Notes and References". Each of the six chapters on linear systems and the three chapters on optimal control contains a variety of exercises for the purpose of illustrating certain related view-points, improving the understanding of the material, or filling in the details of some proofs in the text. For this reason, the reader is encouraged to work on these problems and refer to the "answers and hints" which are included at the end of the text if any difficulty should arise.

This book is designed to serve two purposes: it is written not only for self-study but also for use in a one-quarter or one-semester introductory course in linear systems and control theory for upper-division undergraduate or first-year graduate engineering and mathematics students. Some of the chapters may be covered in one week and others in at most two weeks. For a fifteen-week semester, the instructor may also wish to spend a couple of weeks on the topics discussed in the "Notes and References" section, using the cited articles as supplementary material.

The authors are indebted to Susan Trussell for typing the manuscript and are very grateful to their families for their patience and understanding.

College Station
Texas, May 1988

Charles K. Chui
Guanrong Chen

Contents

1. State-Space Descriptions

Although the history of linear system theory can be traced back to the last century, the so-called state-space approach was not available till the early 1960s. An important feature of this approach over the traditional frequency domain considerations is that both time-varying and time-invariant linear or nonlinear systems can be treated systematically. The purpose of this chapter is to introduce the state-space concept.

1.1 Introduction

A typical model that applied mathematicians and system engineers consider is a "machine" with an "input-output" relation placed at the two terminals (Fig. 1.1). This machine is also called a *system* which may represent certain biological, economical, or physical systems, or a mathematical description in terms of an algorithm, a system of integral or differential equations, etc. In many applications, a system is described by the totality of input-output relations (u, v) where u and v are functions or, when discretized, sequences, and may be either scalar or vector-valued. It should be emphasized that the collection of all input-output ordered pairs is not necessarily single-valued. As a simple example, consider a system given by the differential equation $v''+v=u$. In this situation, the totality of all input-output relations that determines the system is the set

$$S=\{(u, v): v''+v=u\}$$

and it is clear that the same input u gives rise to infinitely many outputs v. For example, $(1, \sin t+1)$, $(1, \cos t+1)$, and even $(1, a\cos t+b\sin t+1)$ for arbitrary constants a and b, all belong to S. To avoid such an unpleasant situation and to give a more descriptive representation of the system, the "state" of the system is considered. The *state* of a system explains its past, present, and future situations. This is done by introducing a minimum number of variables which are called *state variables* that represent the present situation, using the past information, namely the initial state, and describe the future behavior of the system completely. The column vector of the state variables, in a given order, is called a *state vector*.

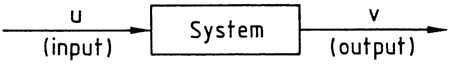

Fig. 1.1

Let us return to the simple example of the system described by the differential equation $v''+v=u$ with a specified initial state. Introducing the state vector

$$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ,$$

where x_1 and x_2 are state variables satisfying the initial state $x_1(a)=b$ and $x_2(a)=c$, we can give a "state-space" description of this system by using a system of two equations:

$$\dot{\boldsymbol{x}}=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\boldsymbol{x}+\begin{bmatrix} 0 \\ 1 \end{bmatrix}u$$
$$v=[1\ \ 0]\boldsymbol{x} , \tag{1.1}$$

where $\dot{\boldsymbol{x}}$ denotes the derivative of the state vector $\boldsymbol{x}$. The definition of *state-space* will be better understood later in Sect. 1.4. Here, the first equation in (1.1) gives the input-state relation while the second equation describes the state-output relation. The so-called *state-space equations* (1.1) could be obtained by setting the state variables x_1 and x_2 to be v and v' respectively. However, without the knowledge of such substitutions, it may not be immediately clear that the input-output relation follows from the state-space equations (1.1). To demonstrate how this is done more generally, we rewrite (1.1) as

$$\dot{\boldsymbol{x}}=A\boldsymbol{x}+Bu$$
$$v=C\boldsymbol{x} \tag{1.2}$$

where A, B, C are 2×2, 2×1, 1×2 matrices and let $p(\lambda)$ be the characteristic polynomial of A. In this example, $p(\lambda)=\lambda^2+1$, so that by the Cayley–Hamilton Theorem, we have

$$p(A)=A^2+I=0 .$$

Hence, differentiating the second equation in (1.2) twice (the number of times of differentiation will equal the degree of the characteristic polynomial of the square matrix A), and utilizing the first equation in (1.2) repeatedly, we have

$$C\boldsymbol{x}=v$$

$$CA\boldsymbol{x}=v'-CBu$$

$$CA^2\boldsymbol{x}=v''-CBu'-CABu .$$

Therefore, the identity $p(A) = A^2 + I = 0$ can be used to eliminate $\boldsymbol{x}$, yielding:

$$(v'' - CBu' - CABu) + v = CA^2\boldsymbol{x} + C\boldsymbol{x} = C(A^2 + I)\boldsymbol{x} = 0 \quad \text{or}$$

$$\begin{aligned} v'' + v &= C(Bu' + ABu) \\ &= [1 \ \ 0] \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} u' + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \right) \\ &= u \ . \end{aligned}$$

