كلية الدراسات العليا

Sudan University of Science and Technology
College of Graduate Studies

جامعة السودان للعلوم والتكنولوجيا

Transfer Functions and Mathematical Modeling of Dynamic Systems with MATLAB

الدوال الناقلة والنمذجة الرياضية للانظمة الحركية مع MATLAB

A Thesis submitted in Fulfillment Requirements for the Degree of Ph.D in Mathematics

By:

Musa Adam Abdallah Ahmed

Supervisor:

Abdelrahim Bashir Hamid

2018

Dedication

To My Family

Parent,

Brothers,

Sisters,

wife and son.

To My Friends

To Classmates (My batch)

Who always support me

And

Who always stand beside and help me

Acknowledgments

First, I thank Allah for guiding me and taking care of me all the time. My life is so blessed because of his majesty.

Thanks to me **supervisor Dr.** ABDELRAHIM BASHIR HAMID University of Gezira for his advice and he did not Stingy any information for me, Thanks to **Co. supervisor Dr.** SAMAR SAMIR her help in this research.

for encouraging and Also , I would like to thank my family especially my Parents supporting me all the time.

I wish to express my considerable gratitude to are best Friends.

Abstract

This study introduces the mechanical elements that are the main components in the dynamic modeling of mechanical systems. It discusses the electrical elements of voltage and current source, resistor, capacitor, inductor, and operational amplifier. We derive mathematical models for electrical systems. The transfer function, enables connecting the input to the output of a dynamic system into the Laplace domain for single-input, single output, multiple-input and multiple output systems. It is shown that how to derive the transfer function for SISO systems and the transfer function matrix for MIMO systems by starting from the system time-domain mathematical model or using complex impedances. The transfer function is further utilized to determine the free and forced time responses with zero initial conditions of mechanical and electrical. The concept of state space is an approach that can be used to model and determine the response of dynamic systems in the time domain. Using a vector-matrix formulation, the state space procedure is an alternate method of characterizing mostly MIMO systems defined by a large number of coordinates (DOFs). Different state space algorithms are applied, depending on whether the input has time derivatives or no time derivatives. Methods of conversion between state space and transfer function. MATLAB is quite useful to transform the system model from transfer function to state space, and vice versa. We also used the partial-fraction expansion of the ratio of two polynomials, block diagram algebra.

الخلاصة

تقدم هذه الدراسة العناصر الميكانيكية التي هي المكونات الرئيسية في النمذجة الديناميكية للأنظمة الميكانيكية. ويناقش العناصر الكهربائية للجهد والمصدر الحالي، يتم استخدام المقاوم، والمكثف، المغو، والمضخم العملياتي لاشتقاق نماذج رياضية للأنظمة الكهربائية. الدالة الناقلة، والتي تمكن من توصيل المدخلات إلى مخرجات النظام الديناميكي في مجال لابلاس من أجل المدخلات الفردية، والمخرجات الفردية، والأنظمة المتعددة المدخلات، والمتعددة المخرجات. أظهرت هذه الطريقة كيفية اشتقاق الدالة الناقلة لأنظمة SISO ومصفوفة الدالة الناقلة لأنظمة MIMO بالبدء من النموذج الرياضي للنطاق الزمني للنظام أو باستخدام المعوقات المعقدة. يتم استخدام الدالة الناقلة كذلك لتحديد الاستجابات الحرة للوقت مع عدم وجود شروط أولية ميكانيكية وكهربائية. مفهوم فضاء الحالة كنهج يمكن استخدامه لنمذجة وتحديد استجابة الأنظمة الديناميكية في المجال الزمني. باستخدام طريقة مصفوفة المتجهات، يكون إجراء الفضاء الخارجي طريقة بديلة لتوصيف معظم أنظمة MIMO المحددة بواسطة عدد كبير من الإحداثيات. (DOFs) يتم تطبيق خوارزميات مختلفة لحالة الفضاء، اعتمادًا على ما إذا كان الإدخال يحتوي على مشتقات زمنية أو لا توجد مشتقات زمنية. برنامج الحاسب الالي ماتلاب مفيد للغاية لتحويل نموذج النظام من الدوال الناقلة إلى فضاء الحالة، والعكس بالعكس. ولنا للحصول على التوسع الجزئي للنسبة بين اثنين من كثيرات الحدود، الجبر التخطيطي للكتلة.

Introduction

Engineering system dynamics is a discipline that focuses on deriving mathematical models based on simplified physical representations of actual systems, such as mechanical, electrical, fluid, or thermal, and on solving the mathematical models (most often consisting of differential equations).The resulting solution (which reflects the system response or behavior) is utilized in design or analysis before producing and testing the actual system. Because dynamic systems are characterized by similar mathematical models, a unitary approach can be used to characterize individual systems pertaining to different fields as well as to consider the interaction of systems from multiple fields as in coupled-field problems.

The main objective of this study is introduces the mechanical elements (inertia, stiffness, damping, and forcing) that are the main components in the dynamic modeling of mechanical systems. By using the lumped-parameter approach, you have learned to model the dynamics of basic, single degree-of-freedom mechanical systems. Newton's second law of motion is applied to derive the mathematical models for the natural, free damped, and forced responses of basic translatory and rotary mechanical systems. We discuss the electrical elements of voltage and current source, resistor, capacitor, inductor, and operational amplifier. Ohm's law, Kirchhoff's laws, the energy method, the mesh analysis method, and the node analysis method are applied to derive mathematical models for electrical systems.

The transfer function, which enables connecting the input to the output of a dynamic system into the Laplace domain for single-input, single output and multiple-input, multiple output systems. It is shown how to derive the transfer function for SISO systems and the transfer function matrix for MIMO systems by starting from the system time-domain mathematical model or using complex impedances. The transfer function is further utilized to determine the free and forced time responses with zero initial conditions of mechanical and electrical systems.

The concept of state space as an approach that can be used to model and determine the response of dynamic systems in the time domain. Using a vector-matrix formulation, the state space procedure is an alternate method of characterizing mostly MIMO systems defined by a large number of coordinates (DOFs). Different state space algorithms are applied, depending on whether the input has time derivatives or no time derivatives. Methods of calculating the free response with nonzero initial conditions and the forced response using the state space approach also are presented. Methods of converting between state space and transfer function.

MATLAB is quite useful to transform the system model from transfer function to state space, and vice versa. It is also used to obtain the partial-fraction expansion of the ratio of two polynomials, block diagram algebra, introduced for solving problems through symbolic calculation and plotting time responses.

Contents

Dedication	I
Acknowledgement	II
Abstract	III
Abstract Arabic	IV
Introduction	V
The Contents	VI
List of tables	IX
Chapter (1) **Mathematical Modeling**	
1-1 Definitions	1
1-2 Mathematical Modeling of Mechanical Systems	2
1-3 Mathematical Modeling of Electrical Systems	16
1-4 Analogous Systems	23
1-5 Linear Systems	24
1-6 Linear time-invariant and Linear time-Varying Systems	25
Chapter (2) **Transfer Function and Impulse Response Function**	
2-1 Transfer Function (Single-Input, Single-Output Systems)	26
2-2 Transfer Function (Multivariable Systems)	28
2-3 Stability of Linear Control Systems	29
2-4 Bounded-input, Bounded-output (BIBO Stability-Continuous-Data Systems	29
2-5 Relationship Between characteristic equation Roots and Stability	30
2-6 Zero-input and Asymptotic Stability of Continuous-Data Systems	31
2-7 Transfer Functions of Cascaded Elements	35
2-8 Impulse Response Function	44
Chapter (3) **Automatic Control Systems**	
3-1 Block Diagrams	50
3-1-1 Relation between Mathematical Equations and Block Diagrams	55
3-1-2 Block Diagram Reduction	58
3-1-3 Block Diagram of Multi-Input Systems-Special Case:	61
3-1-4 Block Diagrams and Transfer Functions of Multivariable Systems	62
3-2 Signal-flow Graph (SFGs)	64
3-3 Automatic controllers	73
Chapter (4) **Modeling in State Space**	
4-1 Modern Control Theory	79
4-1-1 Correlation Between Transfer Functions and State-Space Equations	83
4-2 Block Diagrams, Transfer Functions, and State Diagrams	86
4-2-1 Transfer Functions (Multivariable Systems)	86
4-2-2 Block Diagrams and Transfer Functions of Multivariable Systems	87
4-2-3 State Diagram	89
4-3 Vector-Matrix Representation of State Equations	93

4-4 State-Transition Matrix	95
4-5 State-Transition Equation	97
4-6 Relationship between State Equations and High-Order Differential Equation	102
4-7 Relationship Between State Equations and Transfer Functions	103
Chapter (5) **State Space representations of Dynamic Systems and Linearization of Nonlinear Mathematical Models**	
5-1 State Space representations of Dynamic Systems	105
5-2 State Space to Transfer Function	113
5-2 State Space to Transfer Function	114
5-3-1 Characteristic Equation from a Differential Equation	114
5-3-2 Characteristic Equation from a Transfer Function	115
5-3-3 Characteristic Equation from State Equations	115
5-3-4 Eigenvalues	115
5-3-5 Eigenvectors	116
5-3-6 Generalized Eigenvectors	117
5-4 Similarity Transformation	118
5-4-1 Invariance Properties of the Similarity Transformations	119
5-4-2 Controllability Canonical Form (CCF)	119
5-4-3 Observability Canonical Form (OCF)	121
5-4-4 Diagonal Canonical Form (DCF)	122
5-4-5 Jordan Canonical Form (JCF)	123
5-5 Decompositions of Transfer Function	124
5-5-1 Direct Decomposition	124
5-5-2 Cascade Decomposition	128
5-5-3 Parallel Decomposition	129
5-6 Nonuniqueness of a Set of State Variables	130
5-7 Linearizations of Nonlinear Mathematical Models	131
5-7-1 Nonlinear Systems	131
5-7-2 Linearization of Nonlinear Systems	131
5-7-3 Linearization Using Taylor Series: Classical Representation	132
5-7-4 Linearization Using the State Space Approach	132
5-7-5 Linear Approximation of Nonlinear Mathematical Models	133
Chapter (6) **Transformation of Mathematical Models with MATLAB**	
6-1 Transformation from Transfer Function to State Space Representations	137
6-2 Transformation from State Space Representations to Transfer Function	138
6-3 MATLAB Implementation	141
6-3-1 Series Connection	141
6-3-2 Parallel Connection	142
6-3-3 Feedback Connection	142
6-4 Partial-Fraction Expansion with MATLAB	142
6-5 Transient Response Analysis with MATLAB	144
6-5-1 Step Response	144
6-5-2 Impulse Response	147
6-5-3 Impulse Input	148

6-5-4 Obtaining response to arbitrary input	149
6-6 Block Diagram Algebra Using MATLAB	151
6-7 State-Variable models with MATLAB	152
6-7-1 LTI Objects And The ss(A,B,C,D) Function	153
6-7-2 The ss(sys) And ssdata(sys) Functions	153
6-7-3 Relating State Variables To The Original Variables	154
6-7-4 The `tfdata` Function	154
6-7-5 Linear ODE Solvers	157
6-7-6 The `initial` Function	157
6-7-7 The `impulse, step`, And `lsim` Functions	159
Chapter (6) **Existence results**	
7-1 Differential Equation to State Space	162
7-2 Methods for Electrical Systems Modeling	169
7-2-1 The Mesh Analysis Method	169
7-2-2 The Node Analysis Method	171
7-3 Free Response	172
7-3-1 Natural Response	172
7-3-2 Single-DOF Conservative Electrical Systems	173
Comments and conclusions	175
References	176

List of tables

TABLE 1-1 Basic Rotational Mechanical System Properties and Their Units	8
TABLE 1-2. Summary of elements involved in linear electrical systems	18
TABLE 1-3 Force Voltage Analogy	24
TABLE 2-1 Stability Conditions of Linear Continuous-Data Time-Invariant SISO Systems	32
TABLE 3-1 Block diagrams and their SFG equivalent representations	69
Table 6-1 summarizes these functions.	144
Table 6-2 summarizes these functions.	147
Table 6-3 summarizes these functions.	149
Table 6-4 LTI object functions.	156
Table 6-5 Basic syntax of linear solvers for state variable models.	160

Chapter One

Mathematical Modeling

1-1 Definitions

Before we can discuss control systems, some basic terminologies must be defined.

Definition 1.1: [6] A system is defined as a combination of components (elements) that act together to perform a certain objective. System dynamics deal with:

a) the mathematical modeling of dynamic systems and
b) response analysis of such systems with a view toward understanding the dynamic nature of each system and improving the system's performance.

A system is a combination of components that act together and perform a certain objective. A system need not be physical. The concept of the system can be applied to abstract, dynamic phenomena such as those encountered in economics. The word system should, therefore, be interpreted to imply physical, biological, economic, and the like, systems.

Definition 1.2: [6] Static Systems have an output response to an input that does not change with time.

Definition 1.3: [6] Dynamic Systems have a response to an input that is not instantaneously proportional to the input or disturbance and that may continue after the input is held constant. Dynamic systems can respond to input signals, disturbance signals, or initial conditions.

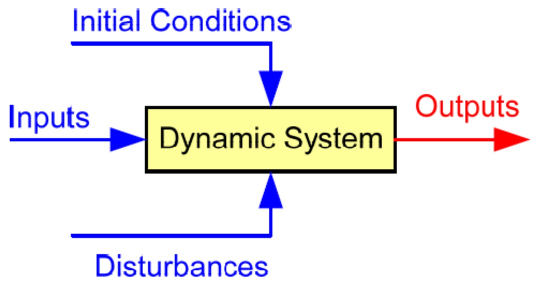


Figure 1-1 Excitation and response of a system

Definition 1-4: [6] Dynamic Systems may be observed in common devices employed in everyday living, Figure 1-2, as well as in sophisticated engineering systems such as those in spacecraft that took astronauts to the moon. Dynamic Systems are found in all major engineering disciplines and include mechanical, electrical, fluid and thermal systems.

Definition 1-5: [6] Controlled Variable and Control Signal or Manipulated Variable. The controlled variable is the quantity or condition that is measured and controlled. The control signal or manipulated variable is the quantity or condition that is varied by the controller so as to affect the value of the controlled variable. Normally, the controlled variable is the output of the system. Control means measuring the value of the controlled variable of the system and applying the control signal to the system to correct or limit deviation of the measured value from a desired value.

In studying control engineering, we need to define additional terms that are necessary to describe control systems.

Definition 1-6: [6] Plants. A plant may be a piece of equipment, perhaps just a set of machine parts functioning together, the purpose of which is to perform a particular operation. We shall call any physical object to be controlled (such as a mechanical device, a heating furnace, a chemical reactor, or a spacecraft) a plan.

Definition 1-7: [6] Processes. The Merriam-Webster Dictionary defines a process to be a natural, progressively continuing operation or development marked by a series of gradual changes that succeed one another in a relatively fixed way and lead toward a particular result or end; or an artificial or voluntary, progressively continuing operation that consists of a series of controlled actions or movements systematically directed toward a particular result or end.

Definition 1-8: [6] Disturbances. A disturbance is a signal that tends to adversely affect the value of the output of a system. If a disturbance is generated within the system, it is called internal, while an external disturbance is generated outside the system and is an input.

Definition 1-9: [6] Feedback Control. Feedback control refers to an operation that, in the presence of disturbances, tends to reduce the difference between the output of a system and some reference input and does so on the basis of this difference. Here only unpredictable disturbances are so specified, since predictable or known disturbances can always be compensated for within the system.

1-2 Mathematical Modeling of Mechanical Systems

Mathematical models. may assume many different forms. Depending on the particular system and the particular circumstances, one mathematical model may be better suited than other models. For example, in optimal control problems, it is advantageous to use state-space representations. On the other hand, for the transient-response or frequency-response analysis of single-input, single-output, linear, time-invariant systems, the transfer-function representation may be more convenient than any other. Once a mathematical model of a system is obtained, various analytical and computer tools can be used for analysis and synthesis purposes.

Mechanical systems may be modeled as systems of lumped masses (rigid bodies) or as distributed mass (continuous) systems. The latter are modeled by partial differential equations, whereas the former are represented by ordinary differential equations. In reality all systems are continuous, but, in most cases, it is easier and therefore preferred to approximate them with lumped mass models and ordinary differential equations.

Definition 1-10: Mass is considered a property of an element that stores the kinetic energy of translational motion. If W denotes the weight of a body, then M is given by

$$M = W/g \qquad (1-1)$$

where g is the acceleration of free fall of the body due to gravity ($g = 32.174\ \mathrm{ft/sec^2}$ in British units, and $g = 9.8066\ \mathrm{m/sec^2}$ in SI units).

The equations of a linear mechanical system are written by first constructing a model of the system containing interconnected linear elements and then by applying Newton's law of motion to the **free-body diagram (FBD)**. For translational motion, the equation of motion is Eq. (1-2), and for rotational motion, Eq. (1-17) is used.

The motion of mechanical elements can be described in various dimensions as **translational, rotational,** or a combination of both. The equations governing the motion of mechanical systems are often directly or indirectly formulated from **Newton's law of motion.**

Translational Motion: The motion of translation is defined as a motion that takes place along a straight or curved path. The variables that are used to describe translational motion are **acceleration, velocity,** and **displacement.**

Newton's law of motion states that the algebraic sum of external forces acting on a rigid body in a given direction is equal to the product of the mass of the body and its acceleration in the same direction. The law can be expressed as

$$\sum_{external} forces = ma \qquad (1-2)$$

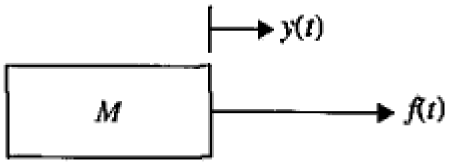


Figure 1-2 Force-mass system.

where *M* denotes the mass, and *a* is the acceleration in the direction considered. Fig. 1-2 illustrates the situation where a force is acting on a body with mass *M*. The force equation is written as

$$f(t) = Ma(t) = M\frac{d^2y(t)}{dt^2} = M\frac{dv(t)}{dt} \qquad (1-3)$$

where $a(t)$ is the acceleration, $v(t)$ denotes linear velocity, and $y(t)$ is the displacement of mass M, respectively.

For linear translational motion, in addition to the mass, the following system elements are also involved.

Linear spring: In practice, a linear spring may be a model of an actual spring or a compliance of a cable or a belt. In general, a spring is considered to be an element that stores potential energy.

$$f(t) = Ky(t) \qquad (1-4)$$

where *K* is the spring constant, or simply stiffness. Eq. (1-4) implies that the force acting on the spring is directly proportional to the displacement (deformation) of the spring. The model representing a linear spring element is shown in Fig. 1-2. If the spring is preloaded with a preload tension of *T*, then Eq. (1-4) should be modified to

$$f(t) - T = Ky(t) \qquad (1-5)$$

Friction for translation motion: Whenever there is motion or tendency of motion between two physical elements, frictional forces exist. The frictional forces encountered in physical systems are usually of a nonlinear nature. The characteristics of the frictional forces between two contacting surfaces often depend on such factors as the composition of the surfaces, the pressure between the surfaces, and their relative velocity among others, so an exact mathematical description of the frictional force is difficult. Three different types of friction are commonly used in practical systems: **viscous friction, static friction,** and **Coulomb friction:** These are discussed separately in the following paragraphs.

Viscous friction. Viscous friction represents a retarding force that is a linear relationship between the applied force and velocity. The schematic diagram element for viscous friction is often represented by a dashpot, such as that shown in Fig. 1-3. The mathematical expression of viscous friction is

$$f(t) = B\frac{dy(t)}{dt} \qquad (1-6)$$

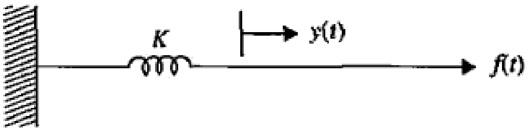


Figure 1-3 Force-spring system.

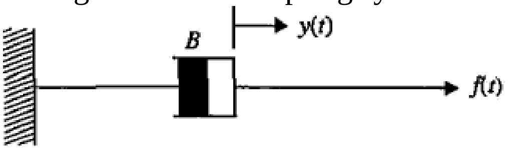


Figure 1-4 Dashpot for viscous friction.

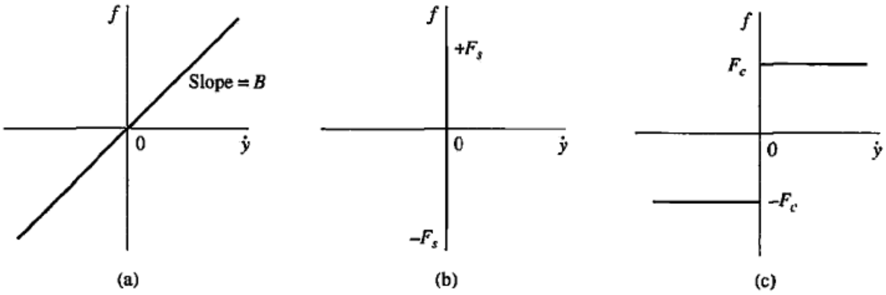


Figure 1-5 Graphical representation of linear and nonlinear frictional forces, (a) Viscous friction, (b) Static friction, (c) Coulomb friction.

where B is the viscous frictional coefficient. Fig. 1-5(a) shows the functional relation between the viscous frictional force and velocity.

Static friction: Static friction represents a retarding force that tends to prevent motion from beginning. The static frictional force can be represented by the expression

$$f(t) = \pm(F_s)|_{\dot{y}=0} \qquad (1-7)$$

which is defined as a frictional force that exists only when the body is stationary but has a tendency of moving. The sign of the friction depends on the direction of motion or the initial direction of velocity. The force-to-velocity relation of static friction is illustrated in Fig. 1-5(b). Notice that, once motion begins, the static frictional force vanishes and other frictions take over.

Coulomb friction: Coulomb friction is a retarding force that has constant amplitude with respect to the change of velocity, but the sign of the frictional force changes with the reversal of the direction of velocity. The mathematical relation for the Coulomb friction is given by

$$f(t) = F_c \frac{\left(\frac{dy(t)}{dt}\right)}{\left|\left(\frac{dy(t)}{dt}\right)\right|} \quad (1-8)$$

where F_C is the **Coulomb friction coefficient**. The functional description of the friction-to-velocity relation is shown in Fig. 1-5(c).

It should be pointed out that the three types of frictions cited here are merely practical models that have been devised to portray frictional phenomena found in physical systems. They are by no means exhaustive or guaranteed to be accurate. In many unusual situations, we have to use other frictional models to represent the actual phenomenon accurately.
ball bearings used in spacecraft systems. It turns out that rolling dry friction has nonlinear hysteresis properties that make it impossible for use in linear system modeling.

Example 1-1: [59] Consider the mass-spring-friction system shown in Fig. 1-6(a). The linear motion concerned is in the horizontal direction. The free-body diagram of the system is shown in Fig. 1-6(b). The force equation of the system is

$$f(t) - B\frac{dy(t)}{dt} - Ky(t) = M\frac{d^2y(t)}{dt^2} \quad (1-9)$$

The last equation may be rearranged by equating the highest-order derivative term to the rest of the terms:

$$\frac{d^2y(t)}{dt^2} = -\frac{B}{M}\frac{dy(t)}{dt} - \frac{K}{M}y(t) + \frac{1}{M}f(t) \quad (1-10)$$

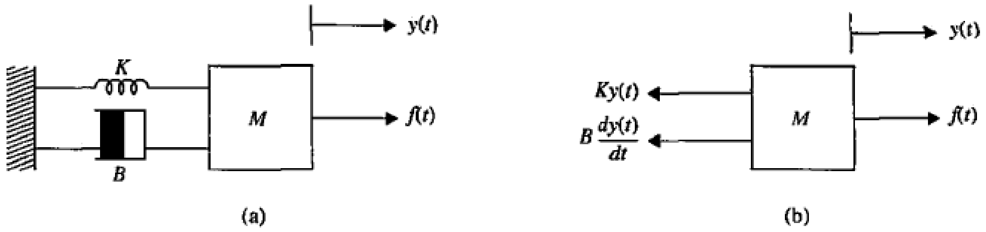


Figure 1-6 (a) Mass-spring-friction system, (b) Free-body diagram.

where $\dot{y}(t) = \left(\frac{dy(t)}{dt}\right)$ and $\ddot{y}(t) = \left(\frac{d^2y(t)}{dt^2}\right)$ represent velocity and acceleration, respectively. Or, alternatively, the former equation may be rewritten into an input-output form as

$$\ddot{y}(t) + \frac{B}{M}\dot{y}(t) + \frac{K}{M}y(t) = \frac{1}{M}f(t) \quad (1-11)$$

where $y(t)$ is the output and $\frac{f(t)}{M}$ is considered the input

Example 1-2: [59] As another example of writing the dynamic equations of a mechanical system with translational motion, consider the system shown in Fig. 1-7(a). Because the spring is deformed when it is subject to a *force* $f(t)$, two displacements, y_1 and y_2, must be assigned to the end points of the spring. The free-body diagrams of the system are shown in Fig. 1-7(b). The force equations are

$$f(t) = K[y_1(t) - y_2(t)] \quad (1-12)$$

$$-K[y_2(t) - y_1(t)] - B\frac{dy_2(t)}{dt} = M\frac{d^2y_2(t)}{dt^2} \quad (1-13)$$

These equations are rearranged in input-output form as

$$\frac{d^2y_2(t)}{dt^2}+\frac{B}{M}\frac{dy_2(t)}{dt}+\frac{K}{M}y_2(t)=\frac{K}{M}y_1(t) \qquad (1-14)$$

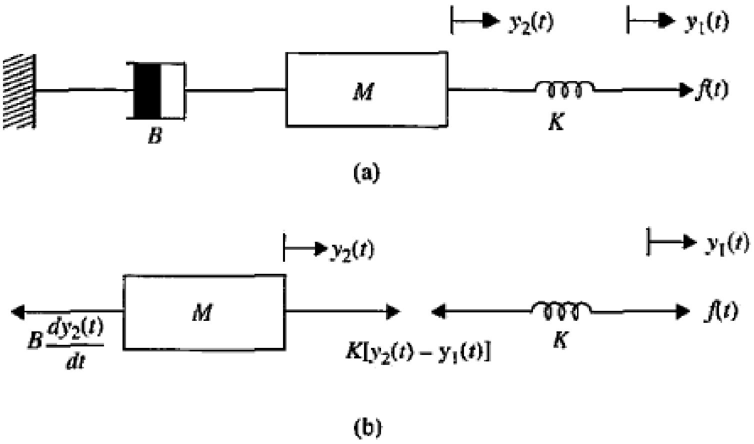


Figure 1-7 Mechanical system for Example 1-2. (a) Mass-spring-damper system, (b) Free-body diagram.

Example 1-3: [59] Consider the two degrees of freedom (2-DOF) spring-mass system, with two masses m_1 and m_2, two springs k_1 and k_2 and two forces f_1 and f_2, as shown in Fig. 1-8.

To avoid any confusion, we first draw the free-body diagram (FBD) of the system by assuming the masses are displaced in the positive direction, so that $y_1 > y_2 > 0$ (i.e., springs are both in tension). The FBD of the system is shown in Fig. 1-9. Applying Newton's second law to the masses M_1 and M_2, we have

$$f_1(t)-k_1y_1+k_2(y_1-y_2)=M_1\ddot{y}_1$$
$$f_2(t)-k_2(y_1-y_2)=M_2\ddot{y}_2 \qquad (1-15)$$

Rearranging the equations into the standard input-output form, we have

$$M_1\ddot{y}_1+(k_1+k_2)y_1-k_2y_2=f_1(t)$$
$$M_2\ddot{y}_2-k_2y_1+k_2y_2=f_2(t) \qquad (1-16)$$

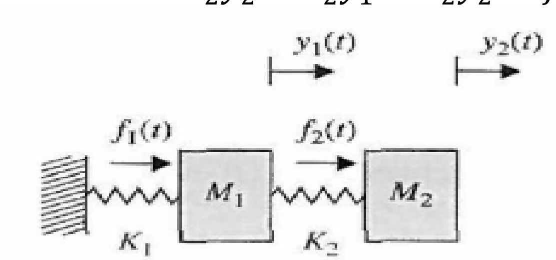


Figure 1-8 A 2-DOF spring-mass system.

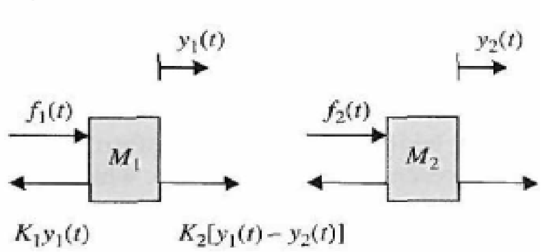


Figure 1-9 FBD of the 2-DOF spring-mass system.

Rotational Motion: The rotational motion of a body can be defined as motion about a fixed axis. The extension of Newton's law of motion for rotational motion states that the algebraic sum of moments or torque about a fixed axis is equal to the product of the inertia and the angular acceleration about the axis. Or

$$\sum torques = J\alpha \qquad (1-17)$$

where J denotes the inertia and α is the angular acceleration. The other variables generally used to describe the motion of rotation are **torque T, angular velocity ω,** and **angular displacement θ**. The elements involved with the rotational motion are as follows:

Inertia. Inertia, J, is considered a property of an element that stores the kinetic energy of rotational motion. The inertia of a given element depends on the geometric composition about the axis of rotation and its density. For instance, the inertia of a circular disk or shaft, of radius *r* and mass *M*, about its geometric axis is given by

$$J = \frac{1}{2}Mr^2 \qquad (1-18)$$

When a torque is applied to a body with inertia *J*, as shown in Fig. 1-10, the torque equation is written

$$T(t) = J\alpha(t) = J\frac{d\omega(t)}{dt} = J\frac{d^2\theta(t)}{dt^2} \qquad (1-19)$$

where $\theta(t)$is the angular displacement; $\omega(t)$, the angular velocity; and $\alpha(t)$, the angular acceleration.

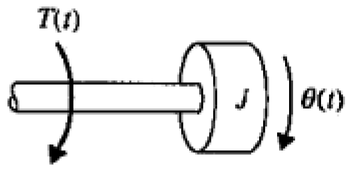


Figure 1-10 Torque-inertia system.

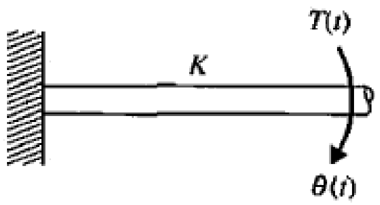


Figure 1-11 Torque torsional spring system.

Torsional spring: As with the linear spring for translational motion, a **torsional** spring **constant *K*,** in torque-per-unit angular displacement, can be devised to represent the compliance of a rod or a shaft when it is subject to an applied torque. Fig. 1-11 illustrates a simple torque-spring system that can be represented by the equation

$$T(t) = K\theta(t) \qquad (1-20)$$

If the torsional spring is preloaded by a preload torque of TP, Eq. (1-20) is modified to

$$T(t) - TP = K\theta(t) \qquad (1-21)$$

Friction for rotational motion: The three types of friction described for translational motion can be carried over to the motion of rotation. Therefore, Eqs. (1-6), (1-7), and (1-8) can be replaced, respectively, by their counterparts:

Viscous friction.

$$T(t) = B\frac{d\theta(t)}{dt} \qquad (1-22)$$

Static friction.

$$T = \pm F_s|_{\dot{\theta}=0} \qquad (1-23)$$

Coulomb friction.

$$T(t) = F_c \frac{\frac{d\theta(t)}{dt}}{\left|\frac{d\theta(t)}{dt}\right|} \quad (1-24)$$

Example 1-4: [59] The rotational system shown in Fig. 1-12(a) consists of a disk mounted on a shaft that is fixed at one end. The moment of inertia of the disk about the axis of rotation is *J*. The edge of the disk is riding on the surface, and the viscous friction coefficient between the two surfaces is *B*. The inertia of the shaft is negligible, but the torsional spring constant is *K*.

Assume that a torque is applied to the disk, as shown; then the torque or moment equation about the axis of the shaft is written from the free-body diagram of Fig. 1-12(b):

$$T(t) = J\frac{d^2\theta(t)}{dt^2} + B\frac{d\theta(t)}{dt} + K\theta(t) \quad (1-25)$$

Notice that this system is analogous to the translational system in Fig. 1-8. The state equations may be written by defining the state variables as $x_1(t) = \theta(t)$ and $.x_2 = dx_1(t)/dt$.

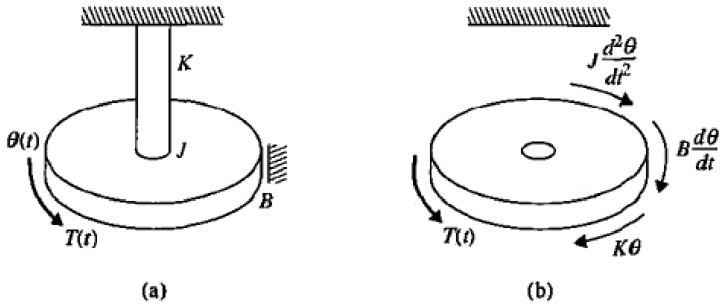


Figure 1-12 Rotational system for Example 1-4.

TABLE 1-1 Basic Rotational Mechanical System Properties and Their Units

Parameter	Symbol Used	SI Units	Other Units	Conversion Factors
Inertia	J	kg-m^2	slug-ft^2 lb-ft-sec^2 oz-in.-sec^2	1 g-cm = 1.417 x 10^{-5} oz-in.-sec2 1 lb-ft-sec^2 = 192 oz-in.-sec^2 = 32.2 Ib-ft^2 1 oz-in.-sec^2 = 386 oz-in^2 1 g-cm-sec^2 = 980 g-cm^2
Angular Displacement	T	Radian	Radian	1 rad=$\frac{180}{\pi} = 57.3deg$
Angular Velocity	O	radian/sec	radian/sec	$1rpm = \frac{2\pi}{60} = 0.107rad/sec$ 1rpm=6deg/sec
Torque	T	(N-m) dyne-cm	Ib-ft oz-in	1g-cm=0.0139oz-in 1Ib-ft=192oz-in 1oz-in=0.00521 Ib-ft
Spring Constant	K	N-m/rad	ft-Ib/rad	
Viscous Friction Coefficient	B	N-m/rad/sec	ft-Ib/rad/sec	
Energy	Q	J(joules)	Btu Calorie	1J=1N-m 1Btu=1055J 1cal=4.184J

Example 1-5:[59] Fig. 1-13(a) shows the diagram of a motor coupled to an inertial load through a shaft with a spring constant *K. A* non-rigid coupling between two mechanical components in a control system often causes torsional resonances that can be transmitted to all parts of the system. The system variables and parameters are defined as follows:

$$T_m(t) = motor\ torque$$
$$B_m = motor\ viscous - friction\ coefficient$$
$$K = spring\ constant\ of\ the\ shaft$$
$$\theta_m(t) = motor\ displacement$$
$$\omega_m(t) = motor\ velocity$$

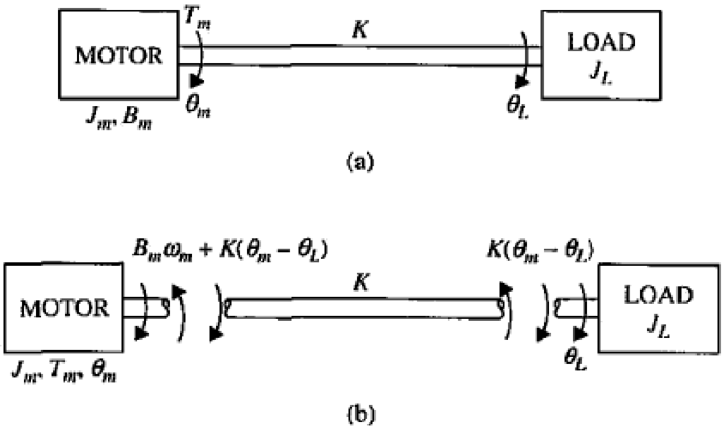


Figure 1-13 (a) Motor-load system, (b) Free-body diagram.

$$J_m = motor\ inertia$$
$$\theta_L(t) = load\ displacement$$
$$\omega_L = load\ velocity$$
$$J_L = load\ inertia$$

The free-body diagrams of the system are shown in Fig. 1-13(b). The torque equations of the system are

$$\frac{d^2\theta_m(t)}{dt^2} = -\frac{B_m}{J_m}\frac{d\theta_m(t)}{dt} - \frac{K}{J_m}[\theta_m(t) - \theta_L(t)] + \frac{1}{J_m}T_m(t) \qquad (1-26)$$

$$K[\theta_m(t) - \theta_L(t)] = J_L\frac{d^2\theta_L(t)}{dt^2} \qquad (1-27)$$

In this case, the system contains three energy-storage elements in J_m,J_L, and *K*. Thus, there should be three state variables. Care should be taken in constructing the state diagram and assigning the state variables so that a minimum number of the latter are incorporated. Eqs. (1-26) and (1-27) are rearranged as

$$\frac{d^2\theta_m(t)}{dt^2} = -\frac{B_m}{J_m}\frac{d\theta_m(t)}{dt} - \frac{K}{J_m}[\theta_m(t) - \theta_L(t)] + \frac{1}{J_m}T_m(t) \qquad (1-28)$$

$$\frac{d^2\theta_L(t)}{dt^2} = \frac{K}{J_L}[\theta_m(t) - \theta_L(t)] \qquad (1-29)$$

Parallel axis theorem: Sometimes it is necessary to calculate the moment of inertia of a homogeneous rigid body about an axis other than its geometrical axis.

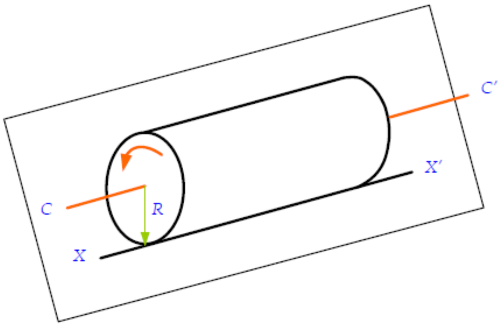


Figure 1-14 Homogeneous cylinder rolling on a flat surface

As an example to that, consider the system shown in Figure 1-14, where a cylinder of mass m and a radius R rolls on a flat surface. The moment of inertia of the cylinder is about axis CC' is

$$J_c = \frac{1}{2} m R^2 \qquad (1-30)$$

The moment of inertia J_x of the cylinder about axis xx' is

$$J_x = J_c + mR^2 = \frac{1}{2} mR^2 + mR^2 = \frac{3}{2} = mR^2 \qquad (1-31)$$

Forced response and natural response. The behavior determined by a forcing function is called a forced response, and that due initial conditions is called natural response. The period between initiation of a response and the ending is referred to as the transient period. After the response has become negligibly small, conditions are said to have reached a steady state.

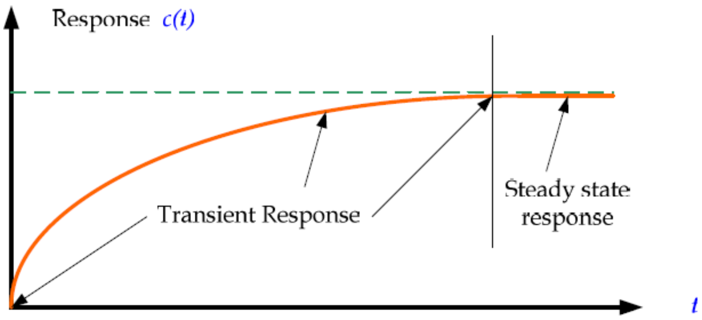


Figure 1-15 Transient and steady state response

Parallel Springs: For the springs in parallel, Figure 1-16, the equivalent spring constant k_{eq} is obtained from the relation

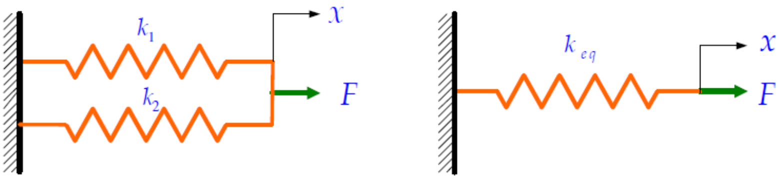


Figure 1-16 Parallel spring elements

$$F = k_1 x + k_2 x = (k_1 + k_2)x = k_{eg} x$$

Where

$$k_{eg} = k_1 + k_2 \ (for\ parallel\ springs) \qquad (1-32)$$

This formula can be extended to *n* springs connected side-by-side as follows:

$$k_{eg} = \sum_{i=1}^{n} k_i \quad \text{(for parallel springs)} \qquad (1-33)$$

Series Springs: [6] For the springs in series, Figure 1-16, the force in each spring is the same. Thus

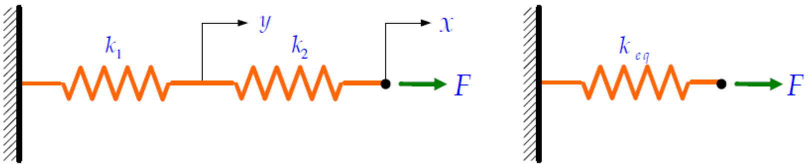


Figure 1-17 Series spring elements

$$F = k_1 y \quad , \quad F = k_2(x - y) \qquad (1-34)$$

Eliminating from these two equations yields

$$F = k_2\left(x - \frac{F}{k_1}\right)$$

Or

$$F = k_2 x - \frac{k_2}{k_1}F \ \Rightarrow \ k_2 x = F + \frac{k_2}{k_1}F = \left(\frac{k_1 + k_2}{k_1}\right)F$$

Or

$$x = \left(\frac{k_1 + k_2}{k_2 k_1}\right)F \Leftrightarrow F = \left(\frac{k_2 k_1}{k_1 + k_2}\right)x = \left(\frac{1}{\frac{1}{k_1} + \frac{1}{k_2}}\right)x$$

Where

$$\frac{1}{k_{eg}} = \frac{1}{k_1} + \frac{1}{k_2} \quad \text{(for series springs)} \qquad (1-35)$$

which can be extended to the case of *n* springs connected end-to-end as follows

$$\frac{1}{k_{eg}} = \sum_{i=1}^{n} \frac{1}{k_i} \quad \text{(for series springs)} \qquad (1-36)$$

Conversion between Translational and Rotational Motions: In motion-control systems, it is often necessary to convert rotational motion into translational motion. For instance, a load may be controlled to move along a straight line through a rotary motor-and-lead screw assembly, such as that shown in Fig. 1-20. Fig. 1-21 shows a similar situation in which a rack-and-pinion assembly is used as a mechanical linkage. Another familiar system in motion control is the control of a mass through a pulley by a rotary motor, as shown in Fig. 1-22. The systems shown in Figs. 1-20,1-21, and 1-22 can all be represented by a simple system with an equivalent inertia connected directly to the drive motor. For instance, the mass in Fig. 1-22 can be regarded as a point mass that moves about the pulley, which has a radius r. By disregarding the inertia of the pulley, the equivalent inertia that the motor sees is

$$J = Mr^2 = W/g\, r^2 \qquad (1-37)$$

If the radius of the pinion in Fig. 1-19 is r, the equivalent inertia that the motor sees is also given by Eq. (1-37).

Now consider the system of Fig. 1-18. The lead of the screw, L, is defined as the linear distance that the mass travels per revolution of the screw. In principle, the two systems in Fig. 1-19 and Fig. 1-20 are equivalent. In Fig. 1-19, the distance traveled by the mass per revolution of the pinion is $2\pi r$. By using Eq. (1-37) as the equivalent inertia for the system of Fig. 1-18, we have

$$J = W/g\,(L/2\pi)^2 \qquad (1-38)$$

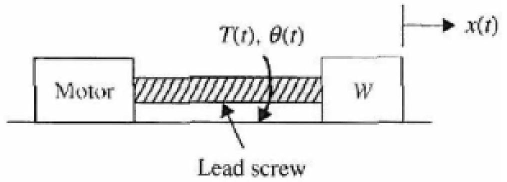


Figure 1-18 Rotary-to-linear motion control system (lead screw).

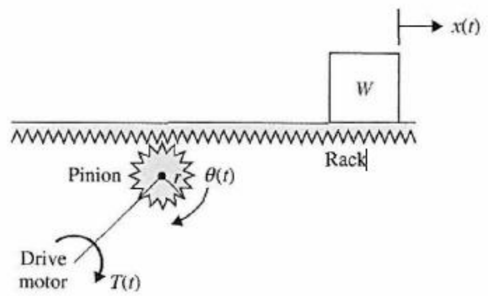


Figure 1-19 Rotary-to-linear motion control system (rack and pinion).

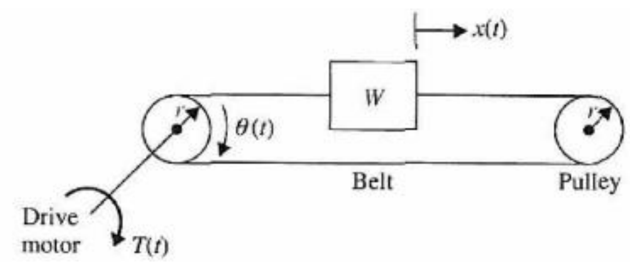


Figure 1-20 Rotary-to-linear motion control system (belt and pulley).

Gear Trains: A gear train, lever, or timing belt over a pulley is a mechanical device that transmits energy from one part of the system to another in such a way that force, torque, speed, and displacement may be altered. These devices can also be regarded as matching devices used to attain maximum power transfer. Two gears are shown coupled together in Fig. 1-21. The inertia and friction of the gears are neglected in the ideal case considered.

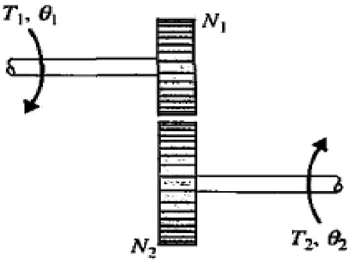


Figure 1-21 Gear train.

The relationships between the torques T_1 and T_2, angular displacement θ_1 and θ_2, and the teeth numbers N_1 and N_2 of the gear train are derived from the following facts:

1. The number of teeth on the surface of the gears is proportional to the radii r_1, and r_2 of the gears; that is,
$$r_1 N_2 = r_2 N_1 \tag{1-39}$$
2. The distance traveled along the surface of each gear is the same. Thus,
$$\theta_1 r_1 = \theta_2 r_2 \tag{1-40}$$
3. The work done by one gear is equal to that of the other since there are assumed to be no losses. Thus,
$$T_1 \theta_1 = T_2 \theta_2 \tag{1-41}$$

If the angular velocities of the two gears ω_1 and ω_2 are brought into the picture, Eqs. (1-39) through (1-41) lead to
$$\frac{T_1}{T_2} = \frac{\theta_2}{\theta_1} = \frac{N_1}{N_2} = \frac{\omega_2}{\omega_1} = \frac{r_1}{r_2} \tag{1-42}$$
In practice, gears do have inertia and friction between the coupled gear teeth that often cannot be neglected. An equivalent representation of a gear train with viscous friction, Coulomb friction, and inertia considered as lumped parameters is shown in Fig. 1-21, where T denotes the applied torque, T_1 and T_2 are the transmitted torque, F_{c1} and F_{c2} are the Coulomb friction coefficients, and B_1 and B_2 are the viscous friction coefficients. The torque equation for gear 2 is
$$T_2(t) = J_2 \frac{d^2\theta_2(t)}{dt^2} + B_2 \frac{d\theta_2(t)}{dt} + F_{c2} \frac{\omega_2}{|\omega_2|} \tag{1-43}$$

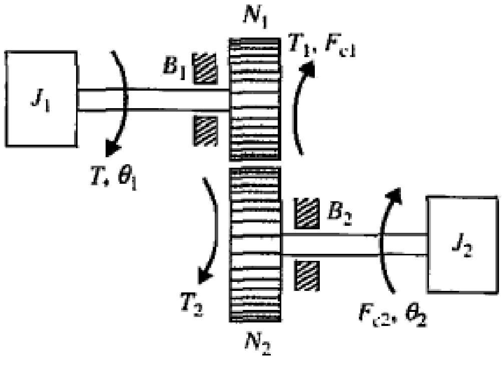


Figure 1-22 Gear train with friction and inertia.

The torque equation on the side of gear 1 is
$$T(t) = J_1 \frac{d^2\theta_1(t)}{dt^2} + B_1 \frac{d\theta_1(t)}{dt} + F_{c1} \frac{\omega_1}{|\omega_1|} + T_1(t) \tag{1-44}$$

Using Eq. (1-42), Eq. (1-43) is converted to

$$T_1(t) = N_1/N_2\, T_2(t)$$
$$= \left(\frac{N_1}{N_2}\right)^2 J_2 \frac{d^2\theta_1(t)}{dt^2} + \left(\frac{N_1}{N_2}\right)^2 B_2 \frac{d\theta_1(t)}{dt} + \frac{N_1}{N_2} F_{c2} \frac{\omega_2}{|\omega_2|} \qquad (1-45)$$

Eq. (1-45) indicates that it is possible to reflect inertia, friction, compliance, torque, speed, and displacement from one side of a gear train to the other. The following quantities are obtained when reflecting from gear 2 to gear 1 :

$$\begin{aligned} \text{Inertia: } & (N_1/N_2)^2 J_2 \\ \text{Viscous} - \text{friction coefficient: } & (N_1/N_2)^2 B_2 \\ Torque\text{: } & N_1/N_2\ T_2 \\ Angular\ displacement\text{: } & N_1/N_2\ \theta_2 \\ Angular\ velocity\text{: } & N_1/N_2\ \omega_2 \\ Coulomb\ friction\ torque\text{: } & \frac{N_1}{N_2} F_{c2} \frac{\omega_1}{|\omega_1|} \end{aligned} \qquad (1-46)$$

Similarly, gear parameters and variables can be reflected from gear 1 to gear 2 by simply interchanging the subscripts in the preceding expressions. If a torsional spring effect is present, the spring constant is also multiplied by $(N_1/N_2)^2$ in reflecting from gear 2 to gear 1. Now substituting Eq. (1-45) into Eq. (1-44), we get

$$T(t) = J_{1e} \frac{d^2\theta_1(t)}{dt^2} + B_{1e} \frac{d\theta_1(t)}{dt} + T_F \qquad (1-47)$$

Where

$$J_{1e} = J_1 + (N_1/N_2)^2 J_2 \qquad (1-48)$$
$$B_{1e} = B_1 + (N_1/N_2)^2 B_2 \qquad (1-49)$$
$$T_F = F_{c1} \frac{\omega_1}{|\omega_1|} + \frac{N_1}{N_2} F_{c2} \frac{\omega_2}{|\omega_2|} \qquad (1-50)$$

Backlash and Dead Zone (Nonlinear Characteristics): Backlash and dead zone are commonly found in gear trains and similar mechanical linkages where the coupling is not perfect. In a majority of situations, backlash may give rise to undesirable inaccuracy, oscillations, and instability in control systems. In addition, it has a tendency to wear out the mechanical elements. Regardless of the actual mechanical elements, a physical model of backlash or dead zone between an input and an output member is shown in Fig. 1-23. The model can be used for a rotational system as well as for a translational system. The amount of backlash is *b/2* on either side of the reference position.

In general, the dynamics of the mechanical linkage with backlash depend on the relative inertia-to-friction ratio of the output member. If the inertia of the output member is very small compared with that of the input member, the motion is controlled predominantly by friction. This means that the output member will not coast whenever there is no contact between the two members. When the output is driven by the input, the two members will travel together until the input member reverses its direction; then the output member will be at a standstill until the backlash is taken up on the other side, at which time it is assumed that the output member instantaneously takes on the velocity of the input member. The transfer characteristic between the input and output displacements of a system with backlash with negligible output inertia is shown in Fig. 1-24.

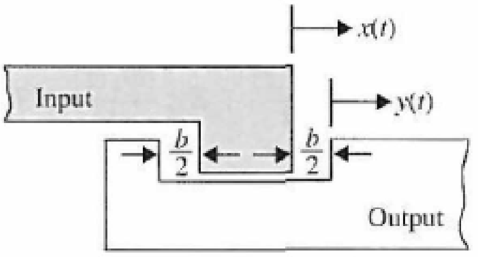


Figure 1-23 Physical model of backlash between two mechanical elements.

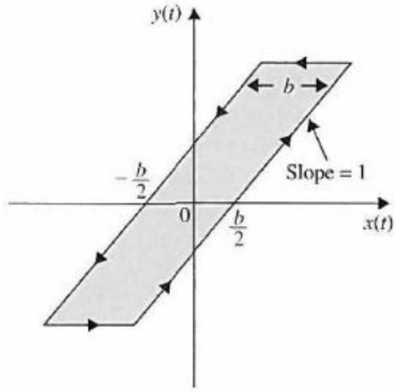


Figure 1-24 Input-output characteristic of backlash.

Example 1-6: [6] Let us obtain the equivalent viscous-friction coefficient b_{eg} for each of the damper systems shown in Figures 1–25(a) and (b). An oil-filled damper is often called a dashpot. A dashpot is a device that provides viscous friction, or damping. It consists of a piston and oil-filled cylinder. Any relative motion between the piston rod and the cylinder is resisted by the oil because the oil must flow around the piston (or through orifices provided in the piston) from one side of the piston to the other. The dashpot essentially absorbs energy. This absorbed energy is dissipated as heat, and the dashpot does not store any kinetic or potential energy.

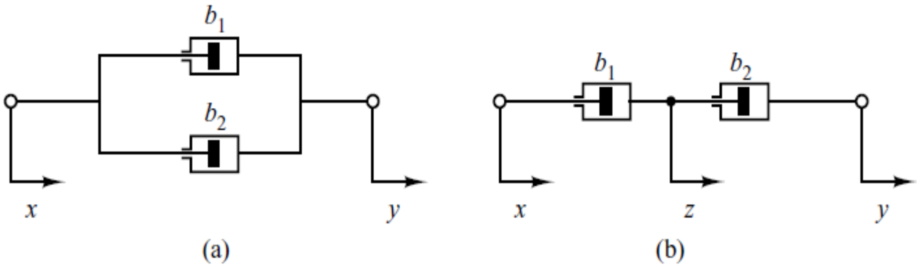


Figure 1–25 (a) Two dampers connected in parallel; (b) two dampers connected in series.

Solution.

(a) The force f due to the dampers is

$$f = b_1(\dot{y} - \dot{x}) + b_2(\dot{y} - \dot{x}) = (b_1 + b_2)(\dot{y} - \dot{x}) \qquad (1-51)$$

In terms of the equivalent viscous-friction coefficient b_{eg} , force f is given by

$$f = b_{eg}(\dot{y} - \dot{x}) \qquad (1-52)$$

Hence

$$b_{eg} = b_1 + b_2 \qquad (1-53)$$

(b) The force f due to the dampers is

$$f = b_1(\dot{z} - \dot{x}) = b_2(\dot{y} - \dot{z}) \qquad (1-54)$$

where z is the displacement of a point between damper b_1 and damper b_2 . (Note that the same force is transmitted through the shaft.) From Equation (1–54), we have

$$(b_1 + b_2)\dot{z} = b_2\dot{y} + b_1\dot{x} \qquad (1-55)$$

or

$$\dot{z} = \frac{1}{b_1 + b_2}(b_2\dot{y} + b_1\dot{x}) \qquad (1-56)$$

In terms of the equivalent viscous-friction coefficient b_{eg}, force f is given by

$$f = b_{eg}(\dot{y} - \dot{x}) \qquad (1-57)$$

By substituting Equation (1–56) into Equation (1–54), we have

$$f = b_2(\dot{y} - \dot{z}) = b_2\left[\dot{y} - \frac{1}{b_1 + b_2}(b_2\dot{y} + b_1\dot{x})\right]$$

$$= \frac{b_1 b_2}{b_1 + b_2}(\dot{y} - \dot{x}) \qquad (1-58)$$

Thus,

$$f = b_{eg}(\dot{y} - \dot{x}) = \frac{b_1 b_2}{b_1 + b_2}(\dot{y} - \dot{x}) \qquad (1-59)$$

Hence,

$$b_{eg} = \frac{b_1 b_2}{b_1 + b_2} = \frac{1}{\frac{1}{b_1} + \frac{1}{b_2}} \qquad (1-60)$$

1-3 Mathematical Modeling of Electrical Systems

Modeling of Passive Electrical Elements: Current and Voltage Current and voltage are the primary variables used to describe a circuit's behavior. Current is the flow of electrons. It is the time rate of change of electrons passing through a defined area, such as the cross-section of a wire. Because electrons are negatively charged, the positive direction of current flow is opposite to that of electron flow. The mathematical description of the relationship between the number of electrons (called charge q) and current i is

$$i = \frac{dq}{dt} \quad or \quad q(t) = \int i dt \qquad (1-61)$$

The unit of charge is the coulomb (C) and the unit of current is ampere (A), which is one coulomb per second.

Energy is required to move a charge between two points in a circuit. The work per unit charge required to do this is called voltage. The unit of voltage is *volt* (V), which is defined to be joule per coulomb. The voltage difference between two points in a circuit is a measure of the energy required to move charge from one point to the other.

Active and Passive Elements. Circuit elements may be classified a active or passive.

Passive Element. an element that contains no energy sources (i.e. the element needs power from another source to operate); these include resistors, capacitors and inductors.

Active Element: an element that acts as an energy source; these include batteries, generators, solar cells, and op-amps.

Current Source and Voltage Source A voltage source is a device that causes a specified voltage to exist between two points in a circuit. The voltage may be time varying or time invariant (for a sufficiently long time). Figure 1-26(a) is a schematic diagram of a voltage source. Figure 1-26(b) shows a voltage source that has a constant value for an indefinite time. Often the voltage is denoted by E or V . A battery is an example of this type of voltage.

A current source causes a specified current to flow through a wire containing this source. Figure 1-26(c) is a schematic diagram of a current source

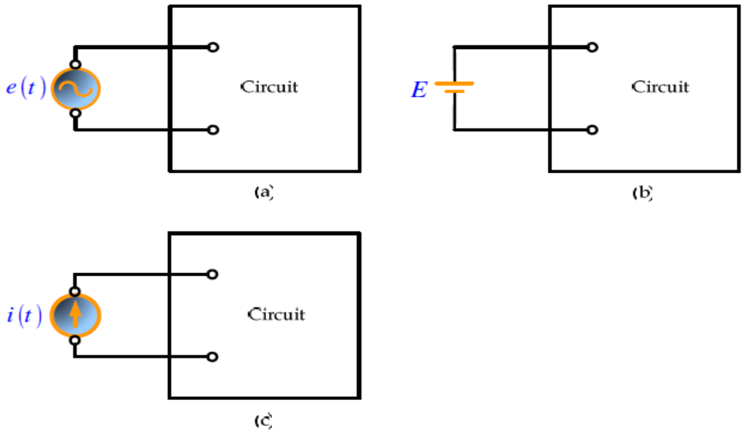


Figure 1-26 (a) Voltage source; (b) constant voltage source; (c) current source

Resistance: Ohm's law states that the voltage drop, $e_R(t)$, across a resistor R is proportional to the current $i(t)$ going through the resistor. Or

$$e_R = i(t)R \qquad (1-62)$$

where e_R is the voltage across the resistor and $i(t)$ is the current through the resistor. The unit of resistance is the ohm (Ω) , where

$$ohm = \frac{volt}{ampere}$$

Capacitance Elements: Two conductors separated by a nonconducting medium form a capacitor, so two metallic plates separated by a very thin dielectric material form a capacitor. The capacitance C is a measure of the quantity of charge that can be stored for a given voltage across the plates. The capacitance C of a capacitor can thus be given by

$$C = q/e_C \qquad (1-63)$$

where q is the quantity of charge stored and e_C is the voltage across the capacitor. The unit of capacitance is the farad (F) , where

$$\text{farad} = \frac{\text{ampere} - \text{second}}{\text{volt}} = \frac{\text{coulomb}}{\text{volt}}$$

Notice that, since $i = dq/\,dt$ and $e_C = q/C$, we have

$$i = C\frac{de_C}{dt} \qquad (1-64)$$

or

$$e_C = \frac{1}{C}\int i(t)dt \qquad (1-65)$$

Although a pure capacitor stores energy and can release all of it, real capacitors exhibit various losses. These energy losses are indicated by a power factor , which is the ratio of energy lost per cycle of ac voltage to the energy stored per cycle. Thus, a small-valued power factor is desirable.

Inductance Elements: If a circuit lies in a time varying magnetic field, an electromotive force is induced in the circuit. The inductive effects can be classified as self inductance and mutual inductance.