1.2 An Example of Input-Output Relations

More generally, if the characteristic polynomial of an $n \times n$ matrix A in an input-state equation such as (1.2) is

$$p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \ldots + a_n \ ,$$

then the above procedure gives

$$\begin{aligned} &C\boldsymbol{x} = v \\ &CA\boldsymbol{x} = v' - CBu \\ &CA^2\boldsymbol{x} = v'' - CBu' - CABu \\ &\ldots \\ &CA^n\boldsymbol{x} = v^{(n)} - CBu^{(n-1)} - CABu^{(n-2)} - \ldots - CA^{n-1}Bu \ , \end{aligned}$$

so that, by setting $a_0 = 1$, we have:

$$\sum_{k=0}^{n} a_k \left(v^{(n-k)} - C \sum_{j=0}^{n-k-1} A^j B u^{(n-k-j-1)} \right) = Cp(A)\boldsymbol{x} = 0 \ .$$

That is, the input-output relation can be given by

$$\sum_{j=0}^{n} a_j v^{(n-j)} = C \sum_{k=0}^{n} a_k \sum_{j=0}^{n-k-1} A^j B u^{(n-k-j-1)} \tag{1.3}$$

with $a_0 = 1$.

A slightly more general form of (1.3) is given by

$$Lv = Mu$$

$$L = \sum_{j=0}^{n} a_j \frac{d^{n-j}}{dt^{n-j}} \ , \quad a_0 = 1 \tag{1.4}$$

$$M = \sum_{k=0}^{m} b_k \frac{d^{m-k}}{dt^{m-k}} \ , \quad m \leq n \ .$$

However, the system with input-output relations described by (1.4) does not necessarily have a state-space description given by (1.2) (Exercise 1.2). We also remark in passing that even if it has such a description, the matrices A, B and C are not unique (Exercise 1.3).

1.3 An Example of State-Space Descriptions

A more general state-space description of a system with input-output pairs (u, v) is given by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ v &= Cx + Du \end{aligned} \tag{1.5}$$

where A, B, C, D are matrices with appropriate dimensions. By eliminating the state vector x and its derivative with the help of the Cayley-Hamilton Theorem as above, it is not difficult to see that the input-output pair (u, v) in (1.5) satisfies the relation $Lv = Mu$ in (1.4) with appropriate choices of constants a_j and b_k (Exercise 1.4). To see the converse, that is, to show that the input-output relations in (1.4) have a state-space description as given in (1.5), we follow the standard technique of transforming an nth order linear differential equation to a first order vector differential equation as was done in the simple example discussed earlier by choosing the matrix A to be

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & 1 \\ -a_n & \cdots & & -a_2 & -a_1 \end{bmatrix}.$$

Of course there are other choices of A. But with this "so-called" standard choice, it is clear that the matrix C must be given by

$$C = [1 \;\; 0 \ldots 0] \;.$$

Hence, by setting $B = [\beta_1 \;\ldots\; \beta_n]^T$ and $D = [\beta_0]$ we see that the variables of the vector $x = [x_1 \;\ldots\; x_n]^T$ in (1.5) satisfy the equations:

$$\begin{aligned} x_1' &= x_2 + \beta_1 u \\ x_2' &= x_3 + \beta_2 u \\ &\cdots \\ x_{n-1}' &= x_n + \beta_{n-1} u \\ x_n' + a_1 x_n + &\ldots + a_n x_1 = \beta_n u \\ v &= x_1 + \beta_0 u \;. \end{aligned}$$

That is, the state variables are defined by

$$x_1 = v - \beta_0 u$$

$$x_2 = x_1' - \beta_1 u = v' - (\beta_0 u' + \beta_1 u)$$

$$x_3 = x_2' - \beta_2 u = v'' - (\beta_0 u'' + \beta_1 u' + \beta_2 u)$$

$$\dots$$

$$x_n = x_{n-1}' - \beta_{n-1} u = v^{(n-1)} - (\beta_0 u^{(n-1)} + \dots + \beta_{n-1} u)$$

and must satisfy the constraint:

$$x_n' + a_1 x_n + \dots + a_n x_1 = \beta_n u \ ,$$

or equivalently,

$$\sum_{j=0}^{n} a_j v^{(n-j)} = \left(\sum_{i=0}^{n} a_i \beta_{n-i}\right) u + \left(\sum_{i=0}^{n-1} a_i \beta_{n-i-1}\right) u' + \dots + (a_1 \beta_0 + a_0 \beta_1) u^{(n-1)} + a_0 \beta_0 u^{(n)} \ . \tag{1.6}$$

Hence, the constants $\beta_0, \dots, \beta_n$ are uniquely determined by the linear matrix equation

$$\begin{bmatrix} a_0 & a_1 & \dots & a_n \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ 0 & \dots & 0 & a_0 \end{bmatrix} \begin{bmatrix} \beta_n \\ \vdots \\ \beta_0 \end{bmatrix} = \begin{bmatrix} b_m \\ \vdots \\ b_{m-n} \end{bmatrix}$$

where $a_0 = 1$ and $b_j = 0$ for $j < 0$. We remark that the highest derivative of u in (1.6) is n, and hence the order m of the differential operator M in (1.4) is not allowed to exceed n.

1.4 State-Space Models

A system with the state-space description given by (1.5) is usually called a single-input/single-output *time-invariant* system; that is, the matrices A, B, C and D in (1.5) are constant matrices and the input and output functions are scalar-valued. In general, we have to work with *time-varying* systems, and in addition, the input and output functions may happen to be vector-valued; in other words, we may have a multi-input/multi-output system. The state-space description of such a system is given by

$$\begin{aligned} \dot{\boldsymbol{x}} &= A(t)\boldsymbol{x} + B(t)\boldsymbol{u} \\ \boldsymbol{v} &= C(t)\boldsymbol{x} + D(t)\boldsymbol{u} \ . \end{aligned} \tag{1.7}$$

The digital version of (1.7) is

$$\begin{aligned} \boldsymbol{x}_{k+1} &= A_k \boldsymbol{x}_k + B_k \boldsymbol{u}_k \\ \boldsymbol{v}_k &= C_k \boldsymbol{x}_k + D_k \boldsymbol{u}_k \ , \end{aligned} \tag{1.8}$$

where $\{\boldsymbol{u}_k\}$ and $\{\boldsymbol{v}_k\}$ are input and output sequences of the discretized (or digital) system, respectively. Of course (1.8) is only an approximation of (1.7), for instance, by setting $\boldsymbol{u}_k = \boldsymbol{u}(kh)$, $\boldsymbol{v}_k = \boldsymbol{v}(kh)$, and $\boldsymbol{x}_k = \boldsymbol{x}(kh)$ where h is a sampling time unit. A natural choice of the matrices A_k, B_k, C_k and D_k is given by

$$A_k = hA(kh) + I$$

$$B_k = B(kh)$$

$$C_k = C(kh) \quad \text{and}$$

$$D_k = D(kh) \ .$$

A small sampling time unit is necessary to give a good approximation. We will be dealing with the state-space descriptions (1.7, 8) for continuous-time and discrete-time systems, respectively. The vector space, spanned by the state vectors which are generated by all "admissible" inputs and initial states, is called the *state-space*. For a better understanding, see Exercises 2.2–4.

It will be clear from Exercise 2.5 that the outputs in the state-space descriptions (1.7, 8) are linear in the state vectors for zero input and linear in the inputs for zero initial state. For this reason, the systems we consider here are called *linear systems*. In the subject of *control theory*, linear systems are also called *linear dynamic systems*, the state-space descriptions (1.7, 8), *dynamic equations*, and the matrices $A(t)$, $B(t)$, $C(t)$, and $D(t)$ in (1.7) or A_k, B_k, C_k, and D_k in (1.8) are called *system* (or *dynamic*), *control*, *observation* (or *output*), and *transfer matrices*, respectively.

Exercises

1.1 Give a state-space description for the input-output relations $v'' + av' + bv = u$ by using the state variables $x_1 = \alpha v + \beta v'$ and $x_2 = \gamma v + \delta v'$ where $\alpha\delta - \beta\gamma \neq 0$.

1.2 Determine all constants a, b and c so that the linear system with input-output relations $v'' + v' = au + bu' + cu''$ has a state-space description of the form given by (1.2).