Self inductance, or simply inductance, L is the proportionality constant between the induced voltage e_L volts and the rate of change of current (or change in current per second) $di/\,dt$ amperes per second; that is,

$$L = \frac{e_L}{di/\,dt} \qquad (1-66)$$

The unit of inductance is the henry (H). An electrical circuit has an inductance of 1 henry when a rate of change of 1 ampere per second will induce an emf of 1 volt:

$$\text{henry} = \frac{\text{volt}}{\text{ampere/second}} = \frac{\text{Weber}}{\text{ampere}}$$

The voltage e_L across the inductor L is given by

$$e_L = L\frac{di_L}{dt} \qquad (1-67)$$

Where L_i is the current through the inductor. The current $i_L(t)$ can thus be given by

$$i_L(t) = \frac{1}{L}\int_0^t e_L dt + i_L(0) \qquad (1-68)$$

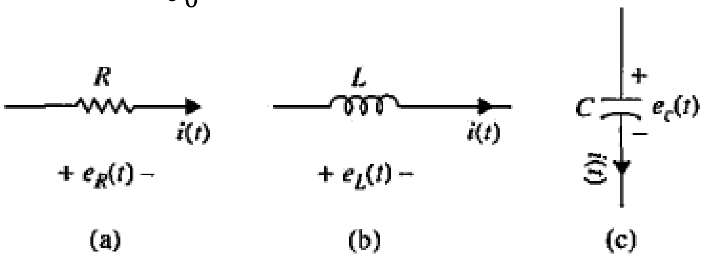
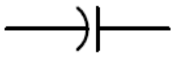


Figure 1-27 Basic passive electrical elements, (a) A resistor. (b) An inductor. (c) A capacitor.

TABLE 1-2. Summary of elements involved in linear electrical systems

Element	Voltage-current	Current-voltage	Voltage-charge	Impedance, $Z(s)=V(s)/I(s)$
Capacitor 	$v(t)=\frac{1}{c}\int_0^t i(\tau)d\tau$	$i(t)=C\frac{dv(t)}{dt}$	$v(t)=\frac{1}{c}q(t)$	$\frac{1}{Cs}$
Resistor 	$v(t)=R\,i(t)$	$i(t)=\frac{1}{R}v(t)$	$v(t)=R\frac{dq(t)}{dt}$	R
Inductor 	$v(t)=L\frac{di(t)}{dt}$	$i(t)=\frac{1}{L}\int_0^t v(\tau)d\tau$	$v(t)=L\frac{d^2q(t)}{dt^2}$	Ls

The following set of symbols and units are used: v(t) = V (Volts), i(t) = A (Amps), q(t) = Q (Coulombs), C = F (Farads), R = / (Ohms), L = H (Henries).

Series Circuit: The combined resistance of series-connected resistors is the sum of the separate resistances. Figure 1-28 shows a simple series circuit. The voltage between points A and B is

$$e = e_1 + e_2 + e_3 \qquad (1-69)$$

Where

$$e_1 = iR_1, \qquad e_2 = iR_2, \qquad e_3 = iR_3 \qquad (1-70)$$

Thus,

$$\frac{e}{i} = R_1 + R_2 + R_3 \qquad (1-71)$$

The combined resistance is given by

$$R = R_1 + R_2 + R_3 \qquad (1-72)$$

In general,

$$R = \sum_{i=1}^{n} R_i \qquad (1-73)$$

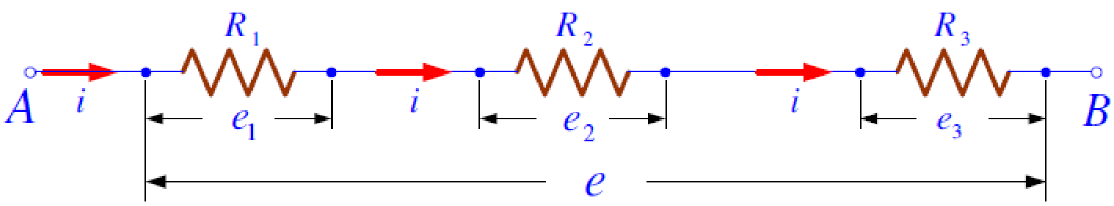


Figure 1-28 Series Circuit

Parallel Circuit. For the parallel circuit shown in figure 1-28,

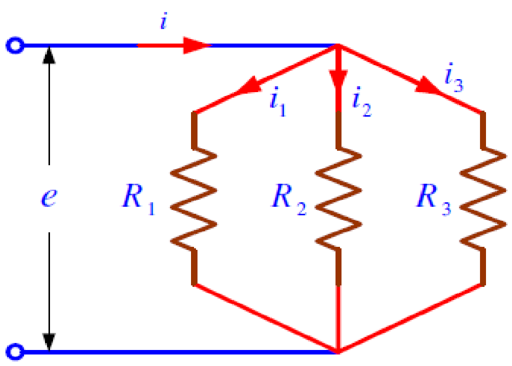


Figure 1-29 Parallel Circuit

$$i_1 = \frac{e}{R_1}, \qquad i_2 = \frac{e}{R_2}, \qquad i_3 = \frac{e}{R_3} \qquad (1-74)$$

Since $= i_1 + i_2 + i_3$, it follows that

$$i = \frac{e}{R_1} + \frac{e}{R_2} + \frac{e}{R_3} \qquad (1-75)$$

where R is the combined resistance. Hence,

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \qquad (1-76)$$

or

$$R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}} = \frac{R_1R_2R_3}{R_1R_2 + R_2R_3 + R_3R_1} \qquad (1-77)$$

In general

$$\frac{1}{R} = \sum_{i=1}^{n} \frac{1}{R_i} \qquad (1-78)$$

Modeling of Electrical Networks: The classical way of writing equations of electric networks is based on the loop method or the node method, both of which are formulated from the two laws of Kirchhoff, which state:

Current Law or Loop Method: A node in an electrical circuit is a point where three or more wires are joined together. Kirchhoff's Current Law (KCL) states that

The algebraic summation of all currents entering a node is zero.

or

The algebraic sum of all currents entering a node is equal to the sum of all currents leaving the same node.

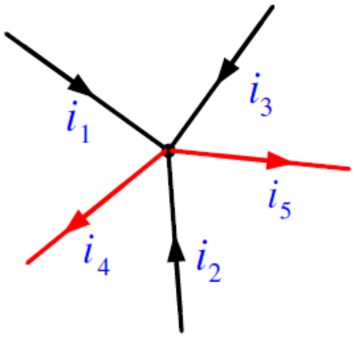


Figure 1-30 Node.

As applied to Figure 1-30, kirchhoff's current law states that

$$i_1 + i_2 + i_3 - i_4 - i_5 = 0 \qquad (1-79)$$

or

$$\underbrace{i_1 + i_2 + i_3}_{\text{Entering currents}} = \underbrace{i_4 + i_5}_{\text{Leaving currents}} \qquad (1-80)$$

Voltage Law or Node Method: The algebraic sum of all voltage drops around a complete closed loop is zero.

or

The sum of the voltage drops is equal to the sum of the voltage rises around a loop.

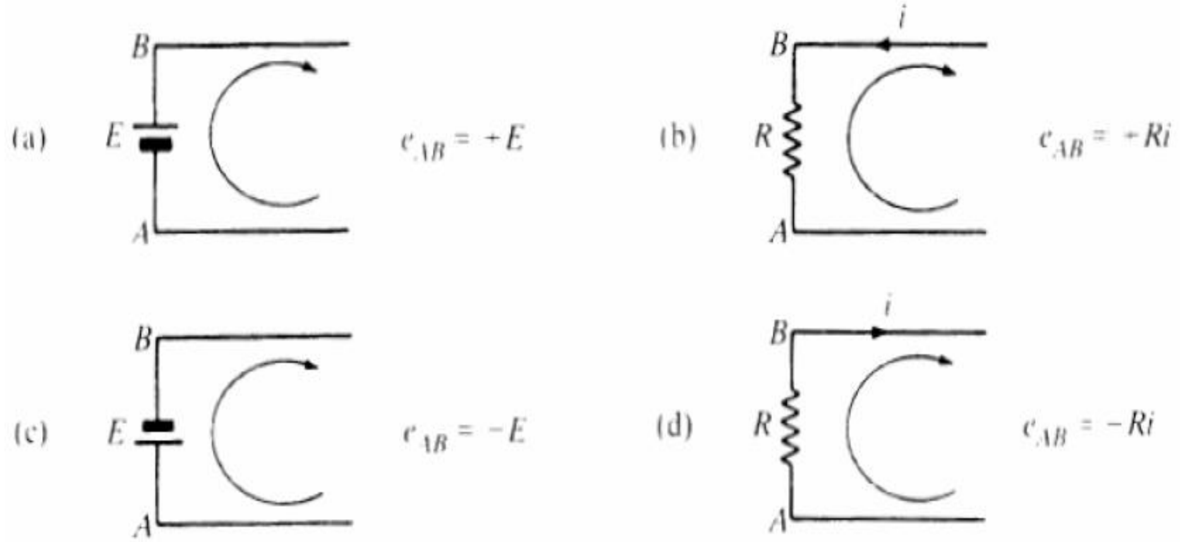


Figure 1-31Diagrams showing voltage rises and voltage drops in circuits. (Note: Each circular arrows shows the direction one follows in analyzing the respective circuit)

A rise in voltage [which occurs in going through a source of electromotive force from the negative terminal to the positive terminal, as shown in Figure 1-31 (a), or in going through a resistance in opposition to the current flow, as shown in Figure 1-30 (b)] should be preceded by a plus sign.

A drop in voltage [which occurs in going through a source of electromotive force from the positive to the negative terminal, as shown in Figure 1-31 (c), or in going through a resistance in the direction of the current flow, as shown in Figure 1-31 (d)] should be preceded by a minus sign.

Figure 1-32 shows a circuit that consists of a battery and an external resistance.

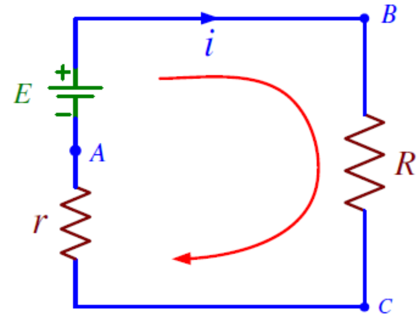


Figure 1-32 Electrical Circ
uit.

Here *E* is the electromotive force, *r* is the internal resistance of the battery, *R* is the external resistance and *i* is the current. Following the loop in the clockwise direction (*A* →*B* →*C* →*D*), we have

$$\vec{e}_{AB} + \vec{e}_{BC} + \vec{e}_{CA} = 0 \qquad (1-81)$$

or

$$E - iR - ir = 0 \qquad (1-82)$$

From which it follows that

$$i = E/R + r \qquad (1-83)$$

***LRC* Circuit.** Consider the electrical circuit shown in Figure 1-33.The circuit consists of an inductance L (henry), a resistance R (ohm), and a capacitance C (farad).

Applying Kirchhoff's voltage law to the system, we obtain the following equations:

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int idt = e_i \qquad (1-84)$$

$$\frac{1}{C}\int idt = e_o \qquad (1-85)$$

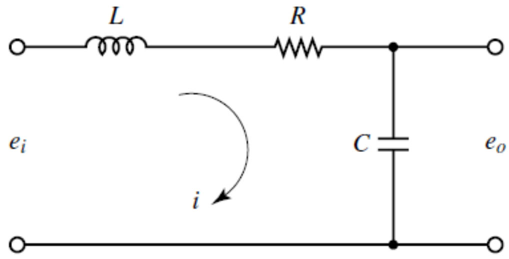


Figure1–33 Electrical circuit.

Equations (1–84) and (1–85) give a mathematical model of the circuit.

Example 1-7: [59] Let us consider the RLC network shown in Fig. 1-34. Using the voltage law

$$e(t) = e_R + e_L + e_C \qquad (1-86)$$

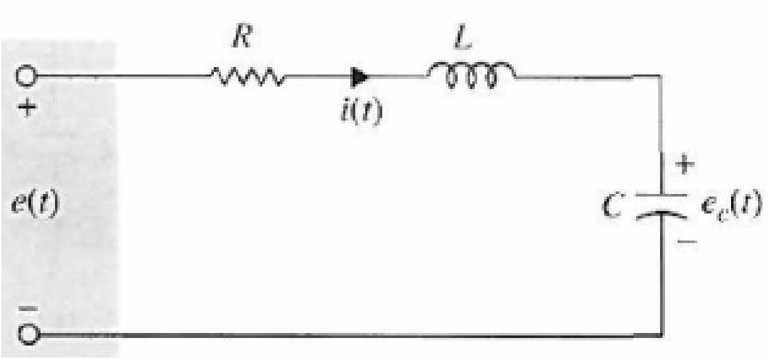


Figure 1-34 RLC network, Electrical schematics.

where

$$e_R = \text{Voltage across the resistor R}$$
$$e_L = \text{Voltage across the inductor}$$
$$e_C = \text{Voltage across the capacitor C}$$

or

$$e(t) = e_C + Ri(t) + L\frac{di(t)}{dt} \qquad (1-87)$$

Using current in C:

$$C\frac{de_C(t)}{dt} = i(t) \qquad (1-88)$$

and taking a derivative of Eq. (1-87) with respect to time, we get the equation of the RLC network as

$$L\frac{d^2i(t)}{dt^2} + R\frac{di(t)}{dt} + \frac{i(t)}{C} = \frac{de(t)}{dt} \qquad (1-89)$$

A practical approach is to assign the current in the inductor *L*, *i(t)*, and the voltage across the capacitor C,$e_C(t)$, as the state variables. The reason for this choice is because the state variables are directly related to the energy-storage element of a system. The inductor stores kinetic energy, and the capacitor stores electric potential energy. By assigning *i(i)* and $e_C(t)$ as state variables, we have a complete description of the past history (via the initial states) and the present and future states of the network.

Example 1-8: [59] Consider the RC circuit shown in Fig. 1-35. Find the differential equation of the system. Using the voltage law

$$e_{in}(t) = e_R(t) + e_c(t) \qquad (1-90)$$

where

$$e_R = iR \qquad (1-91)$$

and the voltage across the capacitor $v_C(t)$ is

$$e_C(t) = \frac{1}{C}\int idt \qquad (1-92)$$

But from Fig. 1-35

$$e_o(t) = \frac{1}{C}\int idt \qquad (1-93)$$

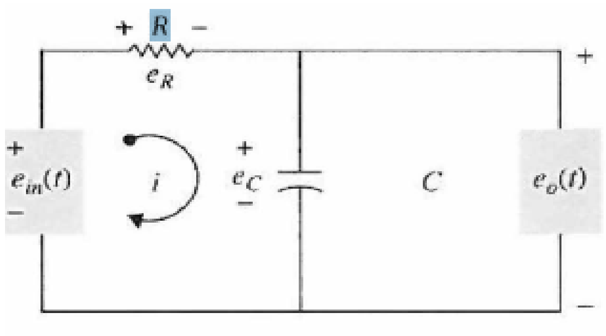


Figure 1-35 Simple electrical *RC* circuit.

If we differentiate Eq. (1-93) with respect to time, we get

$$\frac{i}{C} = \frac{de_o(t)}{dt} \qquad (1-94)$$

or

$$C\dot{e}_o = i \qquad (1-95)$$

This implies that Eq. (1-93) can be written in an input-output form

$$e_{in}(t) = RC\dot{e}_o + e_0(t) \qquad (1-96)$$

1-4 Analogous Systems

Systems that can be represented by the same mathematical model, but that are physically different, are called analogous systems. Thus analogous systems are described by the same differential or integrodifferential equations or transfer functions.

The concept of analogous is useful in practice, for the following reasons:

1. The solution of the equation describing one physical system can be directly applied to analogous systems in any other field.
2. Since one type of system may be easier to handle experimentally than another, instead of building and studying a mechanical system (or a hydraulic system, pneumatic system, or the like), we can build and study its electrical analog, for electrical or electronic system, in general, much easier to deal with experimentally.

Mechanical-Electrical Analogies: Mechanical systems can be studied through their electrical analogs, which may be more easily constructed than models of the corresponding mechanical systems. There are two electrical analogies for mechanical systems: The Force-Voltage Analogy and The Force Current Analogy.

Force Voltage Analogy

Consider the mechanical system of Figure 1-36(a) and the electrical system of Figure 1-36(b).

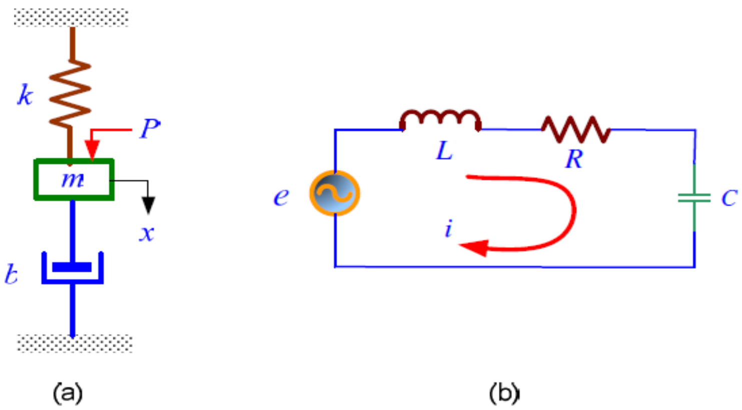


Figure 1-36 Analogous mechanical and electrical systems.

The equation for the mechanical system is

$$m\frac{d^2x}{dt^1}+b\frac{dx}{dt}+kx=p \qquad (1-97)$$

where x is the displacement of mass m , measured from equilibrium position. The equation for the electrical system is

$$L\frac{di}{dt}+Ri+\frac{1}{C}\int idt=e \qquad (1-98)$$

In terms of electrical charge q , this last equation becomes

$$L\frac{d^2q}{dt^2}+R\frac{dq}{dt}+\frac{1}{C}q=e \qquad (1-99)$$

Comparing equations (1-97) and (1-99), we see that the differential equations for the two systems are of identical form. Thus, these two systems are analogous systems. The terms that occupy corresponding positions in the differential equations are called analogous quantities, a list of which appear in Table 1-3.

TABLE 1-3 Force Voltage Analogy

Mechanical Systems	Electrical Systems
Force p (Torque T)	Voltage *e*
Mass m (Moment of inertia *J*)	Inductance *L*
Viscous-friction coefficient *b*	Resistance *R*
Spring constant *k*	Reciprocal of capacitance, 1/ *C*
Displacement *x* (angular displacement θ)	Charge *q*
Velocity $\dot{x}$ (angular velocity $\dot{\theta}$)	Current *i*

1-5 Linear Systems

A system is called linear if the principle of superposition applies. The principle of superposition states that the response produced by the simultaneous application of two different forcing functions is the sum of the two individual responses. Hence, for the linear system, the response to several inputs can be calculated by treating one input at a time and adding the results. It is this principle that allows one to build up complicated solutions to the linear differential equation from simple solutions.

In an experimental investigation of a dynamic system, if cause and effect are proportional, thus implying that the principle of superposition holds, then the system can be considered linear.

A system is linear if it meets the following two criteria:

1. Additivity: If $x_1(t) \rightarrow y_1(t)$ and $x_2(t) \rightarrow y_2(t)$, then
$$x_1(t)+x_2(t) \rightarrow y_1(t)+y_2(t).$$
2. Homogeneity: If $x_1(t) \rightarrow y_1(t)$, then
$$ax_1(t) \rightarrow ay_1(t),$$

where a is a constant. The criteria must apply for all $x_1(t)$ and $x_2(t)$ and for all a.

These two criteria can be combined to yield the principle of superposition. A system satisfies the principle of superposition if, with the inputs and outputs as just defined,

$$a_1x_1(t) + a_2x_2(t) \rightarrow a_1y_1(t) + a_2y_2(t),$$

where a_1 and a_2 are constants. A system is linear if and only if it satisfies the principle of superposition.

1-6 Linear time-invariant and Linear time-Varying Systems

A differential equation is linear if the coefficients are constants or functions only of the independent variable. Dynamic systems that are composed of linear time-invariant lumped-parameter components may be described by linear time-invariant differential equations-that is, constant-coefficient differential equations. Such systems are called linear time-invariant (or linear constant-coefficient) systems. Systems that are represented by differential equations whose coefficients are functions of time are called linear time-varying systems. An example of a time-varying control system is a spacecraft control system. (The mass of a spacecraft changes due to fuel consumption.)

Definition 1-11: A time-variable differential equation is a differential equation with one or more of its coefficients are functions of time, t. For example, the differential equation:

$$t^2\frac{d^2y(t)}{dt^2} + y(t) = u(t) \qquad (1-100)$$

where u and y are dependent variables) is time-variable since the term t^2d^2y/dt^2 depends explicitly on t through the coefficient t^2.

An example of a time-varying system is a spacecraft system which the mass of spacecraft changes during flight due to fuel consumption.

Definition 1-12: A time-invariant differential equation is a differential equation in which none of its coefficients depend on the independent time variable, t. For example, the differential equation:

$$m\frac{d^2y(t)}{dt^2} + b\frac{dy(t)}{dt} + y(t) = u(t) \qquad (1-101)$$

where the coefficients m and b are constants, is time-invariant since the equation depends only implicitly on t through the dependent variables y and u and their derivatives.

Dynamical systems that are described by linear constant coefficient differential equations are called linear time-invariant (LTI) systems.

CHAPTER TWO

Transfer Function and Impulse Response Function

In control theory, functions called transfer functions are commonly used to characterize the input-output relationships of components or systems that can be described by linear, time-invariant, differential equations. We begin by defining the transfer function and follow with a derivation of the transfer function of a differential equation system. Then we discuss the impulse-response function.

Transfer functions (G) are frequently used to characterize the input-output relationships or systems that can be described by Linear Time-Invariant (LTI) differential equations.

The main reasons why transfer functions are useful are as follows:

- **Compact model form:** If the original model is a higher order differential equation, or a set of first order differential equations, the relation between the input variable and the output variable can be described by one transfer function, which is a rational function of the Laplace variable s, without any time-derivatives.
- **Representation of standard models:** Transfer functions are often used to represent standard models of controllers and signal filters.
- **Simple to combine systems:** For example, the transfer function for a combined system which consists of two systems in a series combination is just the product of the transfer functions of each system.
- Simple to calculate time responses: The calculations will be made using the Laplace transform, and the necessary mathematical operations are usually much simpler than solving differential equations.
- Simple to find the frequency response: The frequency response is a function which expresses how sinusoid signals are transferred through a dynamic system. Frequency response is an important tool in analysis and design of signal filters and control systems. The frequency response can be found from the transfer function of the system.

2-1 Transfer Function (Single-Input, Single-Output Systems)

Definition 1: The transfer function of a linear, time-invariant, differential equation system is defined as the ratio of the Laplace transform of the output (response function) to the Laplace transform of the input (driving function) under the assumption that all initial conditions are zero.

Let $G(s)$ denote the transfer function of a single-input, single-output (SISO) system, with input $u(t)$ output $y(t)$, and impulse response $g(t)$. The transfer function $G(s)$ is defined as

$$G(s) = \mathcal{L}[g(t)] \qquad (2-1)$$

Consider the linear time-invariant system defined by the following differential equation:

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}\dot{y} + a_n y = b_0 x^m + b_1 x^{(m-1)} + \cdots + b_{m-1}\dot{x} + b_m x \quad (n \geq m) \quad (2-2)$$

Where y is the output of the system and x is the input.The transfer function of this system is the ratio of the Laplace transformed output to the Laplace transformed input when all initial conditions are zero, or

$$\text{Transfer function} = G(s) = \frac{\mathcal{L}[\text{output}]}{\mathcal{L}[\text{input}]}\Big|_{\text{zero initial conditions}}$$

$$= \frac{Y(s)}{X(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_1 s + a_n} \qquad (2-3)$$

The above equation can be represented by the following graphical representation:

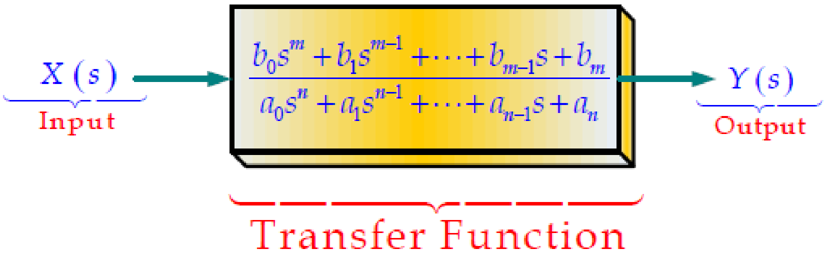


Figure 2-1 Block diagram representation of a transfer function

Proper Transfer Functions. The transfer function is said to be strictly proper if the order of the denominator polynomial is greater than that of the numerator polynomial (i.e., $n > m$). If $n = m$, the transfer function is called proper. The transfer function is improper if $m > n$.

Characteristic Equation. The characteristic equation of a linear system is defined as the equation obtained by setting the denominator polynomial of the transfer function to zero. Thus, the characteristic equation of the system described by Eq. (2-2) is

$$a_0 s^n + a_1 s^{n-1} + \cdots + a_1 s + a_n = 0 \qquad (2-4)$$

Comments on the Transfer Function (TF). The applicability of the concept of the transfer function is limited to linear, time-invariant, differential equation systems. The transfer function approach, however, is extensively used in the analysis and design of such systems. We shall list important comments concerning the transfer function. (Note that a system referred to in the list is one described by a linear, time-invariant, differential equation.)

1. The transfer function of a system is a mathematical model in that it is an operational method of expressing the differential equation that relates the output variable to the input variable.

2. The transfer function is a property of a system itself, independent of the magnitude and nature of the input or driving function

3. The transfer function includes the units necessary to relate the input to the output; however, it does not provide any information concerning the physical structure of the system. (The transfer functions of many physically different systems can be identical.)

4. If the transfer function of a system is known, the output or response can be studied for various forms of inputs with a view toward understanding the nature of the system.

5. If the transfer function of a system is unknown, it may be established experimentally by introducing known inputs and studying the output of the system. Once established, a transfer function gives a full description of the dynamic characteristics of the system, as distinct from its physical description.

A transfer function has the following properties:

- The transfer function is defined only for a linear time-invariant system. It is not defined for nonlinear systems.
- The transfer function between a pair of input and output variables is the ratio of the Laplace transform of the output to the Laplace transform of the input.
- All initial conditions of the system are set to zero.
- The transfer function is independent of the input of the system.

To derive the transfer function of a system, we use the following procedures:

1. Develop the differential equation for the system by using the physical laws, e.g. Newton's laws and Kirchhoff's laws.
2. Take the Laplace transform of the differential equation under the zero initial conditions.
3. Take the ratio of the output $Y(s)$ to the input $U(s)$. This ratio is the transfer function.

2-2 Transfer Function (Multivariable Systems)

The definition of a transfer function is easily extended to a system with multiple inputs and outputs. A system of this type is often referred to as a multivariable system.

In general, if a linear system has p inputs and q outputs, the transfer function between the jth input and the ith output is defined as

$$G_{ij}(s) = \frac{Y_i(s)}{R_j(s)} \qquad (2-5)$$

with $R_k(s) = 0, k = 1,2,\ldots,p, k \neq j$.

When all the p inputs are in action, the ith output transform is written

$$Y_i(s) = G_{i1}(s)R_1(s) + G_{i2}(s)R_2(s) + \cdots + G_{ip}(s)R_p(s) \qquad (2-6)$$

It is convenient to express Eq. (2-6) in matrix-vector form:

$$Y(s) = G(s)R(s) \qquad (2-7)$$

Where

$$Y(s) = \begin{bmatrix} Y_1(s) \\ Y_2(s) \\ \vdots \\ Y_q(s) \end{bmatrix} \qquad (2-8)$$

is the $q \times 1$ transformed output vector;

$$R(s) = \begin{bmatrix} R_1(s) \\ R_2(s) \\ \vdots \\ R_p(s) \end{bmatrix} \qquad (2-9)$$

is the $p \times 1$ transformed input vector; and

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1p}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2p}(s) \\ \vdots & \vdots & \cdots & \vdots \\ G_{q1}(s) & G_{q1}(s) & \cdots & G_{qp}(s) \end{bmatrix} \qquad (2-10)$$

is the $q \times p$ transfer-function matrix.

2-3 Stability of Linear Control Systems

From the studies of linear differential equations with constant coefficients of SISO systems, we learned that the homogeneous solution that corresponds to the transient response of the system is governed by the roots of the characteristic equation. Basically, the design of linear control systems may be regarded as a problem of arranging the location of the poles and zeros of the system transfer function such that the system will perform according to the prescribed specifications.

Among the many forms of performance specifications used in design, the most important requirement is that the system must be stable. An unstable system is generally considered to be useless.

When all types of systems are considered-linear, nonlinear, time-invariant, and time-varying-the definition of stability can be given in many different forms. We shall deal only with the stability of linear SISO time-invariant systems in the following discussions.

For analysis and design purposes, we can classify stability as absolute stability and relative stability. Absolute stability refers to whether the system is stable or unstable; it is a yes or no answer. Once the system is found to be stable, it is of interest to determine how stable it is, and this degree of stability is a measure of relative stability.

In preparation for the definition of stability, we define the two following types of responses for linear time-invariant systems:

i. **Zero-state response.** The zero-state response is due to the input only; all the initial conditions of the system are zero.
ii. **Zero-input response.** The zero-input response is due to the initial conditions only all the inputs are zero.

From the principle of superposition, when a system is subject to both inputs and initial conditions, the total response is written

$$Total\ response = zero - state\ response + \ zero - input\ response$$

The definitions just given apply to continuous-data as well as discrete-data systems.

2-4 Bounded-input, Bounded-output (BIBO Stability-Continuous-Data Systems

Let $u(t)$, $y(t)$, and $g(t)$ be the input, output, and the impulse response of a linear time-invariant system, respectively. With zero initial conditions, the system is said to be BJBO (bounded-input, bounded-output) stable, or simply stable, if its output $y(t)$ is bounded to a bounded input $u(t)$.

The convolution integral relating $u(t), y(t)$, and $g(t)$ is

$$y(t) = \int_0^\infty u(t-\tau)\, g(\tau) d\tau \qquad (2-11)$$

Taking the absolute value of both sides of the equation, we get

$$|y(t)| = \left| \int_0^\infty u(t-\tau)\, g(\tau) d\tau \right| \qquad (2-12)$$

or

$$|y(t)| \leq \int_0^\infty |u(t-\tau)g(\tau)d\tau| \qquad (2-13)$$

If $u(t)$ is bounded,

$$|u(t)| \leq M \qquad (2-14)$$

where M is a finite positive number. Then,

$$|y(t)| \leq M\int_0^\infty |g(\tau)|d\tau \qquad (2-15)$$

Thus, if $y(t)$ is to be bounded, or

$$|y(t)| \leq N < \infty \qquad (2-16)$$

where N is a finite positive number, the following condition must hold:

$$M\int_0^\infty |g(\tau)|d\tau \leq N < \infty \qquad (2-17)$$

Or, for any finite positive Q,

$$\int_0^\infty |g(\tau)|d\tau \leq Q < \infty \qquad (2-18)$$

The condition given in Eq. (2-18) implies that the area under the $|g(\tau)|$-versus-τ-curve must be finite.