1.3 By using Exercise 1.1, show that the matrices A, B, and C in the state-space description (1.2) for the linear system with input-output relations $v'' + av' + bv = 0$ are not unique.

1.4 Determine the constants a_j and b_k in (1.4) for the input-output relations of the linear system (1.5) where A, B, C and D are arbitrary $n \times n$, $n \times 1$, $1 \times n$, and 1×1 matrices.

1.5 (a) Give a state-space description for the two-input and two-output system

$$v_1'' + a_{11} v_1' + a_{12} v_1 + b_{11} v_2' + b_{12} v_2 = \alpha_1 u_1 + \beta_1 u_2$$

$$v_2'' + a_{21} v_1' + a_{22} v_1 + b_{21} v_2' + b_{22} v_2 = \alpha_2 u_1 + \beta_2 u_2 \ .$$

(b) Derive a general state-space description for the normal n-input and n-output system

$$v_1^{(n)} + \sum_{j=1}^{n} \{a_{1j}^1 v_1^{(n-j)} + a_{1j}^2 v_2^{(n-j)} + \ldots + a_{1j}^n v_n^{(n-j)}\} = \sum_{j=1}^{n} \alpha_{1j} u_j \ .$$

$$\ldots$$

$$v_n^{(n)} + \sum_{j=1}^{n} \{a_{nj}^1 v_1^{(n-j)} + a_{nj}^2 v_2^{(n-j)} + \ldots + a_{nj}^n v_n^{(n-j)}\} = \sum_{j=1}^{n} \alpha_{nj} u_j \ .$$

1.6 (a) Give a state-space description for the discrete-time system defined by the difference equation

$$v_{k+2} + v_{k+1} + v_k = u_k \ .$$

$\left(\textit{Hint}: \text{Let } x_{1,k} = v_k, x_{2,k} = v_{k+1} \text{ and } \mathbf{x}_k = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} \right).$

(b) Derive a general state-space description for the discrete-time system defined by the difference equation

$$a_0 v_{k+n} + a_1 v_{k+n-1} + \ldots + a_n v_k = b_0 u_{k+m} + \ldots + b_m u_k \ ,$$

where $a_0 = 1$, $m \leq n$, and m, n are arbitrary positive integers.

2. State Transition Equations and Matrices

In this chapter, we will discuss the solution of the state-space equation assuming that the initial state as well as all the governing matrices are given. Both continuous-time and discrete-time systems will be considered. It is clear that only the input-state equation has to be solved.

2.1 Continuous-Time Linear Systems

From the theory of ordinary differential equations, if $A(t)$ is an $n \times n$ matrix whose entries are continuous functions on an interval J which contains t_0 in its interior, then the initial value problem

$$\begin{aligned} \dot{\boldsymbol{x}} &= A(t)\boldsymbol{x} \\ \boldsymbol{x}(t_0) &= \boldsymbol{e}_i \ , \end{aligned} \tag{2.1}$$

where $\boldsymbol{e}_i = [0 \ldots 0 \ 1 \ 0 \ldots 0]^{\mathrm{T}}$, the entry 1 being the ith component, has a unique solution which we will denote by $\phi_i(t, t_0)$. Let $\Phi(t, t_0)$ be the $n \times n$ matrix with $\phi_i(t, t_0)$ as its ith column. Since these column vectors are linearly independent, the "fundamental matrix" $\Phi(t, t_0)$ is nonsingular. For convenience, we assume that J is an open interval. Since the above discussion is valid for any t_0 in J, we could consider $\Phi(s, t)$ as a matrix-valued function of two variables in J. Clearly,

$$\Phi(t, t) = I \ ,$$

the identity matrix, for all t in J. Set

$$F(s, t) = \Phi(s, \tau)\Phi^{-1}(t, \tau) \ .$$

Then $F(s, \tau) = \Phi(s, \tau)\Phi^{-1}(\tau, \tau) = \Phi(s, \tau)$, i.e., $F \equiv \Phi$, so that

$$\Phi(s, t) = \Phi(s, \tau)\Phi^{-1}(t, \tau)$$

or, equivalently, $\Phi(s, t)$ satisfies the "transition" property:

$$\Phi(s, \tau) = \Phi(s, t)\Phi(t, \tau) \ , \tag{2.2}$$

where s, t, and τ are in J.