2-5 Relationship Between characteristic equation Roots and Stability

To show the relation between the roots of the characteristic equation and the condition in Eq. (2-18), we write the transfer function G(s), according to the Laplace transform definition, as

$$G(s) = \mathcal{L}[g(\tau)] = \int_0^x g(t)e^{-st}dt \qquad (2-19)$$

Taking the absolute value on both sides of the last equation, we have

$$|G(s)| = \left|\int_0^x g(t)e^{-st}dt\right| \leq \int_0^\infty |g(t)||e^{-st}|dt \qquad (2-20)$$

Because $|e^{-st}| = |e^{-\sigma t}|$, where a is the real part of s, when s assumes a value of a pole of $G(s), G(s) = \infty$, Eq. (2-20) becomes

$$\infty \leq \int_0^x |g(t)||e^{-\sigma t}|dt \qquad (2-21)$$

If one or more roots of the characteristic equation are in the $\text{right} - \text{half s} -$ plane or on the $\text{j}\omega - \text{axis}, \sigma \geq 0$, then

$$|e^{-\sigma t}| \leq M = 1 \qquad (2-22)$$

Eq. (2-21) becomes

$$\infty \leq \int_0^x M|g(t)|dt = \int_0^x |g(t)|dt \qquad (2-23)$$

which violates the BIBO stability requirement. Thus, for BIBO stability, the roots of the characteristic equation, or the poles $G(s)$, cannot be located in the right-half s-plane or on the $\text{j}\omega -$ axis; in other words, they must all lie in the left-half s-plane. A system is said to be unstable if it is not BIBO stable. When a system has roots on the $j\omega - axis$, say, at $s = j\omega_0$ and $s = -j\omega_0$, if the input is a sinusoid, $\sin \omega_0 t$, then the output will be of the form of $t\ \sin \omega_0 t$, which is unbounded, and the system is unstable.

2-6 Zero-input and Asymptotic Stability of Continuous-Data Systems

We shall define zero-input stability and asymptotic stability and establish their relations with BIBO stability.

Zero-input stability refers to the stability condition when the input is zero, and the system is driven only by its initial conditions. We shall show that the zero-input stability also depends on the roots of the characteristic equation.

Let the input of an $nth - order$ system be zero and the output due to the initial conditions be $y(t)$. Then, $y(t)$ can be expressed as

$$y(t) = \sum_{k=0}^{n-1} g_k(t) y^{(k)}(t_0) \qquad (2-24)$$

where

$$y^{(k)}(t_0) = \left.\frac{d^k y(t)}{dt^k}\right|_{t=t_0} \qquad (2-25)$$

and $g_k(t)$ denotes the zero-input response due to $y^{(k)}(t_0)$. The zero-input stability is defined as follows: If the zero-input response $y(t)$, subject to the finite initial conditions, $y^k(t_0)$, reaches zero as t approaches infinity, the system is said to be zero-input stable, or stable; otherwise, the system is unstable.

Mathematically, the foregoing definition can be stated: A linear time-invariant system is zero-input stable if, for any set of finite $y^{(k)}(t_0)$, there exists a positive number M, which depends on $y^{(k)}(t_0)$, such that

$$|y(t)| \leq M < \infty \qquad for\ all\ t \geq t_0 \qquad (2-26)$$

$$\lim_{t\to\infty} |y(t)| = 0 \qquad (2-27)$$

Because the condition in the last equation requires that the magnitude of $y(t)$ reaches zero as time approaches infinity, the zero-input stability is also known at the **asymptotic stability**.

Taking the absolute value on both sides of Eq. (2-24), we get

$$|y(t)| = \left|\sum_{k=0}^{n-1} g_k(t) y^{(k)}(t_0)\right| \leq \sum_{k=0}^{n-1} |g_k(t)|\left|y^{(k)}(t_0)\right| \qquad (2-28)$$

Because all the initial conditions are assumed to be finite, the condition in Eq. (2-26) requires that the following condition be true:

$$\sum_{k=0}^{n-1} |g_k(t)| < \infty \quad for\ all\ \ t \geq 0 \qquad (2-29)$$

Let the n characteristic equation roots be expressed as $s_i = \sigma_i + j\omega,\ i = 1,2.\dots.n$. Then, if m of the n roots are simple, and the rest are of multiple order, $y(t)$ will be of the form:

$$\sum_{i=1}^{m} K_i e^{s_i t} + \sum_{i=0}^{n-m-1} L_i t^i e^{s_i t} \qquad (2-30)$$

where K_i and L_i are constant coefficients. Because the exponential terms $e^{s_i t}$ in the last equation control the response $y(t)$ as $t \to \infty$, to satisfy the two conditions in Eqs. (2-26) and (2-27), the real parts of s_i must be negative. In other words, the roots of the characteristic equation must all be in the left-half s-plane.

From the preceding discussions, we see that, for linear time-invariant systems, BIBO, zero-input, and asymptotic stability all have the same requirement that the roots of the characteristic equation must all be located in the left-half s-plane. Thus, if a system is BIBO stable, it must also be zero-input or asymptotically stable. For this reason, we shall simply refer to the stability condition of a linear system as stable or unstable. The latter condition refers to the condition that at least one of the characteristic equation roots is not in the left-half s-plane. For practical reasons, we often refer to the situation in which the characteristic equation has simple roots on the $j\omega -$ axis and none in the right-half plane as marginally stable or marginally unstable. An exception to this is if the system were intended to be an integrator (or, in the case of control systems, a velocity control system); then the system would have root(s) at $s = 0$ and would be considered stable. Similarly, if the system were designed to be an oscillator, the characteristic equation would have simple roots on the $j\omega - axis$, and the system would be regarded as stable.

Because the roots of the characteristic equation are the same as the eigenvalues of A of the state equations, the stability condition places the same restrictions on the eigenvalues.

Let the characteristic equation roots or eigenvalues of A of a linear continuous-data time-invariant SISO system be $s_i = \sigma_i + j\omega_i,\ i = 1,2,\ldots,n$. If any of the roots is complex, it is in complex-conjugate pairs. The possible stability conditions of the system are summarized in Table 2-1 with respect to the roots of the characteristic equation.

The following example illustrates the stability conditions of systems with reference to the poles of the system transfer functions that are also the roots of the characteristic equation.

TABLE 2-1 Stability Conditions of Linear Continuous-Data Time-Invariant SISO Systems

Stability Condition	Root Values
Asymptotically stable or simply stable	$\sigma_i = 0$ for all $i, i = 1,2,\ldots,n$. (All the roots are in the left-half .s-plane.)
Marginally stable or marginally unstable	$\sigma_i = 0$ for any i for simple roots, and no $\sigma_i > 0$ For $i = 1,2,\ldots,n$ (at least one simple root, no multiple-order roots on the $j\omega -$ axis, and n roots in the right-half $s - plane$; note exceptions)
Unstable	$\sigma_i > 0$ for any i or $\sigma_i = 0$ 0 for any multiple-order root $i = 1,2,\ldots,n$ (at least one simple root in the right-half s-plane or at least one multiple-order

root on the $j\omega - axis$)

Example 2-1: [59] The following closed-loop transfer functions and their associated stability conditions are given.

$$M(s) = \frac{20}{(s+1)(s+2)(s+3)}$$

BIBO or asymptotically stable (or, simply, stable)

$$M(s) = \frac{20(s+1)}{(s-1)(s^2+2s+2)}$$

Unstable due to the pole at s = 1

$$M(s) = \frac{20(s-1)}{(s+2)(s^2+4)}$$

Marginally stable or marginally unstable due to s = ±j2

$$M(s) = \frac{10}{(s^2+4)^2(s+10)}$$

Unstable due to the multiple-order pole at s = ± j2

$$M(s) = \frac{10}{s^4+30s^3+s^2+10s}$$

Stable if the pole at s = 0 is placed intentionally

Example 2-2: Consider the mechanical system shown in Figure 2.2. The displacement x of the mass m is measured from the equilibrium position. In this system, the external force $f\ (t)$ is input and x is the output.

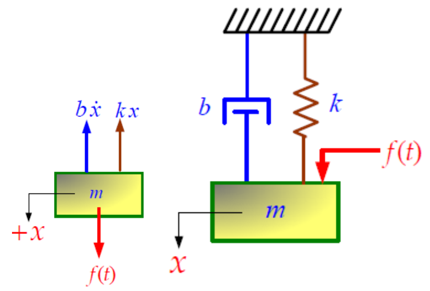


Figure 2-2 Mass -Spring –Damper System and the FBD.

i. The FBD is shown in the Fig. 2-2.

ii. Apply Newton's second law of motion to a system in translation:

$$\underbrace{\sum F}_{\substack{Summation\ of\ all\ forces \\ acting\ on\ the\ system}} = m\ddot{x}$$

$$f(t) - b\dot{x} - kx = m\ddot{x}$$

or

$$m\ddot{x} + b\dot{x} + kx = f(t) \implies \text{Forced Vibration of a second order system}$$

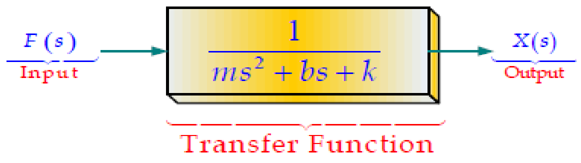
$$m\ddot{x} + b\dot{x} + kx = f(t) \implies \text{Forced Vibration of a second order system}$$

iii. For zero Initial Conditions (I.C's), taking Laplace Transform (LT) of both sides of the above equation yields

$$(ms^2 + bs + k)X(s) = F(s)$$

where $X(s) = \mathcal{L}[x(t)]$ and $F(s) = \mathcal{L}[f(t)]$. From Equation (2-3), the TF for the system is

$$\frac{X(s)}{F(s)} = \frac{Output}{Input} = \frac{1}{(ms^2 + bs + k)}$$



Example 2-3: [6] The transfer functions $X_1(s)/U(s)$ and $X_2(s)/U(s)$ of the mechanical system shown in Figure 2-3.

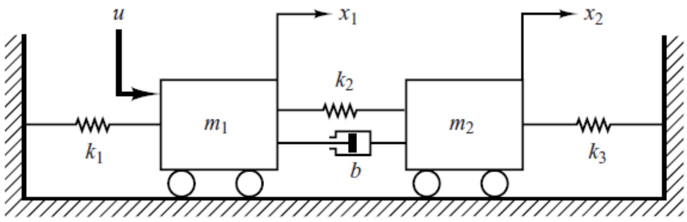


Figure 2-3 Mechanical system.

The equations of motion for the system shown in Figure 2.3 are

$$m_1\ddot{x}_1 = -k_1x_1 - k_2(x_1 - x_2) - b(\dot{x}_1 - \dot{x}_2) + u$$
$$m_2\ddot{x}_2 = -k_3x_2 - k_2(x_2 - x_1) - b(\dot{x}_2 - \dot{x}_1)$$

Simplifying, we obtain

$$m_1\ddot{x}_1 + b\dot{x}_1 + (k_1 + k_2)x_1 = b\dot{x}_2 + k_2x_2 + u$$
$$m_2\ddot{x}_2 + b\dot{x}_2 + (k_2 + k_3)x_2 = b\dot{x}_1 + k_2x_1$$

Taking the Laplace transforms of these two equations, assuming zero initial conditions, we obtain

$$[m_1s^2 + bs + (k_1 + k_2)X_1(s) = (bs + k_2)X_2(s) + U(s) \qquad (2-31)$$
$$[m_2s^2 + bs + (k_2 + k_3)]X_2(s) = (bs + k_2)X_1(s) \qquad (2-32)$$

Solving Equation (2-32) for and substituting it into Equation (2-31) and simplifying, we get

$$[(m_1s^2 + bs + k_1 + k_2)(m_2s^2 + bs + k_2 + k_3) - (bs + k_2)^2]X_1(s)$$
$$= (m_2s^2 + bs + k_2 + k_3)U(s)$$

from which we obtain

$$\frac{X_1(s)}{U(s)} = \frac{m_2s^2 + bs + k_2 + k_3}{(m_1s^2 + bs + k_1 + k_2)(m_2s^2 + bs + k_2 + k_3) - (bs + k_2)^2} \qquad (2-33)$$

From Equations (2–32) and (2–33) we have

$$\frac{X_2(s)}{U(s)} = \frac{bs + k_2}{(m_1s^2 + bs + k_1 + k_2)(m_2s^2 + bs + k_2 + k_3) - (bs + k_2)^2} \qquad (2-34)$$

Equations (2-33) and (2-34) are the transfer functions $X_1(s)/U(s)$ and $X_2(s)/U(s)$, respectively.

2-7 Transfer Functions of Cascaded Elements

Many feedback systems have components that load each other. Consider the system shown in Figure 2-4. Assume that e_i is the input and e_0 is the output. The capacitances C_1 and C_2 are not charged initially.

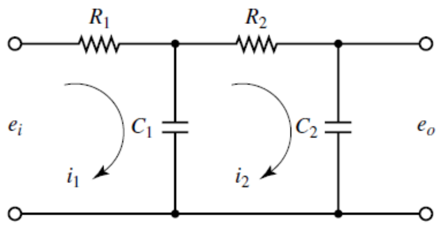


Figure 2–4 Electrical system.

It will be shown that the second stage of the circuit (R_2C_2 portion) produces a loading effect on the first stage (R_1C_1 portion).The equations for this system are

$$\frac{1}{C_1}\int (i_1 - i_2)dt + R_1 i_1 = e_i \qquad (2-35)$$

and

$$\frac{1}{C_1}\int (i_2 - i_1)dt + R_2 i_2 + \frac{1}{C_2}\int i_2 dt = 0 \qquad (2-36)$$

$$\frac{1}{C_2}\int i_2 dt = e_0 \qquad (2-37)$$

Taking the Laplace transforms of Equations (2–35) through (2–37), respectively, using zero initial conditions, we obtain

$$\frac{1}{C_1 s}[I_1(s) - I_2(s)] + R_1 I_1(s) = E_i(s) \qquad (2-38)$$

$$\frac{1}{C_1 s}[I_2(s) - I_1(s)] + R_2 I_2(s) + \frac{1}{C_2 s} I_2(s) = 0 \qquad (2-39)$$

$$\frac{1}{C_2 s} I_2(s) = E_0 \qquad (2-40)$$

Eliminating $I_1(s)$ from Equations (2–38) and (2–39) and writing $E_i(s)$ in terms of $I_2(s)$, we find the transfer function between $E_0(s)$ and $E_i(s)$ to be

$$\frac{E_0(s)}{E_i} = \frac{1}{(R_1C_1s + 1)(R_2C_2 + 1) + R_1C_2s}$$

$$= \frac{1}{R_1C_1R_2s^2 + (R_1C_1 + R_2C_2 + R_1C_2)s + 1} \qquad (2-41)$$

The term R_1C_2s in the denominator of the transfer function represents the interaction of two simple *RC* circuits. $(R_1C_1 + R_2C_2 + R_1C_2)^2 > 4R_1C_1R_2C_2$ Since the two roots of the denominator of Equation (2–41) are real.

The present analysis shows that, if two RC circuits are connected in cascade so that the output from the first circuit is the input to the second, the overall transfer function is not the product of $1/(R_1C_1s + 1)$and $1/(R_2C_2s + 1)$. The reason for this is that, when we derive the transfer function for an isolated circuit, we implicitly assume that the output is unloaded. In other words, the load impedance is assumed to be infinite, which means that no power is being withdrawn at the output. When the second circuit is connected to the output of the first, however, a certain amount of power is withdrawn, and thus the assumption of no loading is violated. Therefore, if the transfer function of this system is obtained under the assumption of no loading, then it is not valid. The degree of the loading effect determines the amount of modification of the transfer function.

Complex Impedances. In deriving transfer functions for electrical circuits, we frequently find it convenient to write the Laplace-transformed equations directly, without writing the differential equations. Consider the system shown in Figure 2–6(a). In this system, Z_1 and Z_2 represent complex impedances. The complex impedance $Z(s)$ of a two-terminal circuit is the ratio of $E(s)$, the Laplace transform of the voltage across the terminals, to $I(s)$, the Laplace transform of the current through the element, under the assumption that the initial conditions are zero, so that $Z(s) = E(s)/I(s)$. If the two-terminal element is a resistance R, capacitance C, or inductance L, then the complex impedance is given by R, $1/Cs$, or Ls, respectively. If complex impedances are connected in series, the total impedance is the sum of the individual complex impedances.

Remember that the impedance approach is valid only if the initial conditions involved are all zeros. Since the transfer function requires zero initial conditions, the impedance approach can be applied to obtain the transfer function of the electrical circuit. This approach greatly simplifies the derivation of transfer functions of electrical circuits.

Consider the circuit shown in Figure 2–6(b). Assume that the voltages e_i and e_o are the input and output of the circuit, respectively. Then the transfer function of this circuit is

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2(s)}{Z_1(s) + Z_2(s)}$$

For the system shown in Figure 1–32,

$$Z_1 = Ls + R, \qquad Z_2 = \frac{1}{Cs}$$

Hence the transfer function $E_o(s)/E_i(s)$ can be found as follows:

$$\frac{E_0(s)}{E_i(s)} = \frac{\frac{1}{Cs}}{Ls + R + \frac{1}{Cs}} = \frac{1}{LCs^2 + RCs + 1}$$

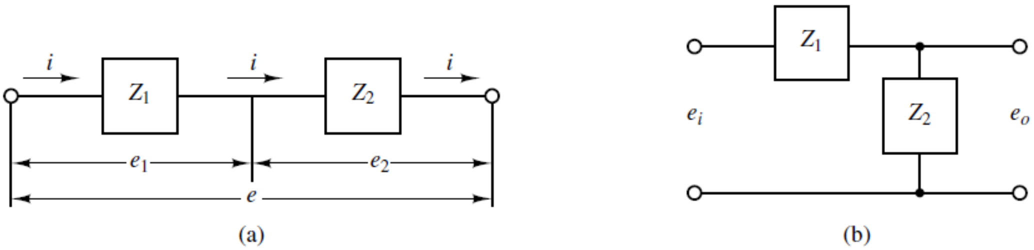


Figure 2-5 Electrical circuits.

Example 2-4: [6] Consider again the system shown in Figure 2–4. Obtain the transfer function $E_o(s)/E_i(s)$ by use of the complex impedance approach. (Capacitors C_1 and C_2 are not charged initially.)

The circuit shown in Figure 2–4 can be redrawn as that shown in Figure 2–6(a), which can be further modified to Figure 2–6(b).

In the system shown in Figure 2–6(b) the current I is divided into two currents I_1 and I_2 . Noting that

$$Z_2 I_1 = (Z_3 + Z_4) I_2, \quad I_1 + I_2 = I$$

we obtain

$$I_1 = \frac{Z_3 + Z_4}{Z_2 + Z_3 + Z_4} I, \quad I_2 = \frac{Z_2}{Z_2 + Z_3 + Z_4} I$$

Noting that

$$E_i(s) = Z_1 I + Z_2 I = \left[Z_1 + \frac{Z_2(Z_3 + Z_4)}{Z_2 + Z_3 + Z_4} \right] I$$

$$E_0(s) = Z_4 I_2 = \frac{Z_2 Z_4}{Z_2 + Z_3 + Z_4} I$$

we obtain

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2 Z_4}{Z_1(Z_2 + Z_3 + Z_4) + Z_2(Z_3 + Z_4)}$$

Substituting $Z_1 = R_1, Z_2 = 1/(C_1 s), Z_3 = R_2$, and $Z_4 = 1/(C_2 s)$ into this last equation, we get

$$\frac{E_0(s)}{E_i(s)} = \frac{\frac{1}{C_1 s}\frac{1}{C_2 s}}{R_1\left(\frac{1}{C_1 s} + R_2 + \frac{1}{C_2 s}\right) + \frac{1}{C_1 s}\left(R_2 + \frac{1}{C_2 s}\right)}$$

$$= \frac{1}{R_1 C_1 R_2 C_2 s^2 + (R_1 C_1 + R_2 C_2 + R_1 C_2)s + 1}$$

which is the same as that given by Equation (2–41).

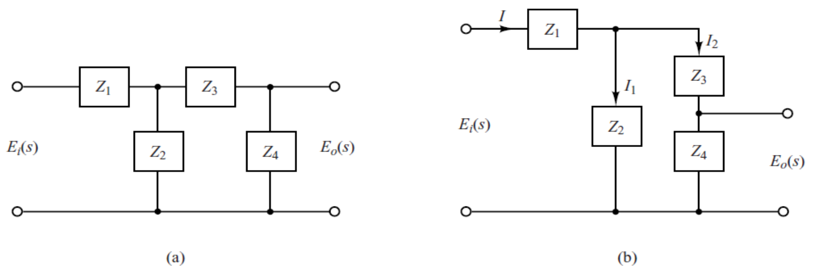


Figure 2-6 (a) The circuit of Figure 2-4 shown in terms of impedances;
(b) Equivalent circuit diagram.

Figure 2-7 (a) System consisting of two nonloading cascaded elements; (b) an equivalent system.

Transfer Functions of Nonloading Cascaded Elements. The transfer function of a system consisting of two nonloading cascaded elements can be obtained by eliminating the intermediate input and output. For example, consider the system shown in Figure 2–7(a).The transfer functions of the elements are

$$G_1 = \frac{X_2(s)}{X_1(s)} \quad \text{and} \quad G_2 = \frac{X_3(s)}{X_2(s)}$$

If the input impedance of the second element is infinite, the output of the first element is not affected by connecting it to the second element. Then the transfer function of the whole system becomes

$$G(s) = \frac{X_3(s)}{X_1(s)} = \frac{X_2(s) X_3(s)}{X_1(s) X_2(s)} = G_1(s) G_2(s)$$

The transfer function of the whole system is thus the product of the transfer functions of the individual elements. This is shown in Figure 2-7(b).

As an example, consider the system shown in Figure 2–8.The insertion of an isolating amplifier between the circuits to obtain nonloading characteristics is frequently used in combining circuits. Since amplifiers have very high input impedances, an isolation amplifier inserted between the two circuits justifies the nonloading assumption.

The two simple RC circuits, isolated by an amplifier as shown in Figure 2–8, have negligible loading effects, and the transfer function of the entire circuit equals the product of the individual transfer functions. Thus, in this case,

$$\frac{E_o(s)}{E_i(s)} = \left(\frac{1}{R_1C_1s+1}\right)(k)\left(\frac{1}{R_2C_2s+1}\right)$$
$$= \frac{k}{(R_1C_1s+1)(R_2C_2s+1)}$$

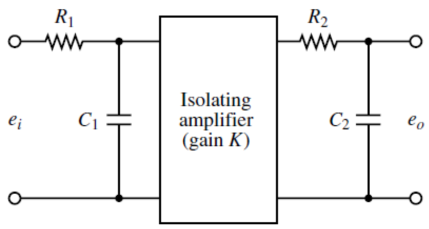


Figure 2-8 Electrical system.

Electronic Controllers. We shall discuss electronic controllers using operational amplifiers. We begin by deriving the transfer functions of simple operational amplifier circuits. Then we derive the transfer functions of some of the operational-amplifier controllers. Finally, we give operational-amplifier controllers and their transfer functions in the form of a table.

Operational Amplifiers. Operational amplifiers, often called op amps, are frequently used to amplify signals in sensor circuits. Op amps are also frequently used for compensation purposes. Figure 2–9 shows an op amp. It is a common practice to choose the ground as 0 volt and measure the input voltages e_1 and e_2 relative to the ground. The input e_1 to the minus terminal of the amplifier is inverted, and the input e_2 to the plus terminal is not inverted. The total input to the amplifier thus becomes $e_2 - e_1$. Hence, for the circuit shown in Figure 2–9, we have

$$e_o = K(e_2 - e_1) = -K(e_1 - e_2)$$

where the inputs e_1 and e_2 may be dc or ac signals and K is the differential gain (voltage gain). The magnitude of K is approximately 105 ~ 106 for dc signals and ac signals with frequencies less than approximately 10 Hz. (The differential gain K decreases with the signal frequency and becomes about unity for frequencies of 1 MHz ~ 50 MHz.) Note that the op amp amplifies the difference in voltages e_1 and e_2 . Such an amplifier is commonly called a differential amplifier. Since the gain of the op amp is very high, it is necessary to have a negative feedback from the output to the input to make the amplifier stable. (The feedback is made from the output to the inverted input so that the feedback is a negative feedback.)

In the ideal op amp, no current flows into the input terminals, and the output voltage is not affected by the load connected to the output terminal. In other words, the input impedance is infinity and the output impedance is zero. In an actual op amp, a very small (almost negligible) current flows into an input terminal and the output cannot be loaded too much. In our analysis here, we make the assumption that the op amps are ideal.

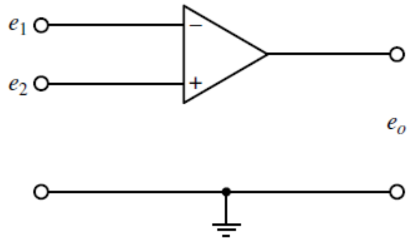


Figure 2-9 Operational amplifier.

Inverting Amplifier. Consider the operational-amplifier circuit shown in Figure 2–10. Let us obtain the output voltage e_o.

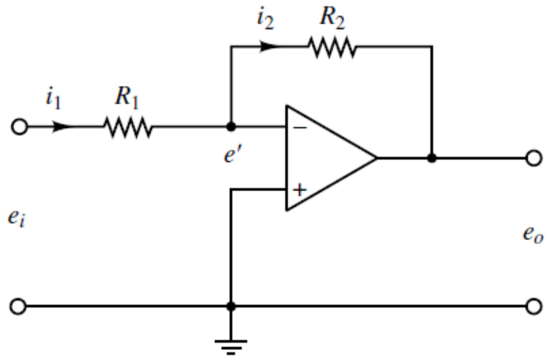


Figure 2-10 Inverting amplifier.

The equation for this circuit can be obtained as follows: Define

$$i_1 = \frac{e_i - e'}{R_1}, \quad i_2 = \frac{e' - e_o}{R_2}$$

Since only a negligible current flows into the amplifier, the current i_1 must be equal to current i_2 . Thus

$$\frac{e_i - e'}{R_1} = \frac{e' - e_o}{R_2}$$

Since $K(0 - e') = e_o$ and $K \gg 1, e'$ must be almost zero, or $e' \doteqdot 0$. Hence we have

$$\frac{e_i}{R_1} = \frac{-e_o}{R_2}$$

or

$$e_o = -\frac{R_2}{R_1} e_i$$

Thus the circuit shown is an inverting amplifier. If $R_1 = R_2$,then the op-amp circuit shown acts as a sign inverter.

Noninverting Amplifier. Figure 2–11(a) shows a noninverting amplifier. A circuit equivalent to this one is shown in Figure 2–11(b). For the circuit of Figure 2–11(b), we have

$$e_o = K\left(e_i - \frac{R_1}{R_1 + R_2} e_o\right)$$

where K is the differential gain of the amplifier. From this last equation, we get

$$e_i = \left(\frac{R_1}{R_1 + R_2} + \frac{1}{K}\right) e_o$$

Since $K \gg 1, if\ \ R_1/(R_1 + R_2) \gg 1/K$,then

$$e_o = \left(1 + \frac{R_2}{R_1}\right) e_i$$

This equation gives the output voltage e_o. Since e_o and e_i have the same signs, the op-amp circuit shown in Figure 2–11(a) is noninverting.

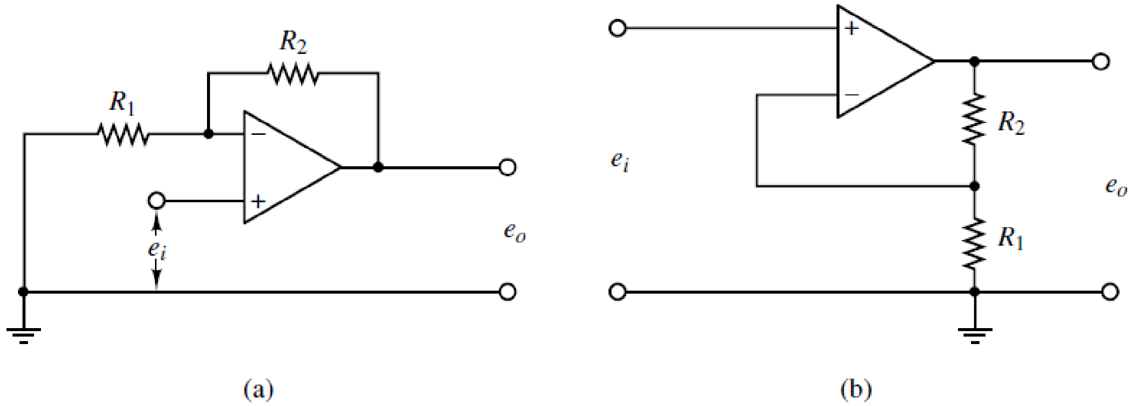


Figure 2-11 (a) Noninverting operational amplifier; (b) equivalent circuit.

Example 2-5:[6] Figure 2-12 shows an electrical circuit involving an operational amplifier. Obtain the output e_o. Let us define

$$i_1 = \frac{e_i - e'}{R_1}, \quad i_2 = C\frac{d(e' - e_o)}{dt}, \quad i_3 = \frac{e' - e_o}{R_2}$$

Noting that the current flowing into the amplifier is negligible, we have

$$i_1 = i_2 + i_3$$

Hence

$$\frac{e_i - e'}{R_1} = C\frac{d(e' - e_o)}{dt} + \frac{e' - e_o}{R_2}$$

Since $e' \doteqdot 0$, we have

$$\frac{e_i}{R_1} = -C\frac{de_o}{dt} - \frac{e_o}{R_2}$$

Taking the Laplace transform of this last equation, assuming the zero initial condition, we have

$$\frac{E_i(s)}{R_1} = -\frac{R_2Cs + 1}{R_2}E_o(s)$$

which can be written as

$$\frac{E_o(s)}{E_i(s)} = -\frac{R_2}{R_1}\frac{1}{R_2Cs + 1}$$

The op-amp circuit shown in Figure 2–12 is a first-order lag circuit. (Several other circuits involving op amps are shown in Table 2–2 together with their transfer functions.)

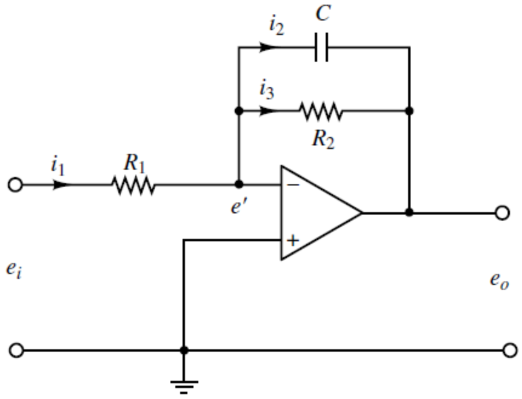


Figure 2-12 First-order lag circuit using operational amplifier.