We now consider the input-state equation with a given initial state $\boldsymbol{x}_0$ at time t_0, namely

$$\begin{aligned} \dot{\boldsymbol{x}} &= A(t)\boldsymbol{x} + B(t)\boldsymbol{u} \\ \boldsymbol{x}(t_0) &= \boldsymbol{x}_0 \ , \end{aligned} \tag{2.3}$$

where $A(t)$ and $B(t)$ are $n \times n$ and $n \times p$ matrices respectively, and $\boldsymbol{u}$ is a p-dimensional column vector. Although weaker conditions are allowed, we will always assume, for convenience, that all entries of $A(t)$ are continuous functions on J and that the entries of $B(t)$ as well as the components of $\boldsymbol{u}$ are piecewise continuous on J. Again from the theory of ordinary differential equations, (2.3) has a unique solution given by

$$\boldsymbol{x}(t) = \Phi(t, t_0)\boldsymbol{x}(t_0) + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)\boldsymbol{u}(\tau)d\tau \ , \tag{2.4}$$

where, as usual, integration is performed componentwise, and $\Phi(t, t_0)$ is the fundamental matrix of the first order homogeneous equation $\dot{\boldsymbol{x}} = A\boldsymbol{x}$ discussed above. In the subject of control theory, one could think of $\boldsymbol{u}$ as the control function that takes an initial state $\boldsymbol{x}(t_0)$ to a state $\boldsymbol{x}(t)$ in continuous time from time t_0 to time t, and "equation" (2.4) describes how this is done. Because of its formulation, this equation is also called the (continuous-time) integral equation of $\boldsymbol{u}$. Note that the solution of this equation for the control function $\boldsymbol{u}$ that takes $\boldsymbol{x}(t_0)$ to $\boldsymbol{x}(t)$ is given by the input-state equation (2.3). The matrix $\Phi(t, t_0)$ that describes this transition process is usually called the *transition matrix* of the linear system.

2.2 Picard's Iteration

In order to have a better understanding of the transition process, it is important to study the transition matrix. We first consider the special case where $A = [a_{ij}]$ is a constant matrix. Denote by $|A|_1$ the l^1 norm of this matrix; that is

$$|A|_1 = \sum_{i,j} |a_{ij}| \ .$$

By Exercise 2.8, we have $|A^2|_1 \leq |A|_1^2, \ldots, |A^n|_1 \leq |A|_1^n, \ldots$, and this allows us to define

$$\mathrm{e}^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$$

since the sequence of partial sums of the infinite series is a Cauchy sequence:

$$\left|\sum_{n=M}^{N} \frac{t^n}{n!} A^n\right|_1 \leq \sum_{n=M}^{N} \frac{|t|^n}{n!} |A^n|_1$$
$$\leq \sum_{n=M}^{N} \frac{(|t||A|_1)^n}{n!}$$

which tends to 0 as M and N tend to infinity independently. (Here, the triangle inequality in Exercise 2.8 has been used.) In addition, it is also clear from this infinite series definition that

$$\frac{d}{dt} e^{tA} = Ae^{tA} \ .$$

Hence, it follows immediately that the solution $\phi_i(t, t_0)$ of (2.1) is given by

$$\phi_i(t, t_0) = e^{(t-t_0)A} e_i \ ;$$

that is, the transition matrix in (2.4) for the system with constant system matrix A is given by

$$\Phi(t, t_0) = e^{(t-t_0)A} \ . \tag{2.5}$$

When $A = A(t)$ is not a constant, that is when time-varying state-space equations are considered, an explicit formulation of the transition matrix is usually difficult to obtain. The following iteration process, usually attributed to Picard, gives an approximation of $\Phi(t, t_0)$. Again, for convenience, we assume that the entries of $A(t)$ are bounded functions in J, so that a positive constant C exists with

$$|A(t)|_1 \leq C < \infty, \quad t \in J \ .$$

We start with the identity matrix. Set

$$P_0(t) = I$$

$$P_1(t) = I + \int_{t_0}^{t} A(s) P_0(s) ds$$

$$\cdots$$

$$P_N(t) = I + \int_{t_0}^{t} A(s) P_{N-1}(s) ds \ .$$

Then for all $t \in J$ and $N > M$, we have

$$|P_N(t) - P_M(t)|_1 = \left| \sum_{k=M}^{N-1} [P_{k+1}(t) - P_k(t)] \right|_1$$

$$= \left| \sum_{k=M}^{N-1} \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2) \ldots \int_{t_0}^{s_k} A(s_{k+1}) ds_{k+1} \ldots ds_1 \right|_1$$