Impedance Approach to Obtaining Transfer Functions. Consider the op-amp circuit shown in Figure 2-13. Similar to the case of electrical circuits we discussed earlier, the impedance approach can be applied to op-amp circuits to obtain their transfer functions. For the circuit shown in Figure 2–13, we have

$$\frac{E_i(s) - E'(s)}{Z_1} = \frac{E'(s) - E_o}{Z_2}$$

Since $E'(s) \doteqdot 0$, we have

$$\frac{E_o(s)}{E_i(s)} = -\frac{Z_2(s)}{Z_1(s)} \qquad (2-42)$$

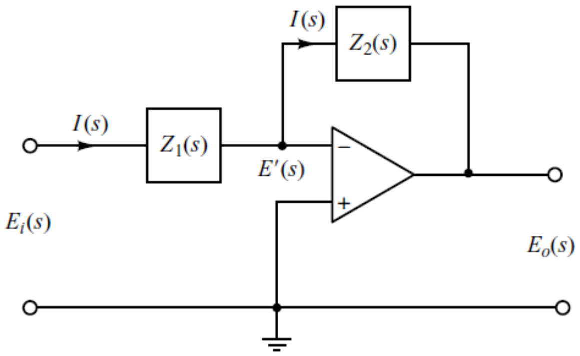


Figure 2-13 Operational amplifier circuit.

Example 2-6: [6] Referring to the op-amp circuit shown in Figure 2–13, obtain the transfer function $E_o(s)/E_i(s)$ by use of the impedance approach.

The complex impedances $Z_1(s)$ and $Z_2(s)$ for this circuit are

$$Z_1(s) = R_1 \quad and \quad Z_2(s) = \frac{1}{Cs + \frac{1}{R_2}} = \frac{R_2}{R_2Cs + 1}$$

The transfer function $E_o(s)/E_i(s)$ is, therefore, obtained as

$$\frac{E_o(s)}{E_i(s)} = -\frac{Z_2(s)}{Z_1(s)} = -\frac{R_2}{R_1}\frac{1}{R_2Cs + 1}$$

which is, the same as that obtained in Example 2-5.

Lead or Lag Networks Using Operational Amplifiers. Figure 2–14(a) shows an electronic circuit using an operational amplifier. The transfer function for this circuit can be obtained as follows: Define the input impedance and feedback impedance as Z_1 and Z_2 , respectively. Then

$$Z_1 = \frac{R_1}{R_1C_1s + 1}, \qquad Z_2 = \frac{R_2}{R_2C_2 + 1}$$

Hence, referring to Equation (2–42), we have

$$\frac{E(s)}{E_i(s)} = -\frac{Z_2}{Z_1} = -\frac{R_2}{R_1}\frac{R_1C_1s+1}{R_2C_2s+1} = -\frac{C_1}{C_2}\frac{s+1/R_1C_1}{s+1/R_2C_2} \qquad (2-43)$$

Notice that the transfer function in Equation (2–43) contains a minus sign. Thus, this circuit is sign inverting. If such a sign inversion is not convenient in the actual application, a sign inverter may be connected to either the input or the output of the circuit of Figure 2–14(a). An example is shown in Figure 2–14(b).The sign inverter has the transfer function of

$$\frac{E_o(s)}{E(s)} = -\frac{R_4}{R_3}$$

The sign inverter has the gain of $-R_4/R_3$. Hence the network shown in Figure 2–14(b) has the following transfer function:

$$\frac{E_o(s)}{E_i(s)} = \frac{R_2R_4R_1C_1s+1}{R_1R_2R_3C_2+1} = \frac{R_4C_1}{R_3C_2}\frac{s+1/R_1C_1}{s+1/R_2C_2}$$

$$= K_c\alpha\frac{Ts+1}{\alpha Ts+1} = K_c\frac{s+1/T}{s+1/\alpha T} \qquad (2-44)$$

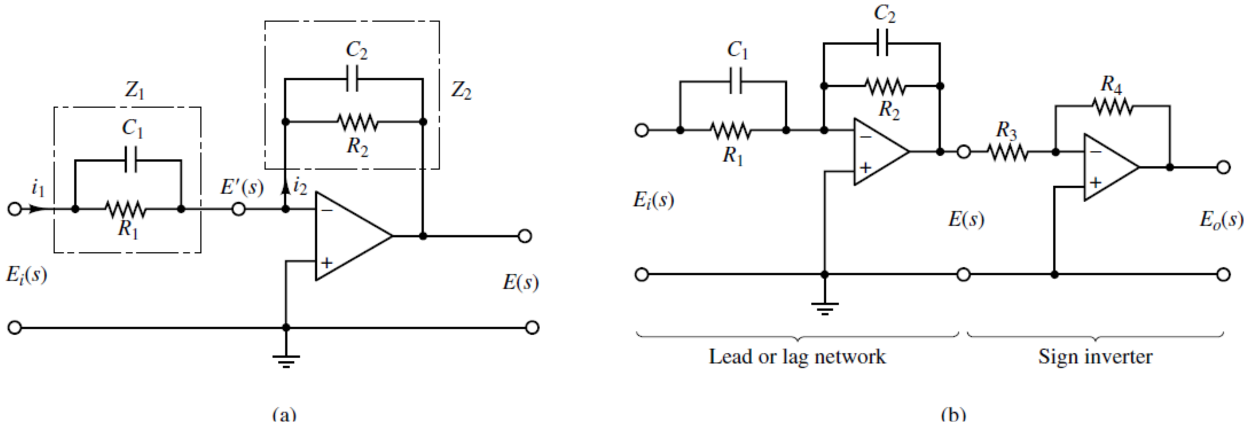


Figure 2-14 (a) Operational-amplifier circuit; (b) operational-amplifier circuit used as a lead or lag compensator.

Where

$$T = R_1C_1, \qquad \alpha T = R_2C_2, \qquad K_c = \frac{R_4C_1}{R_3C_2}$$

Notice that

$$K_c\alpha = \frac{R_4C_1R_2C_2}{R_3C_2R_1C_1} = \frac{R_2R_4}{R_1R_3}, \qquad \alpha = \frac{R_2C_2}{R_1C_1}$$

This network has a dc gain of $K_c\alpha = R_2R_4/(R_1R_3)$.

Note that this network, whose transfer function is given by Equation (2–44), is a lead network if $R_1C_1 > R_2C_2$, or $\alpha < 1$. It is a lag network if $R_1C_1 < R_2C_2$.

PID Controller Using Operational Amplifiers. Figure 2-15 shows an electronic proportional-plus-integral-plus-derivative controller (a PID controller) using operational amplifiers. The transfer function $E(s)/E_i(s)$ is given by

$$\frac{E(s)}{E_i(s)} = -\frac{Z_2}{Z_1}$$

Where

$$Z_1 = \frac{R_1}{R_1C_1+1}, \qquad Z_2 = \frac{R_2C_2s+1}{C_2s}$$

Thus

$$\frac{E(s)}{E_i(s)} = -\left(\frac{R_2C_2s+1}{C_2s}\right)\left(\frac{R_1C_1s+1}{R_1}\right)$$

Noting that

$$\frac{E_o(s)}{E(s)} = -\frac{R_4}{R_3}$$

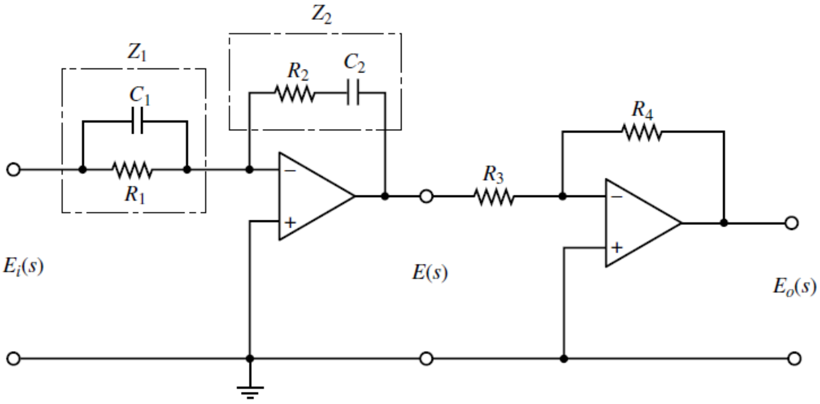


Figure 2-15 Electronic PID controller.

we have

$$\frac{E_o(s)}{E_i(s)} = \frac{E_o(s)}{E(s)}\frac{E(s)}{E_i(s)} = \frac{R_4R_2}{R_3R_1}\frac{(R_1C_1s+1)(R_2C_2s+1)}{R_2C_2s}$$

$$= \frac{R_4R_2}{R_3R_1}\left(\frac{R_1C_1+R_2C_2}{R_2C_2} + \frac{1}{R_2C_2s} + R_1C_1s\right)$$

$$= \frac{R_4(R_1C_1+R_2C_2)}{R_3R_1C_2}\left[1 + \frac{1}{(R_1C_1+R_2C_2)s} + \frac{R_1C_1R_2C_2}{R_1C_1+R_2C_2}s\right] \quad (2-45)$$

Notice that the second operational-amplifier circuit acts as a sign inverter as well as a gain adjuster.

When a PID controller is expressed as

$$\frac{E_o(s)}{E_i(s)} = K_p\left(1 + \frac{T_i}{s} + T_ds\right)$$

K_p is called the proportional gain, T_i is called the integral time, and T_d is called the derivative time. From Equation (2–45) we obtain the proportional gain K_p, integral time T_i, and derivative time T_d to be

$$K_p = \frac{R_4(R_1C_1+R_2C_2)}{R_3R_1C_2}$$

$$T_i = \frac{1}{R_1C_1+R_2C_2}$$

$$T_d = \frac{R_1C_1R_2C_2}{R_1C_1+R_2C_2}$$

When a PID controller is expressed as

$$\frac{E_o(s)}{E_i(s)} = K_p + \frac{K_i}{s} + K_ds$$

K_p is called the proportional gain, K_i is called the integral gain, and K_d is called the derivative gain. For this controller

$$K_p = \frac{R_4(R_1C_1+R_2C_2)}{R_3R_1C_2}$$

$$K_i = \frac{R_4}{R_3 R_1 C_2}$$
$$K_d = \frac{R_4 R_2 C_1}{R_3}$$

2-8 Impulse Response Function

The Transfer Functions (TF) of a Linear Time-Invariant Systems (LTI) system is

$$\text{Transfer Function (TF)} = \mathrm{G(s)} = \frac{\mathrm{Y(s)}}{\mathrm{X(s)}}$$

where $X(s)$ is the Laplace Transform (LT) of the input $x(t)$ and $Y(s)$ is the Laplace Transform (LT) of the output $y\ (t)$ and where we assume all I.C's involved are zero. It follows that

$$Y(s) = G(s)X(s) \qquad (2-46)$$

Now, consider the ***output (response)*** of the system to a ***unit-impulse*** $\delta\ (t)$ ***input*** when all the I.C's are zero. Since

$$L[\delta(t)] = 1$$

the Laplace Transform (LT) of the output of the system is

$$Y(s) = G(s) \qquad (2-47)$$

The inverse Laplace Transform (LT) of the output of the system is given by Equation (2-47) yields the impulse response of the system, i.e;

$$L^{-1}[G(s)] = g(t)$$

is called the ***impulse response function*** or the ***weighting function***, of the system. The impulse-response function $g(t)$ is thus the response of a linear system to a unit impulse input when the I.C's are zero. The LT of $g(t)$ gives the Transfer Function (TF).

Consider that a linear time-invariant system has the input $u(t)$ and output $y(t)$. As shown in Fig. 2-16, a rectangular pulse function $u(t)$ of a very large magnitude $\hat{u}/2\varepsilon$ becomes

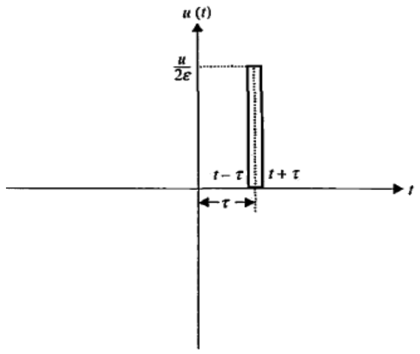


Figure 2-16 Graphical representation an impulse function.

an impulse function for very small durations as $\varepsilon \to 0$. The equation representing Fig. 2-16 is

$$u(t) = \begin{cases} 0 & t \le \tau - \varepsilon \\ \frac{\hat{u}}{2\varepsilon} & \tau - \varepsilon < t < \tau + \varepsilon \\ 0 & t \ge \tau + \varepsilon \end{cases} \qquad (2-48)$$

For $\hat{u} = 1, u(t) = \delta(t)$ is also known as unit impulse or Dirac delta function. For $t = 0$ in Eq. (2-48), using $F(s)$ and noting the actual limits of the integral are defined from $t = 0^- \ to\, t = \infty$, it is easy to verify that the Laplace transform of $\delta(t)$ is unity,.$i.e$ $\mathcal{L}[\delta(t)] = 1$ as $\varepsilon \to 0$.

The important point here is that the response of any system can be characterized by its impulse response $g(t)$, which is defined as the output when the input is a unit-impulse function $\delta(t)$. Once the impulse response of a linear system is known, the output of the system $y(t)$, with any input, $u(t)$, can be found by using the transfer function. We define

$$G(s) = \frac{\mathcal{L}[y(t)]}{\mathcal{L}[u(t)]} = \frac{Y(s)}{F(s)} \qquad (2-49)$$

as the transfer function of the system.

Example 2-7: [59] For the second-order prototype system

$$\frac{d^2y(t)}{dt^2} + 2\zeta\omega_n \frac{dy(t)}{dt} + \omega_n^2 y(t) = \omega_n^2 u(t) \qquad (2-50)$$

Hence,

$$G(s) = \frac{\mathcal{L}[y(t)]}{\mathcal{L}[u(t)]} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s +} \qquad (2-51)$$

is the transfer function of the system in Eq. (1-51). given zero initial conditions, the impulse response $g(t)$ is

$$g(t) = \frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin \omega_n \sqrt{1-\zeta^2 t} \quad t \geq 0 \qquad (2-52)$$

For a unit-step input $u(t) = u_s(t)$, using the convolution properties of Laplace transforms,

$$\mathcal{L}[y(t)] = \mathcal{L}[u_s * g(t)]$$

$$= \mathcal{L}\left[\int_0^t u_s g(t-\tau) d\tau\right] = \frac{G(s)}{s} \qquad (2-53)$$

From the inverse Laplace transform of Eq. (2-53), the output $y(t)$ is therefore

$$\int_0^t u_s g(t-\tau) d\tau$$

or

$$y(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin\left(\omega_n \sqrt{1-\zeta^2 t} + \theta\right) \quad t \geq 0 \qquad (2-54)$$

where, $\theta = \cos^{-1}\zeta$.

Example 2-8: [6] Figure 2–17(a) shows a schematic diagram of an automobile suspension system. As the car moves along the road, the vertical displacements at the tires act as the motion excitation to the automobile suspension system. The motion of this system consists of a translational motion of the center of mass and a rotational motion about the center of mass. Mathematical modeling of the complete system is quite complicated.

A very simplified version of the suspension system is shown in Figure 2–17(b).Assuming that the motion x_i at point P is the input to the system and the vertical motion x_0 of the body is the output, obtain the transfer function $X_0(s)/X_i(s)$.

(Consider the motion of the body only in the vertical direction.) Displacement x_0 is measured from the equilibrium position in the absence of input x_i.

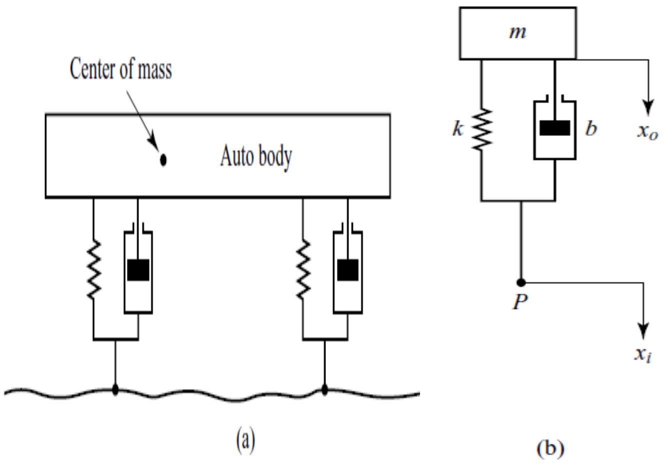


Figure 2-17 (a) Automobile suspension system; (b) simplified suspension system.

Solution. The equation of motion for the system shown in Figure 2–17(b) is

$$m\ddot{x}_o + b(\dot{x}_o - \dot{x}_i) + k(x_o - x_i) = 0$$

or

$$m\ddot{x}_o + b\dot{x}_o + kx_o = b\dot{x}_i + kx_i$$

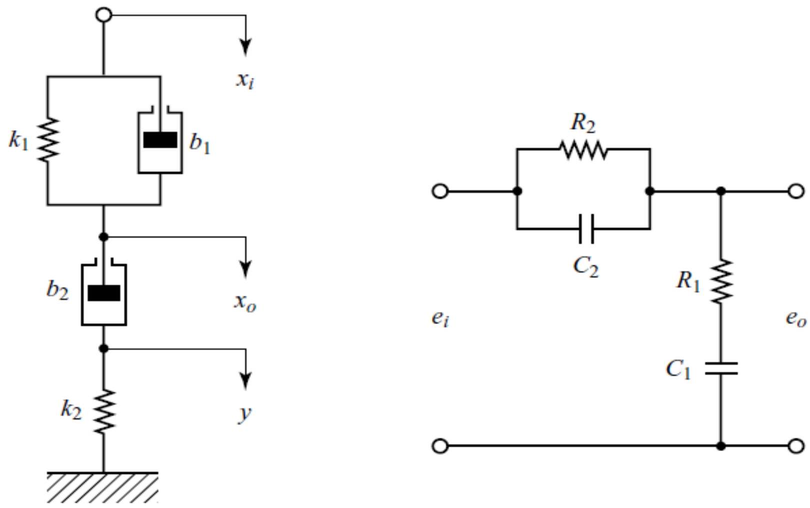
Taking the Laplace transform of this last equation, assuming zero initial conditions, we obtain

$$(ms^2 + bs + k)X_o(s) = (bs + k)X_i(s)$$

Hence the transfer function $X_o(s)/X_i(s)$ is given by

$$\frac{X_o(s)}{X_i(s)} = \frac{bs + k}{ms^2 + bs + k}$$

Example 2-9: [6] We obtain the transfer function $X_0(s)/X_i(s)$ of the mechanical system shown in Figure 2–18(a). Also obtain the transfer function $E_0(s)/E_i(s)$ of the electrical system shown in Figure 2–18(b). Show that these transfer functions of the two systems are of identical form and thus they are analogous systems



(a) (b)

Figure 2-18 (a) Mechanical system; (b) analogous electrical system.

Solution. In Figure 2-18(a) we assume that displacements x_i, x_0, and y are measured from their respective steady-state positions. Then the equations of motion for the mechanical system shown in Figure 2-18(a) are

$$b_1(\dot{x}_i - \dot{x}_0) + k_1(x_i - x_0) = b_2(\dot{x}_0 - \dot{y})$$
$$b_2(\dot{x}_0 - \dot{y}) = k_2 y$$

By taking the Laplace transforms of these two equations, assuming zero initial conditions, we have

$$b_1[sX_i(s) - sX_0(s)] + k_1[X_i(s) - X_0(s)] = b_2[sX_0(s) - sY(s)]$$
$$b_2[sX_0(s) - sY(s)] = k_2 Y(s)$$

If we eliminate *Y(s)* from the last two equations, then we obtain

$$b_1[sX_i(s) - sX_0(s)] + k_1[X_i(s) - X_0(s)] = b_2 sX_0(s) - b_2 s\frac{b_2 sX_0(s)}{b_2 s + k_2}$$

or

$$(b_1 s + k_1)X_i(s) = \left(b_1 s + k_1 + b_2 s - b_2 s\frac{b_2 s}{b_2 s + k_2}\right)X_0(s)$$

Hence the transfer function $X_0(s)/X_i(s)$ can be obtained as

$$\frac{X_0(s)}{X_i(s)} = \frac{\left(\frac{b_1}{k_1}s + 1\right)\left(\frac{b_2}{k_2}s + 1\right)}{\left(\frac{b_1}{k_1}s + 1\right)\left(\frac{b_2}{k_2}s + 1\right) + \frac{b_2}{k_1}s}$$

For the electrical system shown in Figure 2–18(b), the transfer function $E_0(s)/E_i(s)$ is found to be

$$\frac{E_0(s)}{E_i(s)} = \frac{R_1 + \frac{1}{C_1 s}}{\frac{1}{(1/R_2) + C_2 s} + R_1 + \frac{1}{C_1 s}}$$
$$= \frac{(R_1C_1 s + 1)(R_2C_2 s + 1)}{(R_1C_1 s + 1)(R_2C_2 s + 1) + R_2C_1}$$

A comparison of the transfer functions shows that the systems shown in Figures 2-18(a) and (b) are analogous.

Example 2-10: [6] Obtain the transfer function $E_o(s)/E_i(s)$of the op-amp circuit shown in Figure 2–19.

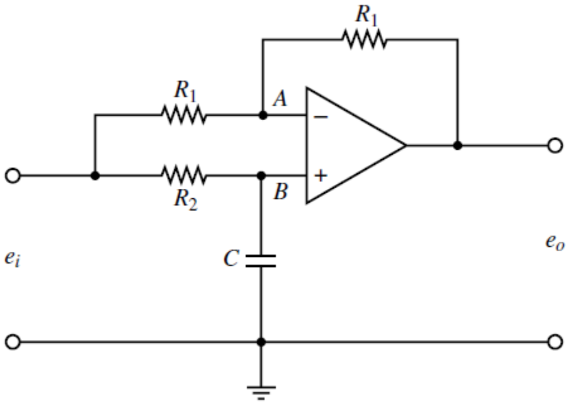


Figure 2-19 Operational amplifier circuit.

Solution. The voltage at point A is

$$e_A = \frac{1}{2}(e_i - e_o) + e_o$$

The Laplace-transformed version of this last equation is

$$E_A(s) = \frac{1}{2}[E_i(s) + E_o(s)]$$

The voltage at point B is

$$E_B(s) = \frac{\frac{1}{Cs}}{R_2 + \frac{1}{Cs}} E_i(s) = \frac{1}{R_2Cs + 1} E_i(s)$$

Since $[E_B(s) - E_A(s)]\, K = E_o(s)$ and $K \gg 1$, we must have $E_A(s) = E_B(s)$.
Thus

$$\frac{1}{2}[E_i(s) + E_o(s)] = \frac{1}{R_2Cs + 1} E_i(s)$$

Hence

$$\frac{E_o(s)}{E_i(s)} = -\frac{R_2Cs - 1}{R_2Cs + 1} = -\frac{s - \frac{1}{R_2C}}{s + \frac{1}{R_2C}}$$

Example 2-11: [6] Obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp system shown in Figure 2-20 in terms of complex impedances Z_1 ,Z_2 ,Z_3 , and Z_4 . Using the equation derived, obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp system shown in Figure 2–20.

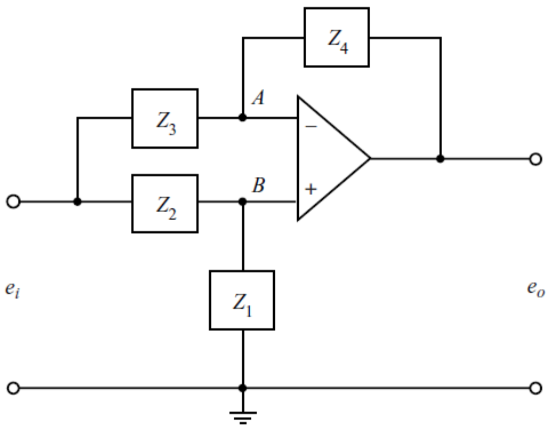


Figure 2–20 Operational amplifier circuit.

Solution. From Figure 2-20, we find

$$\frac{E_i(s) - E_A(s)}{Z_3} = \frac{E_A(s) - E_o(s)}{Z_4} \qquad (2-55)$$

Or

$$E_i(s) - \left(1 + \frac{Z_3}{Z_4}\right) E_A(s) = -\frac{Z_3}{Z_4} E_o(s) \qquad (2-56)$$

Since

$$E_A(s) = E_B(s) = \frac{Z_1}{Z_1 + Z_2} E_i(s) \qquad (2-57)$$

by substituting Equation (2–56) into Equation (2–57), we obtain

$$\left[\frac{Z_4 Z_1 + Z_4 Z_2 - Z_3 Z_1}{Z_4(Z_1 + Z_2)}\right] E_i(s) = -\frac{Z_3}{Z_4} E_o(s)$$

from which we get the transfer function $E_o(s)/E_i(s)$ to be

$$\frac{E_o(s)}{E_i(s)} = -\frac{Z_4 Z_2 - Z_3 Z_1}{Z_3(Z_1 + Z_2)} \qquad (2-58)$$

To find the transfer function $E_o(s)/E_i(s)$ of the circuit shown in Figure 2–20, we substitute

$$Z_1 = \frac{1}{Cs}, \qquad Z_2 = R_2, \qquad Z_3 = R_1, \;\; Z_4 = R_1$$

into Equation (2–58).The result is

$$\frac{E_o(s)}{E_i(s)} = -\frac{R_1 R_2 - R_1 \frac{1}{Cs}}{R_1\left(\frac{1}{Cs} + R_2\right)} = -\frac{R_2 Cs - 1}{R_2 Cs + 1}$$

CHAPTER THREE

Automatic Control Systems

A control system may consist of a number of components. To show the functions performed by each component, in control engineering, we commonly use a diagram called the block diagram. This chapter first explains what a block diagram is. Next, it discusses introductory aspects of automatic control systems, including various control actions. Then, it presents a method for obtaining block diagrams for physical systems, and, finally, discusses techniques to simplify such diagrams.

We also utilize the block diagram reduction techniques and the Mason's gain formula to find the transfer function of the overall control system. The main objectives of this chapter are:

1. To study block diagrams, their components, and their underlying mathematics.
2. To obtain transfer function of systems through block diagram manipulation and reduction.
3. To introduce the signal-flow graphs.
4. To establish a parallel between block diagrams and signal-flow graphs.
5. To use Mason's gain formula for finding transfer function of systems.
6. To introduce state diagrams.

3-1 Block Diagrams

The block diagram modeling may provide control engineers with a better understanding of the composition and interconnection of the components of a system. Or it can be used, together with transfer functions, to describe the cause-and-effect relationships throughout the system. For example, consider a simplified block diagram representation of the heating system in your lecture room, shown in Fig. 3-1, where by setting a desired temperature, also defined as the input, one can set off the furnace to provide heat to the room. The process is relatively straightforward. The actual room temperature is also known as the output and is measured by a sensor within the thermostat. A simple electronic circuit within the thermostat compares the actual room temperature to the desired room temperature (comparator). If the room temperature is below the desired temperature, an error voltage will be generated. The error voltage acts as a switch to open the gas valve and turn on the furnace (or the actuator). Opening the windows and the door in the classroom would cause heat loss and, naturally, would disturb the heating process (disturbance).

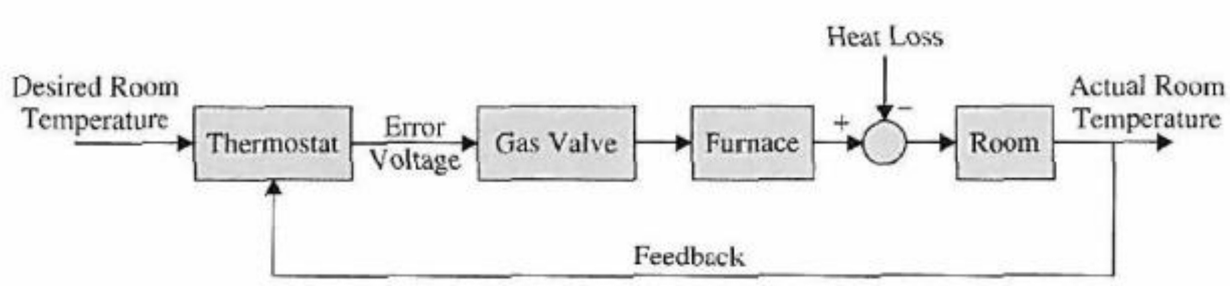
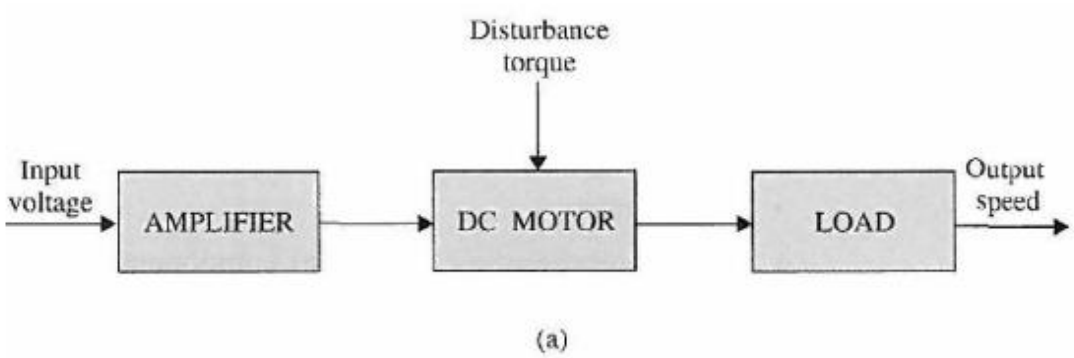


Figure 3-1 A simplified block diagram representation of a heating system.



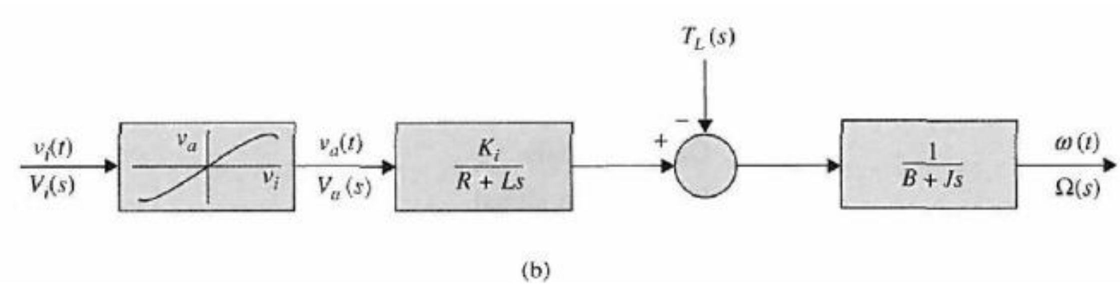


Figure 3-2 (a) Block diagram of a dc-motor control system, (b) Block diagram with transfer functions and amplifier characteristics.