$$\leq \sum_{k=M}^{N-1} \left| \int_{t_0}^{t} \ldots \int_{t_0}^{s_k} ds_{k+1} \ldots ds_1 \right| C^{k+1}$$

$$= \sum_{k=M}^{N-1} \frac{(C|t-t_0|)^{k+1}}{(k+1)!}$$

which tends to zero uniformly on any bounded interval as $M, N \to \infty$ independently. That is, $\{P_N(t)\}$ is a Cauchy sequence of matrix-valued continuously differentiable functions on J. Let $P(t, t_0)$ be its uniform limit. Since

$$\frac{d}{dt} P_N(t) = A(t) P_{N-1}(t)$$

and $P_N(t_0) = I$, it follows from a theorem of Weierstrass that

$$\frac{d}{dt} P(t, t_0) = A(t) P(t, t_0)$$

$$P(t_0, t_0) = I \ .$$

This, of course, means that the columns of $P(t, t_0)$ are the unique solutions $\phi_i(t, t_0)$ of the initial value input-state equations (2.1), so that $P(t, t_0)$ coincides with $\Phi(t, t_0)$. We have now described a simple iteration process that gives a uniform approximation of $\Phi(t, t_0)$. It also allows us to write:

$$\Phi(t, t_0) = I + \int_{t_0}^{t} A(s) ds + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2) ds_2 ds_1 + \ldots \ . \tag{2.6}$$

It is clear that if $A = A(t)$ is a constant matrix, then (2.5) and (2.6) are identical, using the definition of $\exp[(t-t_0)A]$.

2.3 Discrete-Time Linear Systems

We now turn to the discrete-time system. The input-state equation with a given initial state $\boldsymbol{x}_0$ is given by

$$\boldsymbol{x}_{k+1}=A_k\boldsymbol{x}_k+B_k\boldsymbol{u}_k, \quad k=0, 1, \ldots, \tag{2.7}$$

where A_k and B_k are $n\times n$ and $n\times p$ matrices and $\boldsymbol{u}_k$, $k=0, 1, \ldots,$ are p-dimensional column vectors. Writing out (2.7) for $k=0, 1, \ldots,$ respectively, we have

$$\begin{aligned}
&\boldsymbol{x}_1=A_0\boldsymbol{x}_0+B_0\boldsymbol{u}_0\\
&\boldsymbol{x}_2=A_1\boldsymbol{x}_1+B_1\boldsymbol{u}_1\\
&\ldots\\
&\boldsymbol{x}_{k+1}=A_k\boldsymbol{x}_k+B_k\boldsymbol{u}_k\ .
\end{aligned}$$

Hence, by substituting the first equation into the second one, and this new equation into the third one, etc., we obtain

$$\boldsymbol{x}_N=\Phi_{N_0}\boldsymbol{x}_0+\sum_{k=1}^{N}\Phi_{Nk}B_{k-1}\boldsymbol{u}_{k-1} \tag{2.8}$$

where we have defined the "transition" matrices:

$$\begin{aligned}
&\Phi_{kk}=I\\
&\Phi_{jk}=A_{j-1}\ldots A_k \quad \text{for} \quad j>k\ .
\end{aligned} \tag{2.9}$$

In particular, if $A_k=A$ for all k, then $\Phi_{jk}=A^{j-k}$ for $j\geq k$. Equation (2.8) is called the (discrete-time) *state transition equation* corresponding to the input-state equation (2.7) and Φ_{jk} $(j\geq k)$ are called the *transition matrices.* The state transition equation describes the transition rule in discrete-time that the control sequence $\{\boldsymbol{u}_k\}$ takes the initial state $\boldsymbol{x}_0$ to the final state $\boldsymbol{x}_N$. We remark, however, that although the transition matrices Φ_{jk} satisfy the "transition" property

$$\Phi_{ik}=\Phi_{ij}\Phi_{jk} \quad \text{for} \quad i\geq j\geq k\ ,$$

Φ_{jk} is *not defined* for $j<k$, and in fact, even if $A_k=A$ for all k, Φ_{ik} $(i>k)$ is singular if A is. This shows that discrete-time and continuous-time linear systems may have different behaviors. However, if the system matrices $A_k, \ldots, A_{j-1}$, where $k<j$, are nonsingular, it is natural to introduce the notation $\Phi_{kj}=A_k^{-1}\ldots A_{j-1}^{-1}$, so that $\Phi_{kj}=\Phi_{jk}^{-1}$ or $\Phi_{kj}\Phi_{jk}=I$, completing the transition property.