The room temperature is constantly monitored by the output sensor. The process of sensing the output and comparing it with the input to establish an error signal is known as feedback. Note that the error voltage here causes the furnace to turn on, and the furnace would finally shut off when the error reaches zero.

As another example, consider the block diagram of Fig. 3-2 (a), which models an open loop, dc-motor, speed-control system. The block diagram in this case simply shows how the system components are interconnected, and no mathematical details are given. If the mathematical and functional relationships of all the system elements are known, the block diagram can be used as a tool for the analytic or computer solution of the system. In general, block diagrams can be used to model linear as well as nonlinear systems. For example, the input-output relations of the dc-motor control system may be represented by the block diagram shown in Fig. 3-2 (b). In this figure, the input voltage to the motor is the output of the power amplifier, which, realistically, has a nonlinear characteristic. If the motor is linear, or, more appropriately, if it is operated in the linear region of its characteristics, its dynamics can be represented by transfer functions. The nonlinear amplifier gain can only be described in time domain and between the time variables $v_i(t)$ and $v_a(t)$, Laplace transform variables do not apply to nonlinear systems; hence, in this case, $V_i(s)$ and $V_a(s)$ do not exist. However, if the magnitude of $v_i(t)$ is limited to the linear range of the amplifier, then the amplifier can be regarded as linear, and the amplifier may be described by the transfer function

$$\frac{V_a(s)}{V_i(s)} = K \qquad (3-1)$$

where K is a constant, which is the slope of the linear region of the amplifier characteristics.

Alternatively, we can use signal-flow graphs or state diagrams to provide a graphical representation of a control system.

Typical Elements of Block Diagrams in Control Systems. We shall now define the block-diagram elements used frequently in control systems and the related algebra. The common elements in block diagrams of most control systems include:

a) Comparators
b) Blocks representing individual component transfer functions, including:
c) Reference sensor (or input sensor)
d) Output sensor
e) Actuator
f) Controller
g) Plant (the component whose variables are to be controlled)

h) Input or reference signals
i) Output signals
j) Disturbance signal
k) Feedback loops

Fig. 3-3 shows one configuration where these elements are interconnected. You may wish to compare Fig. 3-1 or Fig. 3-2 to Fig. 3-3 to find the control terminology for each system. As a rule, each block represents an element in the control system, and each element can be modeled by one or more equations. These equations are normally in the time domain or preferably (because of ease in manipulation) in the Laplace domain. Once the block diagram of a system is fully constructed, one can study individual components or the overall system behavior.

One of the important components of a control system is the sensing and the electronic device that acts as a junction point for signal comparisons—otherwise known as a comparator. In general, these devices possess sensors and perform simple mathematical operations such as addition and subtraction (such as the thermostat in Fig. 3-1). Three examples of comparators are illustrated in Fig. 3-4. Note that the addition and subtraction operations in Fig. 3-4 (a) and (b) are linear, so the input and output variables of these block diagram elements can be time-domain variables or Laplace-transform variables. Thus, in Fig. 3-4 (a), the block diagram implies

$$e(t) = r(t) - y(t) \qquad (3-2)$$

or

$$E(s) = R(s) - Y(s) \qquad (3-3)$$

As mentioned earlier, **blocks** represent the equations of the system in time domain or the **transfer function** of the system in the Laplace domain, as demonstrated in Fig. 3-5.

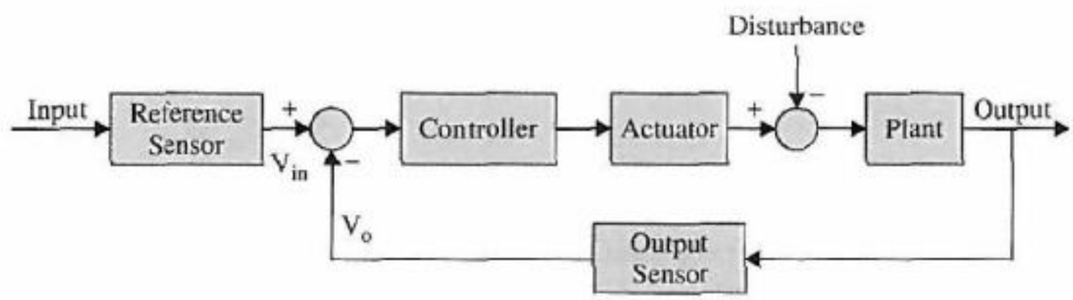


Figure 3-3 Block diagram representation of a general control system.

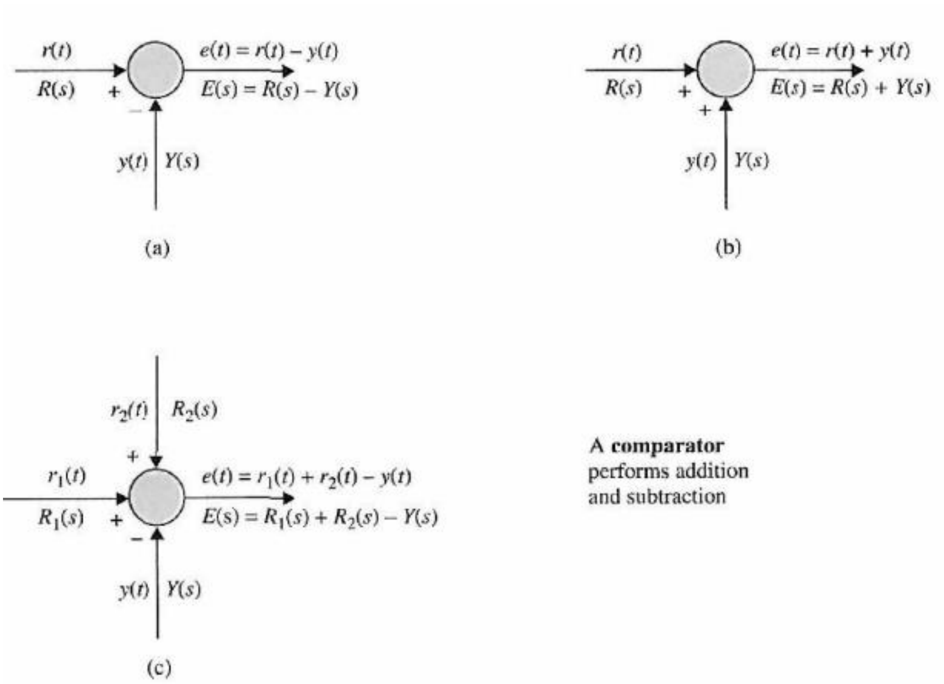


Figure 3-4 Block-diagram elements of typical sensing devices of control systems, (a) Subtraction, (b) Addition, (c) Addition and subtraction.

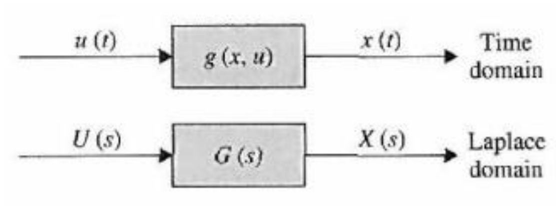


Figure 3-5 Time and Laplace domain block diagrams.

In Laplace domain, the following input-output relationship can be written for the system in Fig. 3-5:

$$X(s) = G(s)U(s) \qquad (3-4)$$

If signal *X(s)* is the output and signal *U(s)* denotes the input, the transfer function of the block in Fig. 3-5 is

$$G(s) = \frac{X(s)}{U(s)} \qquad (3-5)$$

Typical block elements that appear in the block diagram representation of most control systems include **plant, controller, actuator,** and **sensor.**

Example 3-1: [59] Consider the block diagram of two transfer functions $G_1(s)$ and $G_2(s)$ that are connected in series. Find the transfer function $G(s)$ of the overall system.

Solution. The system transfer function can be obtained by combining individual block equations. Hence, for signals *A(s)* and *X(s)*, we have

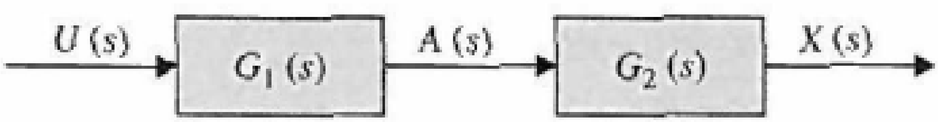


Figure 3-6 Block diagrams $G_1(s)$ and $G_2(s)$ connected in series.

$$X(s) = A(s)G_2(s)$$
$$A(s) = U(s)G_1(s)$$
$$X(s) = G_1(s)G_2(s)U(s)$$

$$G(s) = \frac{X(s)}{U(s)}$$

Hence,

$$G(s) = G_1(s)G_2(s) \qquad (3-6)$$

Hence, using Eq. (3-6), the system in Fig. 3-6 can be represented by the system in Fig. 3-5.

Example 3-2: [59] Consider a more complicated system of two transfer functions $G_1(s)$ and $G_2(s)$ that are connected in parallel, as shown in Fig. 3-7. Find the transfer function $G(s)$ of the overall system.

Solution. The system transfer function can be obtained by combining individual block equations. Note for the two blocks $G_1(s)$ and $G_2(s)$, $A_1(s)$ acts as the input, and $A_2(s)$ and $A_3(s)$ are the outputs, respectively. Further, note that signal *U(s)* goes through a **branch point *P*** to become $A_1(s)$. Hence, for the overall system, we combine the equations as follows.

$$A_1(s) = U(s)$$
$$A_2(s) = A_1(s)G_1(s)$$
$$A_3(s) = A_1(s)G_2(s)$$
$$X(s) = A_2(s) + A_3(s)$$
$$X(s) = U(s)\big(G_1(s) + G_2(s)\big)$$
$$G(s) = \frac{X(s)}{U(s)}$$

Hence,

$$G(s) = G_1(s) + G_2(s) \qquad (3-7)$$

For a system to be classified as a **feedback control** system, it is necessary that the controlled variable be fed back and **compared** with the reference input. After the comparison, an error signal is generated, which is used to **actuate** the control system. As a result, the actuator is activated in the presence of the error to minimize or eliminate that very error. A necessary component of every feedback control system is an **output sensor**, which is used to convert the output signal to a quantity that has the same units as the reference input. A feedback control system is also known a closed-loop system. A system may have multiple feedback loops. Fig. 3-8 shows the block diagram of a linear feedback control system with a single feedback loop. The following terminology is defined with reference to the diagram:

$r(t), R(s) =$ reference input(command)

$y(t), Y(s) =$ output (controlled variable)

$b(t), B(s) =$ feedback signal

$u(t), U(s) =$ actuating signal $=$ error sign $e(t), E(s),$ when $H(s) = 1$

$\mathrm{H(s)} =$ feedback transfer function

$G(s)H(s) = L(s) =$ loop transfer function

$G(s) =$ forward $-$ path transfer function

$$M(s) = \frac{Y(s)}{R(s)}$$

$=$ closed $-$ loop transfer function or system transfer function

The closed-loop transfer function $M(s)$ can be expressed as a function of $G(s)$ and $H(s)$. From Fig. 3-8, we write

$$Y(s) = G(s)U(s) \quad (3-8)$$

and

$$B(s) = H(s)Y(s) \quad (3-9)$$

The actuating signal is written

$$U(s) = R(s) - B(s) \quad (3-10)$$

Substituting Eq. (3-10) into Eq. (3-8) yields

$$Y(s) = G(s)R(s) - G(s)H(s) \quad (3-11)$$

Substituting Eq. (3-9) into Eq. (3-7) and then solving for *Y(s)/R(s)* gives the closed-loop transfer function

$$M(s) = \frac{Y(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \quad (3-12)$$

The feedback system in Fig. 3-8 is said to have a **negative feedback loop** because the comparator **subtracts.** When the comparator **adds** the feedback, it is called **positive feedback,** and the transfer function Eq. (3-12) becomes

$$M(s) = \frac{Y(s)}{R(s)} = \frac{G(s)}{1 - G(s)H(s)} \quad (3-13)$$

If *G* and *H* are constants, they are also called **gains.** If *H* = **1** in Fig. 3-8, the system is said to have a **unity feedback loop,** and if *H = 0*, the system is said to be **open loop.**

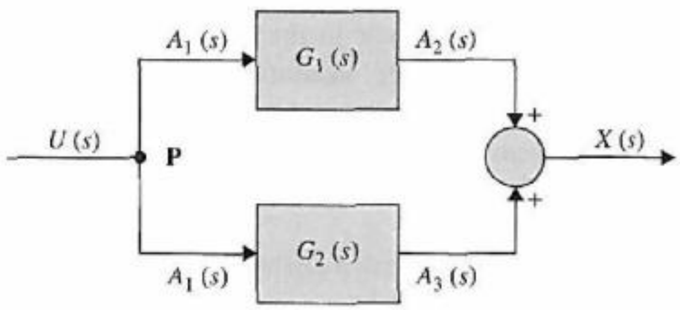


Figure 3-7 Block diagrams $G_1(s)$ and $G_2(s)$ connected in parallel.

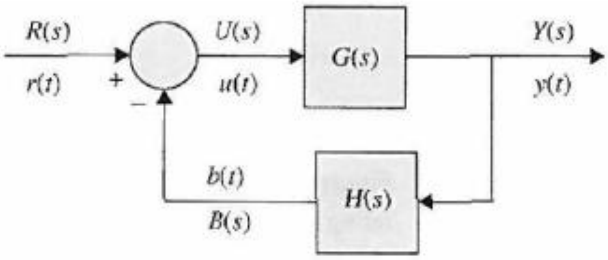


Figure 3-8 Basic block diagram of a feedback control system.

3-1-1 Relation between Mathematical Equations and Block Diagrams

Consider the following second-order prototype system:

$$\ddot{x}(t) + 2\zeta\omega_n\dot{x}(t) + \omega_n^2 x(t) = \omega_n^2 u(t) \quad (3-14)$$

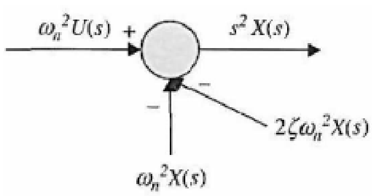


Figure 3-9 Graphical representation of Eq. (3-17) using a comparator.

which has Laplace representation (assuming zero initial conditions $x(0) = \dot{x}(0) = 0$):

$$X(s)s^2 + 2\zeta\omega_n X(s)s + \omega_n^2 X(s) = \omega_n^2 U(s) \quad (3-15)$$

Eq. (3-15) consists of constant damping ratio ζ, constant natural frequency ω_n, input *U(s)*, and output *X(s)*. If we rearrange Eq. (3-15) to

$$\omega_n^2 U(s) - 2\zeta\omega_n X(s)s - \omega_n^2 X(s) = X(s)s^2 \qquad (3-16)$$

it can graphically be shown as in Fig. 3-9.

The signals $2\zeta\omega_n sX(s)$ and $\omega_n^2 X(s)$ may be conceived as the signal *X(s)* going into blocks with transfer functions $2\zeta\omega_n s$ and ω_n^2, respectively, and the signal *X(s)* may be obtained by integrating $s^2X(s)$ twice or by post-multiplying by $\frac{1}{s^2}$, as shown in Fig. 3-10.

Because the signals *X(s)* in the right-hand side of Fig. 3-10 are the same, they can be connected, leading to the block diagram representation of the system Eq. (3-17), as shown in Fig. 3-11. If you wish, you can further dissect the block diagram in Fig. 3-11 by factoring out the term - as in Fig. 3-12(a) to obtain Fig. 3-12(b).

If the system studied here corresponds to the spring-mass-damper seen in Fig. 1-7 (see Chapter 1), then we can designate internal variables *A(s)* and *V(s)*, which represent acceleration and velocity of the system, respectively, as illustrated in Fig. 3-12. The best way to see this is by recalling that $\frac{1}{s}$ is the integration operation in Laplace domain. Hence, if *A(s)* is integrated once, we get *V(s)*, and after integrating *V(s)*, we get the *X(s)* signal.

It is evident that there is no unique way of representing a system model with block diagrams. We may use different block diagram forms for different purposes, as long as the

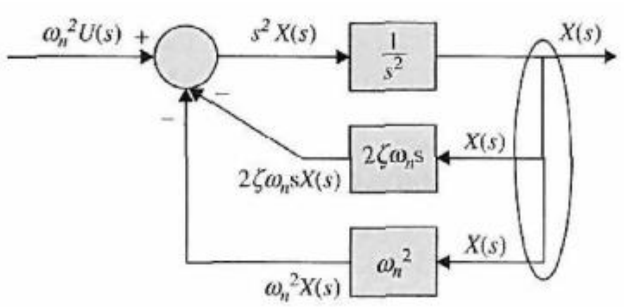


Figure 3-10 Addition of blocks $\frac{1}{s^2}$, $2\zeta\omega_n s$, and ω_n^2 to the graphical representation of Eq. (3-17).

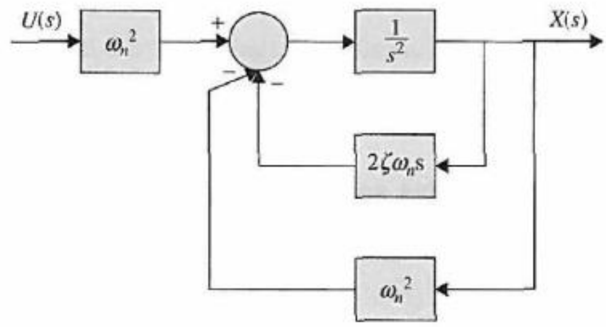
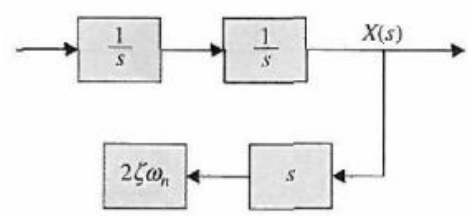


Figure 3-11 Block diagram representation of Eq. (3-17) in Laplace domain.



(a)

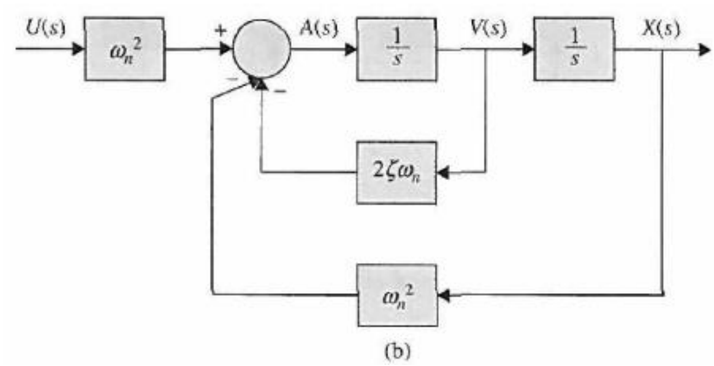


Figure 3-12 (a) Factorization of | term in the internal feedback loop of Fig. 3-11. (b) Final block diagram representation of Eq. (3-17) in Laplace domain.

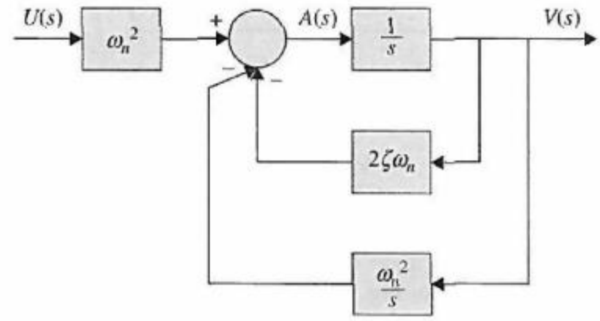


Figure 3-13 Block diagram of Eq. (3-17) in Laplace domain with *V(s)* represented as the output.

overall transfer function of the system is not altered. For example, to obtain the transfer function *V(s)/U(s),* we may yet rearrange Fig. 3-12 to get *V(s)* as the system output, as shown in Fig. 3-13. This enables us to determine the behavior of velocity signal with input *U(s).*

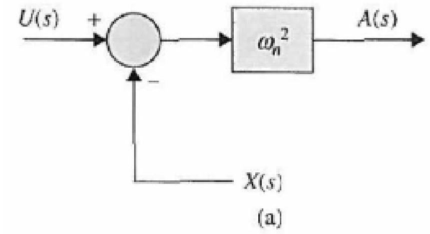
Example 3-3: [59] Find the transfer function of the system in Fig. 3-12 and compare that to the transfer function of system in Eq. (3-15).

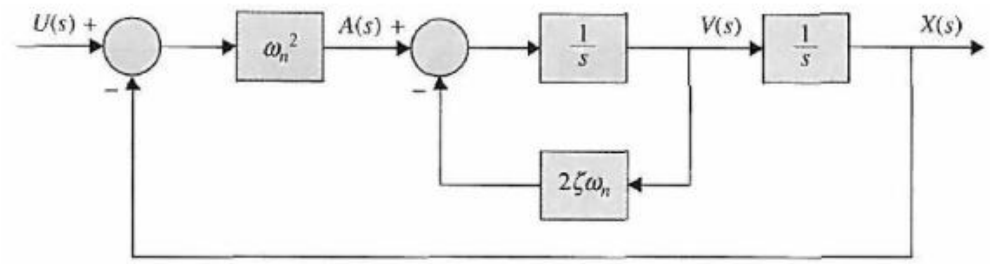
Solutions

The ω_n^2, block at the input and feedback signals in Fig. 3-12(b) may be moved to the right-hand side of the comparator. This is the same as factorization of ω_n^2 as shown:

$$\omega_n^2 U(s) - \omega_n^2 X(s) = \omega_n^2\big(U(s) - X(s)\big) \qquad (3-17)$$

Fig. 3-14(a) shows the factorization operation of Eq. (3-17), which results in a simpler block diagram representation of the system shown in Fig. 3-14 (b). Note that Fig. 3-12(b) and Fig. 3-14(b) are equivalent systems.





(b)

Figure 3-14 (a) Factorization of ω_n^2. (b) Alternative block diagram representation of Eq. (3-17) in Laplace domain.

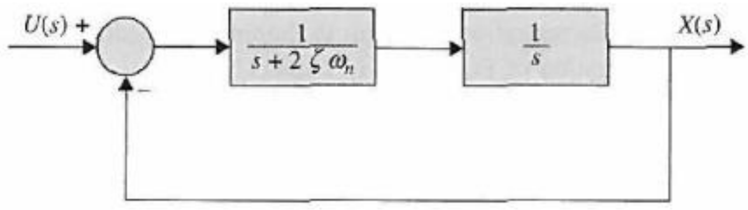


Figure 3-15 A block diagram representation of $\frac{\omega_n^2}{s^2+2\zeta\omega_n+\omega_n^2}$

Considering Fig. 3-12(b), it is easy to identify the internal feedback loop, which in turn can be simplified using Eq. (3-12), or

$$\frac{V(s)}{A(s)} = -\frac{\frac{1}{s}}{1+\frac{2\zeta\omega_n}{s}} = \frac{s}{s+2\zeta\omega_n} \qquad (3-18)$$

After pre- and post-multiplication by ω_n^2 and $\frac{1}{s}$, respectively, the block diagram of the system is simplified to what is shown in Fig. 3-15, which ultimately results in

$$\frac{X(s)}{U(s)} = \frac{\frac{\omega_n^2}{s(s+2\zeta\omega_n)}}{\frac{\omega_n^2}{1+s(s+2\zeta\omega_n)}} = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \qquad (3-19)$$

Eq. (3-19) is also the transfer function of system Eq. (3-15).

Example 3-4: Find the velocity transfer function using Fig. 3-13 and compare that to the derivative of Eq. (3-19).

Solutions. Simplifying the two feedback loops in Fig. 3-13, starting with the internal loop first, we have

$$\frac{V(s)}{U(s)} = \frac{\frac{\frac{1}{s}}{1+\frac{2\zeta\omega_n}{s}}}{\frac{\frac{1}{s}}{1+\frac{2\zeta\omega_n}{s}} - \frac{\omega_n^2}{s}}$$

$$\frac{V(s)}{U(s)} = \frac{s\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \qquad (3-20)$$

Eq. (3-20) is the same as the derivative of Eq. (3-19), which is nothing but multiplying Eq. (3-19) by an *s* term. Try to find the *A(s)/U(s)* transfer function. Obviously you must get: $s^2X(s)/U(s)$.

3-1-2 Block Diagram Reduction

As you might have noticed from the examples in the previous section, the transfer function of a control system may be obtained by manipulation of its block diagram and by its ultimate reduction into one block. For complicated block diagrams, it is often necessary to move a **comparator** or a **branch point** to make the block diagram reduction process simpler. The two key operations in this case are:

1. Moving a branch point from P to Q, as shown in Fig. 3-16(a) and Fig. 3-16(b). This operation must be done such that the signals *Y(s)* and *B(s)* are unaltered. In Fig. 3-16(a), we have the following relations:

$$Y(s) = A(s)G_2(s)$$
$$B(s) = Y(s)H_1(s) \qquad (3-21)$$

In Fig. 3-16(b), we have the following relations:

$$Y(s) = A(s)G_2(s)$$
$$B(s) = A(s)H_1(s)G_2(s) \qquad (3-22)$$

But

$$G_2(s) = \frac{Y(s)}{A(s)} \qquad (3-23)$$
$$\Rightarrow B(s) = Y(s)H_1(s)$$

2. Moving a comparator, as shown in Fig. 3-17(a) and Fig. 3-17(b), should also be done such that the output *Y(s)* is unaltered. In Fig. 3-17(a), we have the following relations:

$$Y(s) = A(s)G_2(s) + B(s)H_1(s) \qquad (3-24)$$

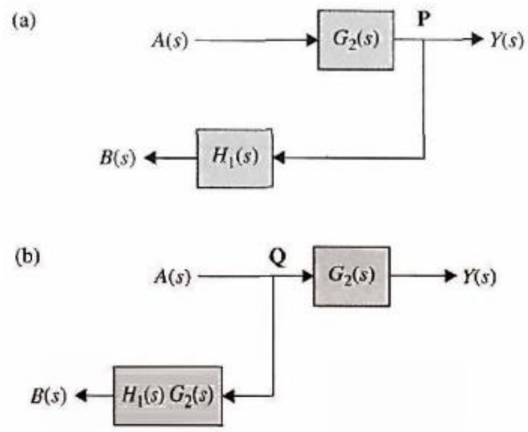


Figure 3-16 (a) Branch point relocation from point P to (b) point Q.

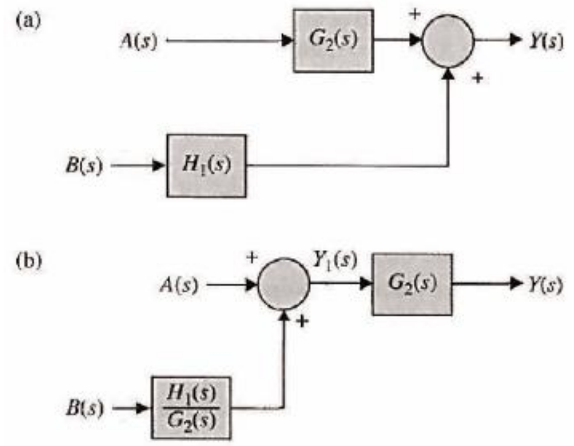


Figure 3-17 (a) Comparator relocation from the right-hand side of block $G_2(s)$ to (b) the left-hand side of block $G_2(s)$.

In Fig. 3-17(b), we have the following relations:

$$Y_1(s) = A(s) + B(s)\frac{H_1(s)}{G_2(s)} \qquad (3-25)$$
$$Y(s) = Y_1(s)G_2(s)$$

so

$$Y(s) = A(s)G_2(s) + B(s)\frac{H_1(s)}{G_2(s)}G_2(s) \qquad (3-26)$$

$$\Rightarrow Y(s) = A(s)G_2(s) + B(s)H_1(s)$$

Example 3-5: [59] Find the input-output transfer function of the system shown in Fig. 3-17(a).

Solution. To perform the block diagram reduction, one approach is to move the branch point at Y_1 to the left of block G_2, as shown in Fig. 3-18(b). After that, the reduction becomes trivial, first by combining the blocks G_2, G_3, and G_4 as shown in Fig. 3-18(c), and then by eliminating the two

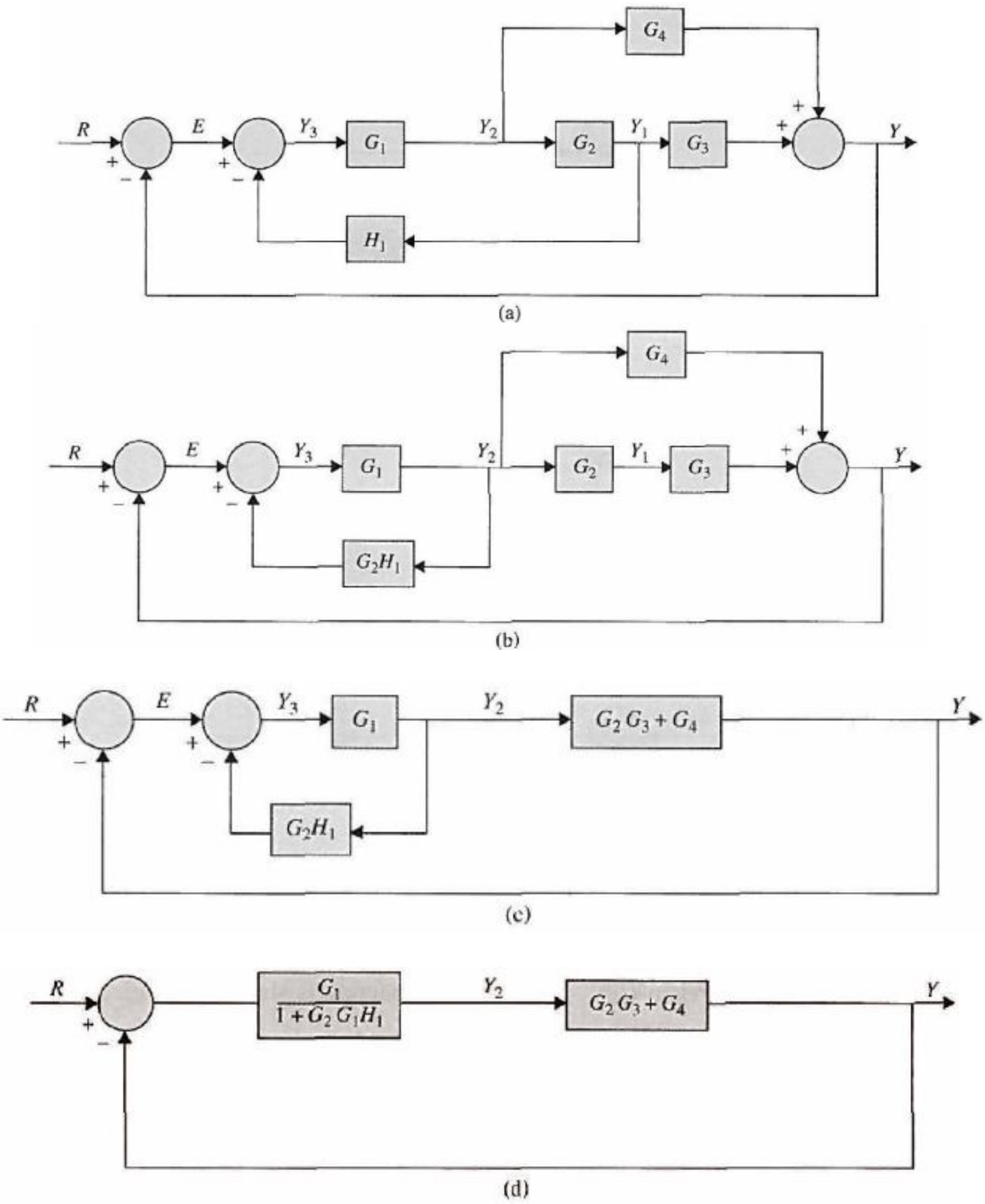


Figure 3-18 (a) Original block diagram, (b) Moving the branch point at Y_1 to the left of block G_2. (c) Combining the blocks G_1, G_2, and G_3. (d) Eliminating the inner feedback loop.

feedback loops. As a result, the transfer function of the final system after the reduction in Fig. 3-18(d) becomes

$$\frac{Y(s)}{E(s)} = \frac{G_1G_2G_3 + G_1G_4}{1 + G_2G_3H_1 + G_1G_2G_3 + G_1G_4} \qquad (3-27)$$

3-1-3 Block Diagram of Multi-Input Systems-Special Case:

Systems with a Disturbance. An important case in the study of control systems is when a disturbance signal is present. Disturbance (such as heat loss in the example in Fig. 3-1) usually adversely affects the performance of the control system by placing a burden on the controller/actuator components. A simple block diagram with two inputs is shown in Fig. 3-19. In this case, one of the inputs, *D(s),* is known as disturbance, while *R(s)* is the reference input. Before designing a proper controller for the system, it is always important to learn the effects of *D(s)* on the system.

We use the method of superposition in modeling a multi-input system.

Super Position: For linear systems, the overall response of the system under multi-inputs is the summation of the responses due to the individual inputs, i.e., in this case,

$$Y_{tatal} = Y_R|_{D=0} + Y_D|_{R=0} \qquad (3-28)$$

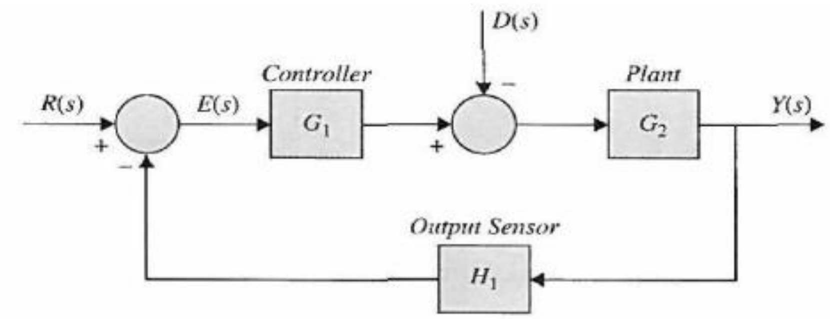


Figure 3-19 Block diagram of a system undergoing disturbance

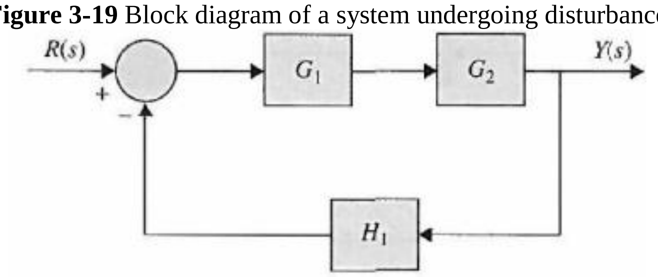


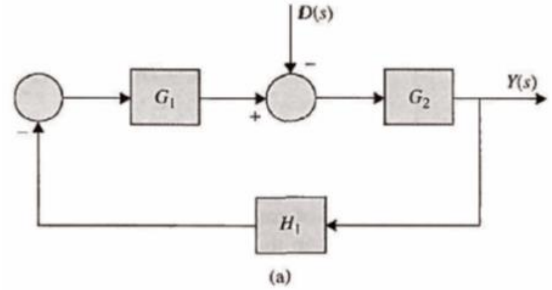
Figure 3-20 Block diagram of the system in Fig. 3-19 when *D(s)* = 0.

When *D(s)* = 0, the block diagram is simplified (Fig. 3-20) to give the transfer function

$$\frac{Y(s)}{R(s)} = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H_1(s)} \qquad (3-29)$$

When *R(s)*=0, the block diagram is rearranged to give (Fig. 3-21):

$$\frac{Y(s)}{D(s)} = \frac{-G_2(s)}{1 + G_1(s)G_2(s)H_1(s)} \qquad (3-30)$$



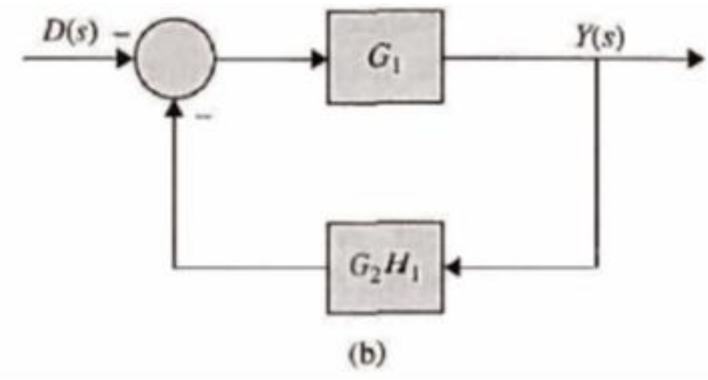


Figure 3-21 Block diagram of the system in Fig. 3-19 when R(s) = 0.

As a result, from Eq. (3-28) to Eq. (3-32), we ultimately get

$$Y_{tatal} = \left.\frac{Y(s)}{R(s)}\right|_{D=0} R(s) + \left.\frac{Y(s)}{D(s)}\right|_{R=0} D(s) \qquad (3-31)$$

$$Y(s) = \frac{G_1G_2}{1+G_1G_2H_1}R(s) + \frac{-G_2}{1+G_1G_2H_1}D(s)$$

Observations: $\left.\frac{Y}{R}\right|_{D=0}$ and $\left.\frac{Y}{D}\right|_{R=0}$ have the same denominators if the disturbance signal goes to the forward path. The negative sign in the numerator of $\left.\frac{Y}{D}\right|_{R=0}$ shows that the disturbance signal interferes with the controller signal, and, as a result, it adversely affects the performance of the system. Naturally, to compensate, there will be a higher burden on the controller.

3-1-4 Block Diagrams and Transfer Functions of Multivariable Systems

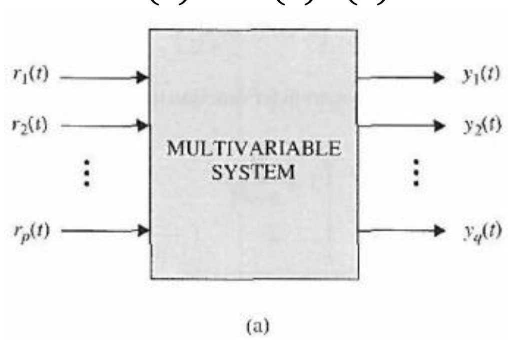
We shall illustrate the block diagram and matrix representations of multivariable systems. Two block-diagram representations of a multivariable system with p inputs and q outputs are shown in Fig. 3-22(a) and (b). In Fig. 3-22 (a), the individual input and output signals are designated, whereas in the block diagram of Fig. 3-22(b), the multiplicity of the inputs and outputs is denoted by vectors. The case of Fig. 3-22(b) is preferable in practice because of its simplicity.

Fig. 3-23 shows the block diagram of a multivariable feedback control system. The transfer function relationships of the system are expressed in vector-matrix form:

$$Y(s) = G(s)U(s) \qquad (3-32)$$

$$U(s) = R(s) - B(s) \qquad (3-33)$$

$$B(s) = H(s)Y(s) \qquad (3-34)$$



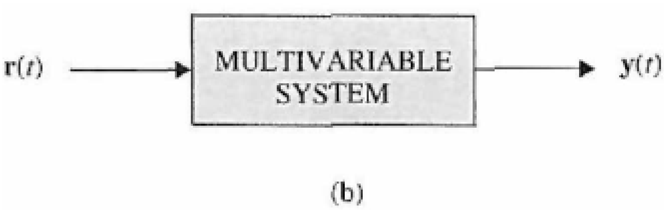


Figure 3-22 Block diagram representations of a multivariable system.

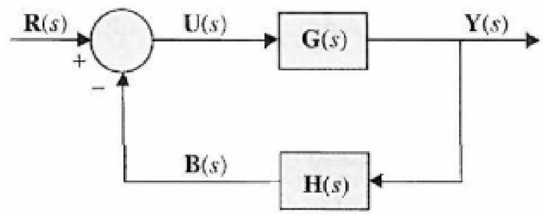


Figure 3-23 Block diagram of a multivariable feedback control system.

where Y(s) is the $q \times 1$ output vector; U(s), R(s), and B(s) are all $p \times 1$ vectors; and G(s) and H(s) are $q \times p$ and p x q transfer-function matrices, respectively. Substituting Eq. (3-11) into Eq. (3-10) and then from Eq. (3-10) to Eq. (3-9), we get

$$Y(s) = G(s)R(s) - G(s)H(s)Y(s) \qquad (3-35)$$

Solving for *Y(s)* from Eq. (3-12) gives

$$Y(s) = [1 + G(s)H(s)]^{-1}G(s)R(s) \qquad (3-36)$$

provided that I + G(s)H(s) is nonsingular. The closed-loop transfer matrix is defined as

$$M(s) = [I + G(s)H(s)]^{-1}G(s) \qquad (3-37)$$

Then Eq. (3-14) is written

$$Y(s) = M(s)R(s) \qquad (3-38)$$

Example 3-6: [59] Consider that the forward-path transfer function matrix and the feedback-path transfer function matrix of the system shown in Fig. 3-23 are

$$G(s) = \begin{bmatrix} \frac{1}{s+1} & -\frac{1}{s} \\ 2 & \frac{1}{s+2} \end{bmatrix}, \qquad H(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (3-39)$$

respectively. The closed-loop transfer function matrix of the system is given by Eq. (3-15), and is evaluated as follows:

$$I + G(s)H(s) = \begin{bmatrix} 1 + \frac{1}{s+1} & -\frac{1}{s} \\ 2 & 1 + \frac{1}{s+2} \end{bmatrix} = \begin{bmatrix} \frac{s+2}{s+1} & -\frac{1}{s} \\ 2 & \frac{s+3}{s+2} \end{bmatrix} \qquad (3-40)$$

The closed-loop transfer function matrix is

$$M(s) = [I + G(s)H(s)]^{-1}G(s) = \frac{1}{\Delta}\begin{bmatrix} \frac{s+3}{s+2} & \frac{1}{s} \\ -2 & \frac{s+2}{s+1} \end{bmatrix}\begin{bmatrix} \frac{1}{s+1} & -\frac{1}{s} \\ 2 & \frac{1}{s+2} \end{bmatrix} \qquad (3-41)$$

Where

$$\Delta = \frac{s+2}{s+1}\frac{s+3}{s+2} + \frac{2}{s} = \frac{s^2+5s+2}{s(s+1)} \qquad (3-42)$$

Thus,

$$M(s) = \frac{s(s+1)}{s^2+5s+2}\begin{bmatrix} \frac{3s^2+9s+4}{s(s+1)(s+2)} & -\frac{1}{s} \\ 2 & \frac{3s+2}{s(s+1)} \end{bmatrix} \qquad (3-43)$$

3-2 Signal-flow Graph (SFGs)

A signal-flow graph (SFG) may be regarded as a simplified version of a block diagram. Besides the differences in the physical appearance of the SFG and the block diagram, the signal-flow graph is constrained by more rigid mathematical rules, whereas the block-diagram notation is more liberal. An SFG may be defined as a graphical means of portraying the input-output relationships among the variables of a set of linear algebraic equations.

Consider a linear system that is described by a set of N algebraic equations:

$$y_j = \sum_{k=1}^{N} a_{kj}\, y_k \qquad j = 1,2,\ldots,N \qquad (3-44)$$

It should be pointed out that these N equations are written in the form of cause-and-effect relations:

$$\text{jth effect} = \sum_{k=1}^{N} (\text{gain from } k \text{ to } j) \times (k\text{th cause}) \qquad (3-45)$$

or simply

$$\text{Output} = \sum (\text{gain}) \times (\text{input}) \qquad (3-46)$$

This is the single most important axiom in forming the set of algebraic equations for SFGs. When the system is represented by a set of integrodifferential equations, we must first transform these into Laplace-transform equations and then rearrange the latter in the form of Eq. (3-31), or

$$Y_j(s) = \sum_{k=1}^{N} G_{kj}(s) Y_k(s) \qquad j = 1,2,\ldots,N \qquad (3-47)$$

Basic Elements of an SFG. When constructing an SFG, junction points, or nodes, are used to represent variables. The nodes are connected by line segments called branches, according to the cause-and-effect equations. The branches have associated branch gains and directions. A signal can transmit through a branch only in the direction of the arrow. In general, given a set of equations such as Eq. (3-31) or Eq. (3-47), the construction of the SFG is basically a matter of

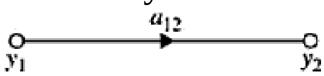


Figure 3-24 Signal flow graph of $y_2 = a_{12} y_1$.

following through the cause-and-effect relations of each variable in terms of itself and the others. For instance, consider that a linear system is represented by the simple algebraic equation

$$y_2 = a_{12} y_1 \qquad (3-48)$$

where y_1 is the input, y_2 is the output, and a_{12} is the gain, or transmittance, between the two variables. The SFG representation of Eq. (3-48) is shown in Fig. 3-24. Notice that the branch directing from node y_1 (input) to node y_2 (output) expresses the dependence of y_2 on y_1 but not the reverse. The branch between the input node and the output node should be interpreted as a unilateral amplifier with gain a_{12}, so when a signal of one unit is applied at the input y_1, a signal of strength $a_{12} y_1$ is delivered at node y_2. Although algebraically Eq. (3-48) can be written as

$$y_1 = \frac{1}{a_{12}} y_2 \qquad (3-49)$$

the SFG of Fig. 3-24 does not imply this relationship. If Eq. (3-49) is valid as a cause-and-effect equation, a new SFG should be drawn with y_2 as the input and y_1 as the output.

Example 3-7: [59] As an example on the construction of an SFG, consider the following set of algebraic equations:

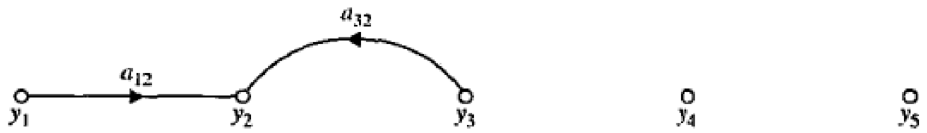
$$\begin{aligned} y_2 &= a_{12}y_1 + a_{32}y_3 \\ y_3 &= a_{23}y_2 + a_{43}y_4 \\ y_4 &= a_{24}y_2 + a_{34}y_3 + a_{44}y_4 \\ y_5 &= a_{25}y_2 + a_{45}y_4 \end{aligned} \qquad (3-50)$$

The SFG for these equations is constructed, step by step, in Fig. 3-25.

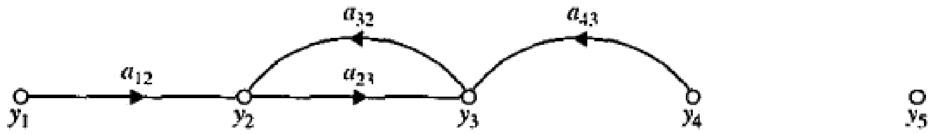
Summary of the Basic Properties of SFG. The important properties of the SFG that have been covered thus far are summarized as follows.

1. SFG applies only to linear systems.
2. The equations for which an SFG is drawn must be algebraic equations in the form of cause-and-effect.
3. Nodes are used to represent variables. Normally, the nodes are arranged from left to right, from the input to the output, following a succession of cause-and-effect relations through the system.
4. Signals travel along branches only in the direction described by the arrows of the branches.
5. The branch directing from node y_k to y_j represents the dependence of y_j upon y_k but not the reverse.
6. A signal y_k traveling along a branch between y_k and y_j is multiplied by the gain of the branch a_{kj} so a signal $a_{kj}y_k$ is delivered at y_j.

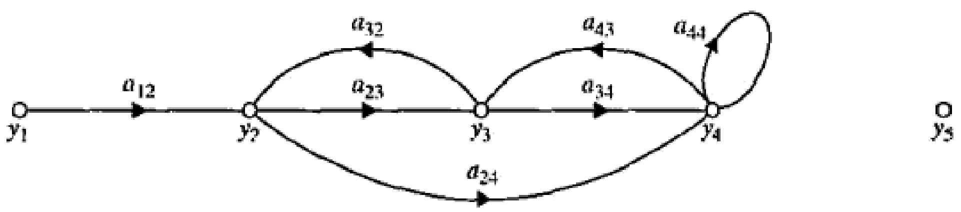
Definitions of SFG Terms. In addition to the branches and nodes defined earlier for the SFG, the following terms are useful for the purpose of identification and execution of the SFG algebra.



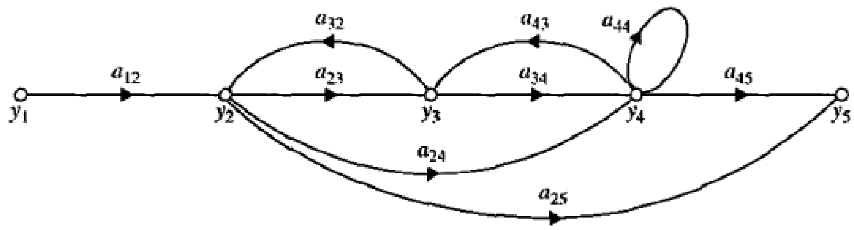
(a) $y_2 = a_{12}y_1 + a_{32}y_3$



(b) $y_2 = a_{12}y_1 + a_{32}y_3 \quad y_3 = a_{23}y_2 + a_{42}y_4$



(c) $y_2 = a_{12}y_1 + a_{32}y_3 \quad y_3 = a_{23}y_2 + a_{42}y_4 \quad y_4 = a_{24}y_2 + a_{34}y_3 + a_{44}y_4$



(d) Complete signal-flow graph

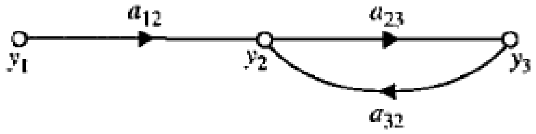
Figure 3-25 Step-by-step construction of the signal-flow graph in Eq. (3-50).

Input Node (Source): An input node is a node that has only outgoing branches (example: node y_1 in Fig. 3-24).

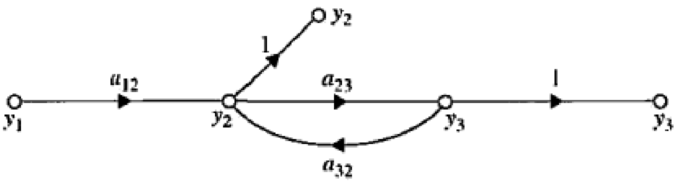
Output Node (Sink): An output node is a node that has only incoming branches: (example: node y_2 in Fig. 3-24). However, this condition is not always readily met by an output node. For instance, the SFG in Fig. 3-26(a) does not have a node that satisfies the condition of an output node. It may be necessary to regard y_2 and/or y_3 as output nodes to find the effects at these nodes due to the input. To make y_2 an output node, we simply connect a branch with unity gain from the existing node y_2 to a new node also designated as y_2, as shown in Fig. 3-26(b). The same procedure is applied to y_3, Notice that, in the modified SFG of Fig. 3-26(b), the equations $y_2 = y_2$ and $y_3 = y_3$ are added to the original equations. In general, we can make any noninput node of an SFG an output by the procedure just illustrated. However, we **cannot** convert a noninput node into an input node by reversing the branch direction of the procedure described for output nodes. For instance, node y_2 of the SFG in Fig. 3-26(a) is not an input node. If we attempt to convert it into an input node by adding an incoming branch with unity gain from another identical node y_2, the SFG of Fig. 3-27 would result. The equation that portrays the relationship at node y_2 now reads

$$y_2 = y_2 + a_{12}y_1 + a_{32}y_3 \qquad (3-51)$$

which is different from the original equation given in Fig. 3-26(a).



(a) Original signal-flow graph



(b) Modified signal-flow graph

Figure 3-26 Modification of a signal-flow graph so that y_2 and y_3 satisfy the condition as output nodes.

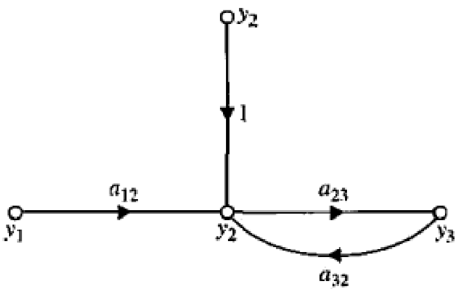


Figure 3-27 Erroneous way to make node y_2 an input node.

Path: A path is any collection of a continuous succession of branches traversed in the same direction. The definition of a path is entirely general, since it does not prevent any node from being traversed more than once. Therefore, as simple as the SFG of Fig. 3-26(a) is, it may have numerous paths just by traversing the branches a_{23} and a_{32} continuously.

Forward Path: A forward path is a path that starts at an input node and ends at an output node and along which no node is traversed more than once. For example, in the SFG of Fig. 3-25(d), y_1 is the input node, and the rest of the nodes are all possible output nodes. The forward path between y_1 and y_2 is simply the connecting branch between the two nodes. There are two forward paths between y_1 and y_3. One contains the branches from y_1 to y_2 to y_3, and the other one contains the branches from y_1 to y_2 to y_4 (through the branch with gain a_{24}) and then back to y_3 (through the branch with gain a_{43}). We should try to determine the two forward paths between y_1 and y_4. Similarly, there are three forward paths between y_1 and y_5.

Path Gain: The product of the branch gains encountered in traversing a path is called the path gain. For example, the path gain for the path $y_1 - y_2 - y_3 - y_4$ in Fig. 3-25(d) is $a_{12}a_{23}a_{34}$.

Loop: A loop is a path that originates and terminates on the same node and along which no other node is encountered more than once. For example, there are four loops in the SFG of Fig. 3-25(d). These are shown in Fig. 3-28.

Forward-Path Gain: The forward-path gain is the path gain of a forward path.

Loop Gain: The loop gain is the path gain of a loop. For example, the loop gain of the loop $y_2 - y_4 - y_3 - y_2$ in Fig. 3-28 is $a_{24}a_{43}a_{32}$.

Nontouching Loops: Two parts of an SFG are nontouching if they do not share a common node. For example, the loops $y_2 - y_3 - y_2$ and $y_4 - y_4$ of the SFG in Fig. 3-25 (d) are nontouching loops.

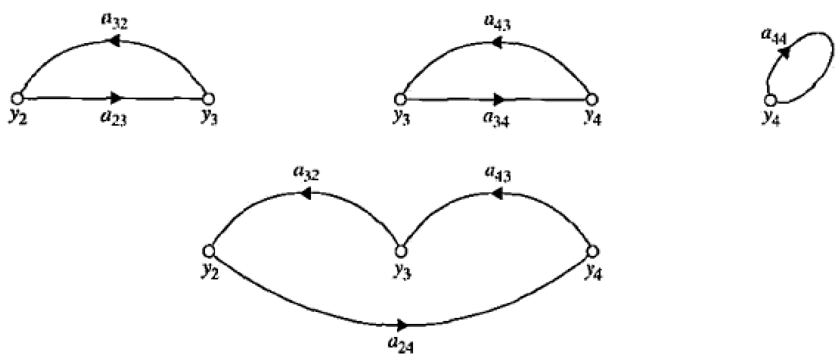


Figure 3-28 Four loops in the signal-flow graph of Fig. 3-25(d).

SFG Algebra. Based on the properties of the SFG, we can outline the following manipulation rules and algebra:

1. The value of the variable represented by a node is equal to the sum of all the signals entering the node. For the SFG of Fig. 3-29, the value of y_1 is equal to the sum of the signals transmitted through all the incoming branches; that is,

$$y_1 = a_{21}y_2 + a_{31}y_3 + a_{41}y_4 + a_{51}y_5 \qquad (3-52)$$

2. The value of the variable represented by a node is transmitted through all branches leaving the node. In the SFG of Fig. 3-29, we have

$$y_6 = a_{16}y_1 \qquad (3-53)$$
$$y_7 = a_{17}y_1$$
$$y_8 = a_{18}y_1$$

3. Parallel branches in the same direction connecting two nodes can be replaced by a single branch with gain equal to the sum of the gains of the parallel branches. An example of this case is illustrated in Fig. 3-30.

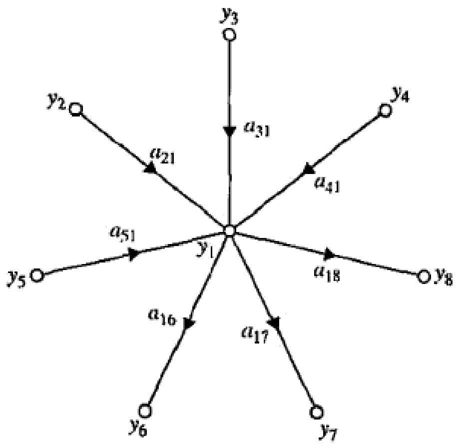


Figure 3-29 Node as a summing point and as a transmitting point.

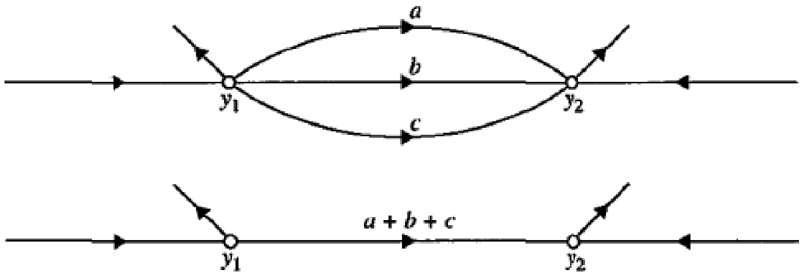


Figure 3-30 Signal-flow graph with parallel paths replaced by one with a single branch.

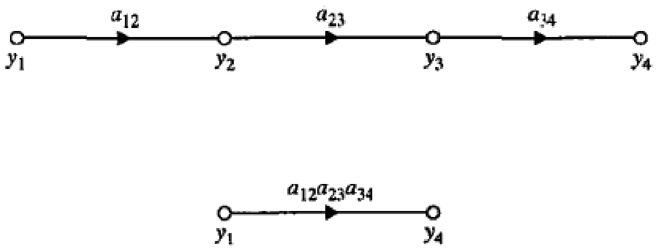


Figure 3-31 Signal-flow graph with cascade unidirectional branches replaced by a single branch.

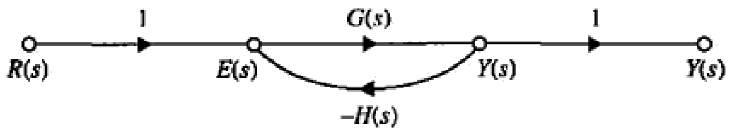


Figure 3-32 Signal-flow graph of the feedback control system shown in Fig. 3-8.

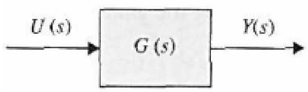
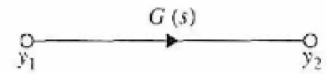
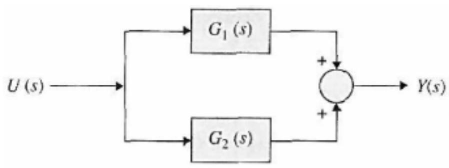
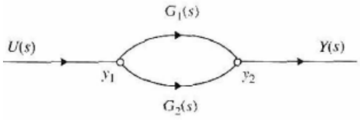
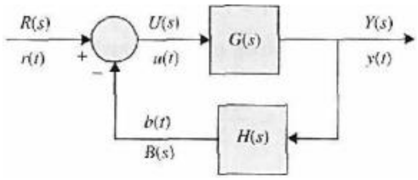
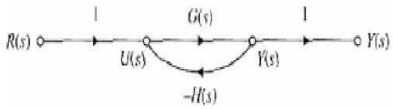
4. A series connection of unidirectional branches, as shown in Fig. 3-31, can be replaced by a single branch with gain equal to the product of the branch gains.

SFG of a Feedback Control System. The SFG of the single-loop feedback control system in Fig. 3-8 is drawn as shown in Fig. 3-32. Using the SFG algebra already outlined, the closed-loop transfer function in Eq. (3-12) can be obtained.

Relation between Block Diagrams and SFGs. The relation between block diagrams and SFGs are tabulated for three important cases, as shown in Table 3-1.

Gain Formula for SFG. Given an SFG or block diagram, the task of solving for the input-output relations by algebraic manipulation could be quite tedious. Fortunately, there is a general gain formula available that allows the determination of the input-output relations of an SFG by inspection.

TABLE 3-1 Block diagrams and their SFG equivalent representations

Simple Transfer Function	Block Diagram	Signal Flow Diagram
$\frac{Y(s)}{U(s)} = G(s)$		
Parallel Feedback		
$\frac{Y(s)}{U(s)} = \frac{G(s)}{1+G(s)H(s)}$		

Given an SFG with N forward paths and K loops, the gain between the input node y_{in} and output node y_{out}

$$M = \frac{y_{out}}{y_{in}} = \sum_{k=1}^{N} \frac{M_k \Delta_k}{\Delta} \qquad (3-54)$$

where

$$y_{in} = \text{input} - \text{node variable}$$
$$y_{out} = \text{output} - \text{node variable}$$
$$M = \text{gain between } y_{in} \text{ and } y_{out}$$
$$N = \text{total number of forward paths between } y_{in} \text{ and } y_{out}$$
$$M_k = \text{gain of the } k\text{th forward paths between } y_{in} \text{ and } y_{out}$$

$$\Delta = 1 - \sum_i L_{i1} + \sum_j L_{j2} - \sum_k L_{k3} + \cdots \qquad (3-55)$$

L_{mr} = gain product of the mth ($m = i,j,k, ...$) possible combination of r nontouching loops ($1 < r < K$).

or $\Delta = 1 -$ (sum of the gains of **all individual** loops) $+$ ($sum\ of\ products\ of\ gains\ of\ all$ $possible\ combinations\ of$ ***two*** $nontouching\ loops$) $-$ ($sum\ of\ products\ of\ gains\ of$ $all\ possible\ combinations\ of\ three\ nontouching\ loops$) $+ \cdots$

$\Delta_k =$ the A for that part of the SFG that is nontouching with the *k*th forward path.

The gain formula in Eq. (3-54) may seem formidable to use at first glance. However, A and Δ_k are the only terms in the formula that could be complicated if the SFG has a large number of loops and nontouching loops.

Care must be taken when applying the gain formula to ensure that it is applied between an **input node** and an **output node.**

Example 3-8: [59] Consider that the closed-loop transfer function $Y(s)/R(s)$ of the SFG in Fig. 3-32 is to be determined by use of the gain formula, Eq. (3-54). The following results are obtained by inspection of the SFG:

1. There is only one forward path between *R(s)* and *Y(s)*, and the forward-path gain is
$$M_1 = G(s) \qquad (3-56)$$
2. There is only one loop; the loop gain is
$$L_{11} = -G(s)H(s) \qquad (3-57)$$
3. There are no nontouching loops since there is only one loop. Furthermore, the forward path is in touch with the only loop. Thus, $\Delta_1 = 1$, and
$$\Delta = 1 - L_{11} = 1 + G(s)H(s) \qquad (3-58)$$

Using Eq. (3-54), the closed-loop transfer function is written

$$\frac{Y(s)}{R(s)} = \frac{M_1\Delta_1}{\Delta} \qquad (3-59)$$

which agrees with Eq. (3-12).

Example 3-9: [59] Consider the SFG shown in Fig. 3-25(d). Let us first determine the gain between y_1 and y_5 using the gain formula.

The three forward paths between y_1 and y_5 and the forward-path gains are

$M_1 = a_{12}a_{23}a_{34}a_{45}$ Forward path: $y_1 - y_2 - y_3 - y_4 - y_5$

$M_2 = a_{12}a_{25}$ F rward path: $y_1 - y_2 - y_5$

$M_3 = a_{12}a_{24}a_{45}$ Forward path: $y_1 - y_2 - y_4 - y_5$

The four loops of the SFG are shown in Fig. 3-28. The loop gains are

$$L_{11} = a_{23}a_{32}, \qquad L_{21} = a_{34}a_{43}, \qquad L_{31} = a_{24}a_{43}a_{32}, \qquad L_{41} = a_{44}$$

There is only one pair of nontouching loops; that is, the two loops are

$$y_2 - y_3 - y_2 \text{ and } y_4 - y_4$$

Thus, the product of the gains of the two nontouching loops is

$$L_{12} = a_{23}a_{32}a_{44} \qquad (3-60)$$

All the loops are in touch with forward paths M_1 and M_3. Thus, $\Delta_1 = \Delta_3 = 1$. Two of the loops are not in touch with forward path M_2 These loops are $y_3 - y_4 - y_3$ and $y_3 - y_4$. Thus,

$$\Delta_2 = 1 - a_{34}a_{43} - a_{44} \qquad (3-61)$$

Substituting these quantities into Eq. (3-54), we have

$$\frac{y_5}{y_1} = \frac{M_1\Delta_1 + M_2\Delta_2 + M_3\Delta_3}{\Delta}$$
$$= \frac{(a_{12}a_{23}a_{34}a_{45}) + (a_{12}a_{25})(1 - a_{34}a_{43} - a_{44}) + a_{12}a_{24}a_{45}}{1 - (a_{23}a_{32} + a_{34}a_{43} + a_{24}a_{43}a_{32} + a_{44}) + a_{23}a_{32}a_{44}} \qquad (3-62)$$

where

$$\Delta = 1 - (L_{11} + L_{21} + L_{31} + L_{41})$$
$$= 1 - (a_{23}a_{32} + a_{34}a_{43} + a_{24}a_{43}a_{32} + a_{44}) + a_{23}a_{32}a_{44} \quad (3-63)$$

We should verify that choosing y_2 as the output,

$$\frac{y_2}{y_1} = \frac{a_{12}(1 - a_{34}a_{43} - a_{44})}{\Delta} \quad (3-64)$$

where Δ is given in Eq. (3-63).

Example 3-10: [59] Consider the SFG in Fig. 3-33. The following input-output relations are obtained by use of the gain formula:

$$\frac{y_2}{y_1} = \frac{1 + G_3H_2 + H_4 + G_3H_2H_4}{\Delta} \quad (3-65)$$

$$\frac{y_4}{y_1} = \frac{G_1G_2(1 + H_4)}{\Delta} \quad (3-66)$$

$$\frac{y_6}{y_1} = \frac{y_7}{y_1} = \frac{G_1G_2G_3G_4 + G_1G_5(1 + G_3H_2)}{\Delta} \quad (3-67)$$

Where

$$\Delta = 1 + G_1H_1 + G_3H_2 + G_1G_2G_3H_3 + H_4 + G_1G_3H_1H_2$$
$$+G_1H_1H_4 + G_3H_2H_4 + G_1G_2G_3H_3H_4 + G_1G_3H_1H_2H_4 \quad (3-68)$$

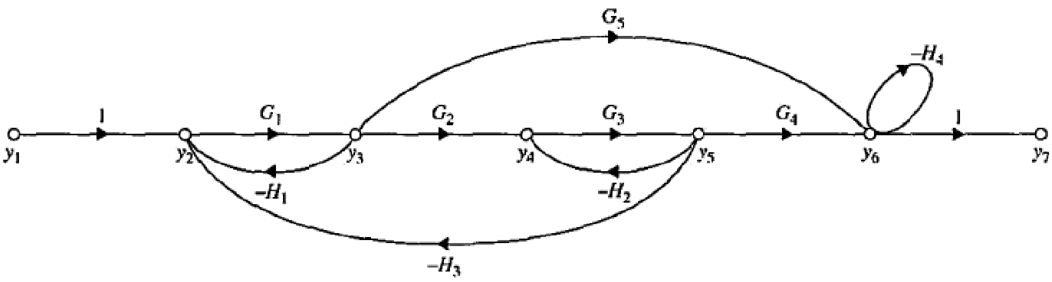


Figure 3-33 Signal-flow graph for Example 3-10.

Application of the Gain Formula between Output Nodes and Noninput Nodes. It was pointed out earlier that the gain formula can only be applied between a pair of input and output nodes. Often, it is of interest to find the relation between an output-node variable and a noninput-node variable. For example, in the SFG of Figure 3-33, it may be of interest to find the relation y_7/y_2, which represents the dependence of y_7 upon y_2, the latter is not an input.

We can show that, by including an input node, the gain formula can still be applied to find the gain between a noninput node and an output node. Let y_{in} be an input and y_{out} be an output node of a SFG. The gain, y_{out}/y_2 where y_2 is not an input, may be written as

$$\frac{y_{out}}{y_2} = \frac{\frac{y_{out}}{y_{in}}}{\frac{y_2}{y_{in}}} = \frac{\frac{\sum M_k\Delta_k|_{from y_{in} to y_{out}}}{\Delta}}{\frac{\sum M_k\Delta_k|_{from y_{in} to y_2}}{\Delta}} \quad (3-69)$$

Because Δ is independent of the inputs and the outputs, the last equation is written

$$\frac{y_{out}}{y_2} = \frac{\sum M_k\Delta_k|_{\mathrm{from}\, y_{\mathrm{in}}\, \mathrm{to}\, y_{\mathrm{out}}}}{\sum M_k\Delta_k|_{\mathrm{from}\, y_{\mathrm{in}}\, \mathrm{to}\, y_2}} \quad (3-70)$$

Notice that Δ does not appear in the last equation.

Example 3-11: [59] From the SFG in Fig. 3-33, the gain between y_2 and y_7 is written

$$\frac{y_7}{y_2} = \frac{y_7/y_1}{y_2/y_1} = \frac{G_1G_2G_3G_4 + G_1G_5(1 + G_3H_2)}{1 + G_3H_2 + H_4 + G_3H_2H_4} \qquad (3-71)$$

Application of the Gain Formula to Block Diagrams. Because of the similarity between the block diagram and the SFG, the gain formula in Eq. (3-54) can directly be applied to the block diagram to determine the transfer function of the system. However, in complex systems, to be able to identify all the loops and nontouching parts clearly, it may be helpful if an equivalent SFG is drawn for the block diagram first before applying the gain formula.

Example 3-12: [59] To illustrate how an equivalent SFG of a block diagram is constructed and how the gain formula is applied to a block diagram, consider the block diagram shown in Fig. 3-34(a). The equivalent SFG of the system is shown in Fig. 3-34(b). Notice that since a node on the SFG is interpreted as the summing point of all incoming signals to the node, the negative feedbacks on the block diagram are represented by assigning negative gains to the feedback paths on the SFG. First we can identify the forward paths and loops in the system and their corresponding gains. That is:

Forward Path Gains: $1.\, G_1G_2G_3;\, 2.\, G_1G_4$

Loop Gains: $1.\, -G_1G_2H_1;\, 2..\, -G_2G_3H_2;\, 3.\, -G_1G_2G_3;\, 4.\, -G_4H_2;\, 5.\, -G_1G_4$

The closed-loop transfer function of the system is obtained by applying Eq. (3-54) to either the block diagram or the SFG in Fig. 3-34. That is

$$\frac{Y(s)}{R(s)} = \frac{G_1G_2G_3 + G_1G_4}{\Delta} \qquad (3-72)$$

where

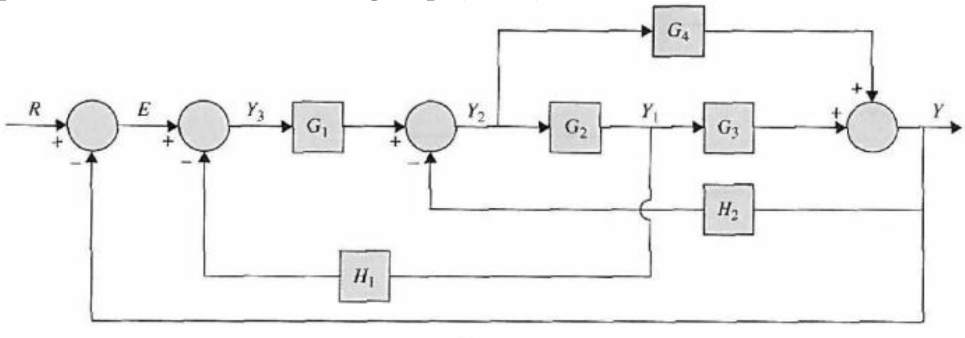
$$\Delta = 1 + G_1G_2H_1 + G_2G_3H_2 + G_1G_2G_3 + G_4H_2 + G_1G_4 \qquad (3-73)$$

Similarly,

$$\frac{E(s)}{R(s)} = \frac{1 + G_1G_2H_1 + G_2G_3H_2 + G_4H_2}{\Delta} \qquad (3-74)$$

$$\frac{E(s)}{R(s)} = \frac{G_1G_2G_3 + G_1G_4}{1 + G_1G_2H_1 + G_2G_3H_2 + G_4H_2} \qquad (3-75)$$

The last expression is obtained using Eq. (3-70).



(a)

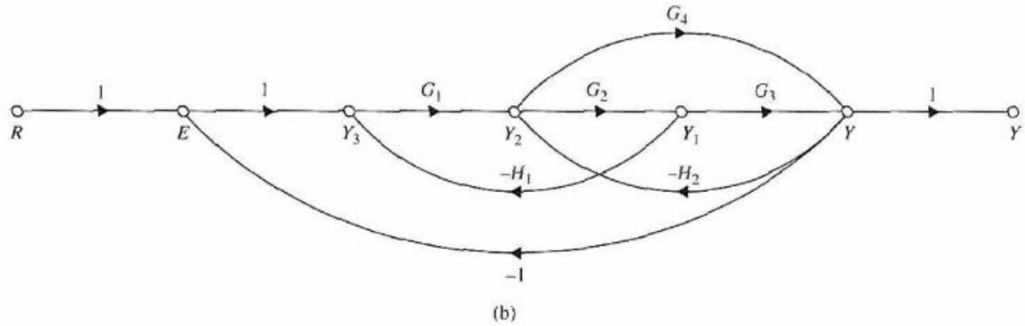


Figure 3-34 (a) Block diagram of a control system, (b) Equivalent signal-flow graph.

Simplified Gain Formula. From Example 3-12, we can see that *all loops and forward paths are touching* in this case. As a general rule, if there are no nontouching loops and forward paths (e.g., $y_2 - y_3 - y_2$ and $y_4 - y_4$ in Example 3-9) in the block diagram or SFG of the system, then Eq. (3-54) takes a far simpler look, as shown next.

$$M = \frac{y_{out}}{y_{in}} = \sum \frac{Forward\ Path\ Gains}{1 - Loop\ Gains} \qquad (3-76)$$

Redo Examples 3-8 through 3-12 to confirm the validity of Eq. (3-76).

3-3 Automatic controllers

An automatic controller compares the actual value of the plant output with the reference input (desired value), determines the deviation, and produces a control signal that will reduce the deviation to zero or to a small value. The manner in which the automatic controller produces the control signal is called the *control action*. Figure 3–8 is a block diagram of an industrial control system, which Consists of an automatic controller, an actuator, a plant, and a sensor (measuring element).

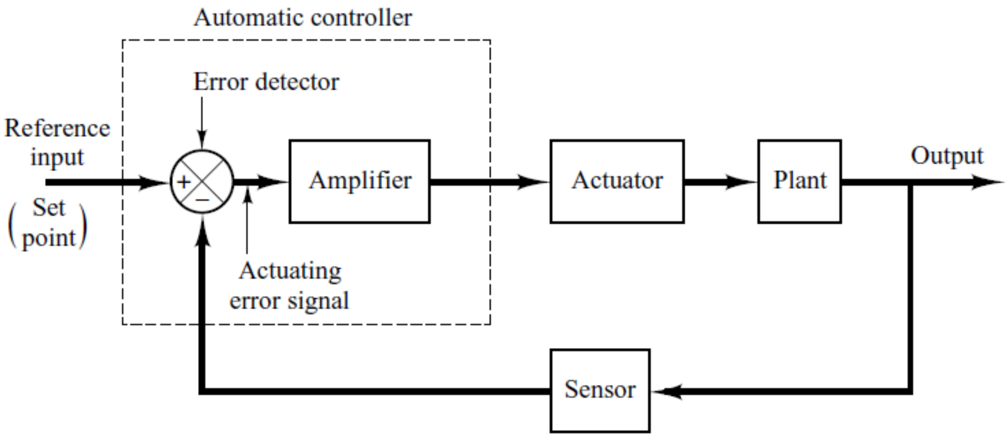


Figure 3–35 Block diagram of an industrial control system, which consists of an automatic controller, an actuator, a plant, and a sensor (measuring element).

The controller detects the actuating error signal, which is usually at a very low power level, and amplifies it to a sufficiently high level. The output of an automatic controller is fed to an actuator, such as an electric motor, a hydraulic motor, or a pneumatic motor or valve. (The actuator is a power device that produces the input to the plant according to the control signal so that the output signal will approach the reference input signal.)

The sensor or measuring element is a device that converts the output variable into another suitable variable, such as a displacement, pressure, voltage, etc., that can be used to compare the output to the reference input signal. This element is in the feedback path of the closed-loop system. The set point of the controller must be converted to a reference input with the same units as the feedback signal from the sensor or measuring element.

Classifications of Industrial Controllers:

Most industrial controllers may be classified according to their control actions as:

1. Two-position or on–off controllers
2. Proportional controllers
3. Integral controllers
4. Proportional-plus-integral controllers
5. Proportional-plus-derivative controllers
6. Proportional-plus-integral-plus-derivative controllers

Most industrial controllers use electricity or pressurized fluid such as oil or air as power sources. Consequently, controllers may also be classified according to the kind of power employed in the operation, such as pneumatic controllers, hydraulic controllers, or electronic controllers. What kind of controller to use must be decided based on the nature of the plant and the operating conditions, including such considerations as safety, cost, availability, reliability, accuracy, weight, and size.

Two-Position or On-Off Control Action. In a two-position control system, the actuating element has only two fixed positions, which are, in many cases, simply on and off. Two-position or on–off control is relatively simple and inexpensive and, for this reason is very widely used in both industrial and domestic control systems.

Let the output signal from the controller be u(t) and the actuating error signal be e(t). In two-position control, the signal u(t) remains at either a maximum or minimum value, depending on whether the actuating error signal is positive or negative, so that

$$u(t) = U_1, \quad for \;\; e(t) > 0$$
$$= U_2, \quad for \;\; e(t) < 0 \qquad (3-77)$$

where U_1 and U_2 are constants. The minimum value U_2 is usually either zero or $-U_1$. Two-position controllers are generally electrical devices, and an electric solenoid-operated valve is widely used in such controllers. Pneumatic proportional controllers with very high gains act as two-position controllers and are sometimes called pneumatic two position controllers.

Figures 3–36(a) and (b) show the block diagrams for two-position or on–off controllers. The range through which the actuating error signal must move before the switching occurs is called the differential gap.

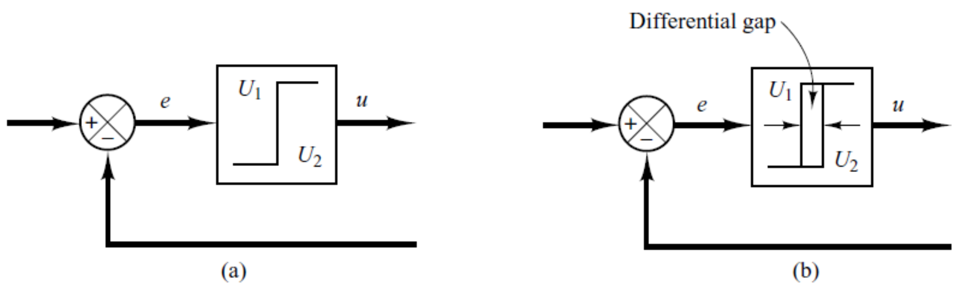


Figure 3–36 (a) Block diagram of an on–off controller; (b) block diagram of an on–off controller with differential gap.

A differential gap is indicated in Figure 3–36(b). Such a differential gap causes the controller output u(t) to maintain its present value until the actuating error signal has moved slightly beyond the zero value. In some cases, the differential gap is a result of unintentional friction and lost motion; however, quite often it is intentionally provided in order to prevent too-frequent operation of the on–off mechanism.

Consider the liquid-level control system shown in Figure 3–37(a), where the electromagnetic valve shown in Figure 3–37(b) is used for controlling the inflow rate. This valve is either open or closed. With this two-position control, the water inflow rate is either a positive constant or zero. As shown in Figure 3–38, the output signal continuously moves between the two limits required to cause the actuating element to move from one fixed position to the other. Notice that the output curve follows one of two exponential curves, one corresponding to the filling curve and the other to the emptying curve. Such output oscillation between two limits is a typical response characteristic of a system under two-position control.

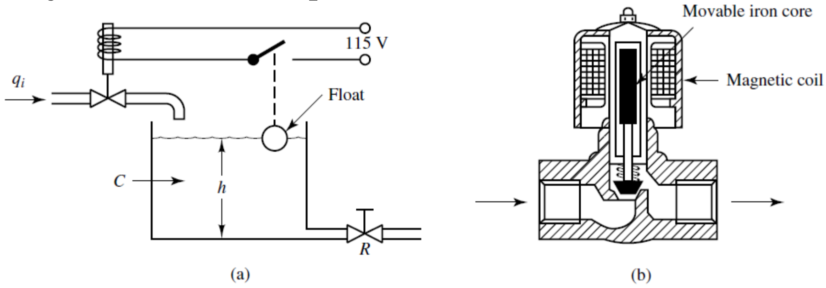


Figure 3–37 (a) Liquid-level control system; (b) electromagnetic valve.

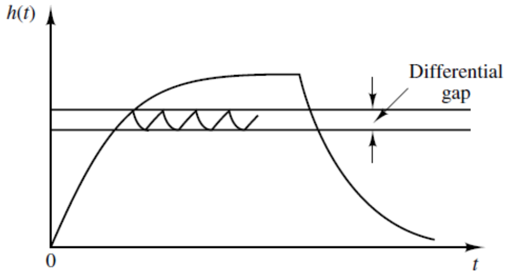


Figure 3–38 Level $h(t)$-versus-t curve for the system shown in Figure 2–8(a).

From Figure 3–38, we notice that the amplitude of the output oscillation can be reduced by decreasing the differential gap. The decrease in the differential gap, however, increases the number of on–off switching per minute and reduces the useful life of the component. The magnitude of the differential gap must be determined from such considerations as the accuracy required and the life of the component.

Proportional Control Action. For a controller with proportional control action, the relationship between the output of the controller $u(t)$ and the actuating error signal $e(t)$ is

$$u(t) = K_p e(t) \qquad (3-78)$$

or, in Laplace-transformed quantities,

$$\frac{u(s)}{E(s)} = K_p \qquad (3-79)$$

where K_p is termed the proportional gain.

Whatever the actual mechanism may be and whatever the form of the operating power, the proportional controller is essentially an amplifier with an adjustable gain.

Integral Control Action. In a controller with integral control action, the value of the controller output $u(t)$ is changed at a rate proportional to the actuating error signal $e(t)$.That is,

$$\frac{du(t)}{dt} = K_i e(t) \qquad (3-80)$$

or

$$u(t) = K_i \int_0^t e(t)dt \qquad (3-81)$$

where K_i is an adjustable constant. The transfer function of the integral controller is

$$\frac{U(s)}{E(s)} = \frac{K_i}{s} \qquad (3-82)$$

Proportional-Plus-Integral Control Action. The control action of a proportional-plus-integral controller is defined by

$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(t)dt \qquad (3-83)$$

or the transfer function of the controller is

$$\frac{U(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s}\right) \qquad (3-84)$$

Where T_i is called the *integral time*.

Proportional-Plus-Derivative Control Action. The control action of a proportional-plus-derivative controller is defined by

$$u(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt} \qquad (3-85)$$

and the transfer function is

$$\frac{U(s)}{E(s)} = K_p(1 + T_d s) \qquad (3-86)$$

Where T_d is called the *derivative time*.

Proportional-Plus-Integral-Plus-Derivative Control Action. The combination of proportional control action, integral control action, and derivative control action is termed proportional-plus-integral-plus-derivative control action. It has the advantages of each of the three individual control actions. The equation of a controller with this combined action is given by

$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(t)dt + K_p T_d \frac{de(t)}{dt} \qquad (3-87)$$

or the transfer function is

$$\frac{U(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s} + T_d s\right) \qquad (3-88)$$

where K_p is the proportional gain, T_i is the integral time, and T_d is the derivative time. The block diagram of a proportional-plus-integral-plus-derivative controller is shown in Figure 3–39.

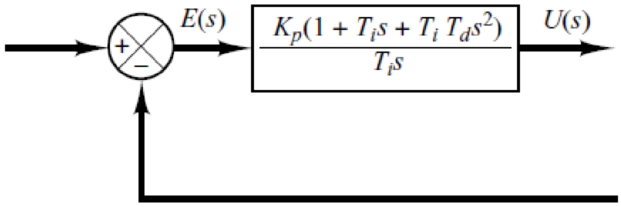


Figure 3-39 Block diagram of a proportional-plus-integral-plus-derivative controller.

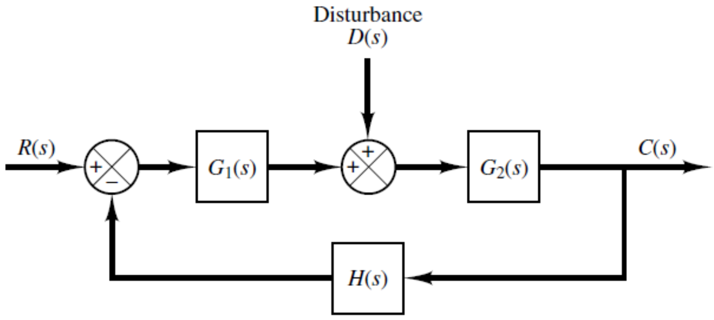


Figure 3-40 Closed-loop system subjected to a disturbance.

Closed-Loop System Subjected to a Disturbance. Figure 3-40 shows a closed loop system subjected to a disturbance. When two inputs (the reference input and disturbance) are present in a linear time-invariant system, each input can be treated independently of the other; and the outputs corresponding to each input alone can be added to give the complete output. The way each input is introduced into the system is shown at the summing point by either a plus or minus sign.

Consider the system shown in Figure 3–40. In examining the effect of the disturbance $D(s)$, we may assume that the reference input is zero; we may then calculate the response $C_D(s)$ to the disturbance only. This response can be found from

$$\frac{C_D(s)}{D(s)} = \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)} \qquad (3-89)$$

On the other hand, in considering the response to the reference input $R(s)$, we may assume that the disturbance is zero. Then the response $C_R(s)$ to the reference input $R(s)$ can be obtained from

$$\frac{C_R(s)}{R(s)} = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)} \qquad (3-90)$$

The response to the simultaneous application of the reference input and disturbance can be obtained by adding the two individual responses. In other words, the response $C(s)$ due to the simultaneous application of the reference input $R(s)$ and disturbance $D(s)$ is given by

$$C(s) = C_R(s) + C_D(s)$$
$$= \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}[G_1(s)R(s) + D(s)] \qquad (3-91)$$

Consider now the case where $|G_1(s)H(s)| \gg 1$ and $|G_1(s)G_2(s)H(s)| \gg 1$. In this case, the closed-loop transfer function $C_D(s)/D(s)$ becomes almost zero, and the effect of the disturbance is suppressed. This is an advantage of the closed-loop system.

On the other hand, the closed-loop transfer function $C_R(s)/R(s)$ approaches $1/H(s)$ as the gain of $G_1(s)G_2(s)H(s)$ increases. This means that if $|G_1(s)G_2(s)H(s)| \gg 1$, then the closed-loop transfer function $C_R(s)/R(s)$ becomes independent of $G_1(s)$ and $G_2(s)$ and inversely proportional to $H(s)$, so that the variations of $G_1(s)$ and $G_2(s)$ do not affect the closed-loop transfer function $C_R(s)/R(s)$. This is another advantage of the closed-loop system. It can easily be seen that any closed-loop system with unity feedback, $H(s) = 1$, tends to equalize the input and output.

Procedures for Drawing a Block Diagram. To draw a block diagram for a system, first write the equations that describe the dynamic behavior of each component. Then take the Laplace transforms of these equations, assuming zero initial conditions, and represent each Laplace-transformed equation individually in block form. Finally, assemble the elements into a complete block diagram.

As an example, consider the RC circuit shown in Figure 3–41(a).The equations for this circuit are

$$i = \frac{e_i - e_o}{R} \qquad (3-92)$$

$$\mathrm{e_o} = \frac{\int i dt}{C} \qquad (3-93)$$

The Laplace transforms of Equations (3–92) and (3–93), with zero initial condition, become

$$I(s) = \frac{E_i(s) - \mathrm{E_o}(s)}{\mathrm{R}} \qquad (3-94)$$

$$E_o(\mathrm{s}) = \frac{I(s)}{Cs} \qquad (3-95)$$

Equation (3–94) represents a summing operation, and the corresponding diagram is shown in Figure 3–41(b). Equation (3–95) represents the block as shown in Figure 3–41(c). Assembling these two elements, we obtain the overall block diagram for the system as shown in Figure 3–41(d).

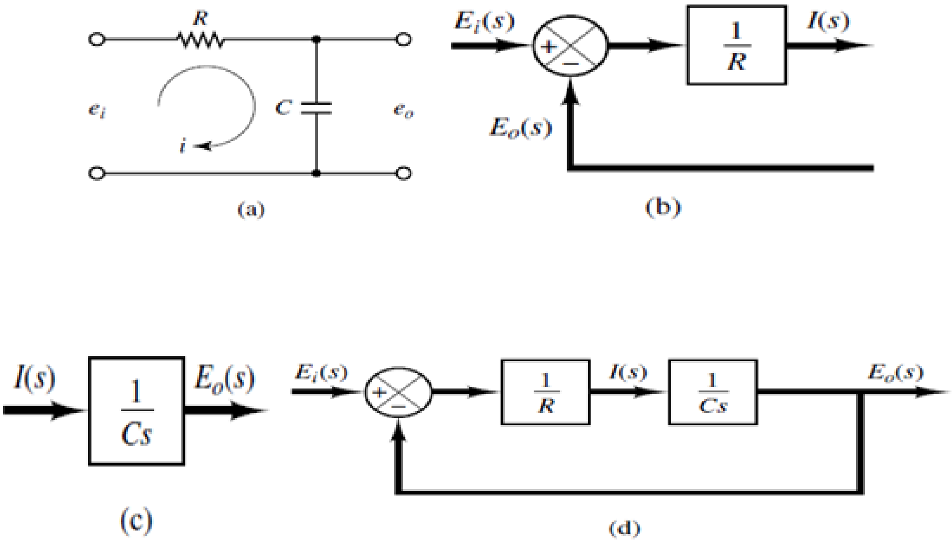


Figure 3–41 (a) RC circuit; (b) block diagram representing Equation (3-94); (c) block diagram representing Equation (3-95); (d) block diagram of the RC circuit.

CHAPTER FOUR

Modeling in State Space

We shall present introductory material on state-space analysis of control systems. And the objective is to introduce the basic methods of state variables and state equations so that we can gain a working knowledge of the subject for further studies when the state-space approach is used for modem and optimal control design. Specifically, the closed-form solutions of linear time-invariant state equations are presented. Various transformations that may be used to facilitate the analysis and design of linear control systems in the state-variable domain are introduced. The relationship between the conventional transfer-function approach and the state-variable approach is established so that the analyst will be able to investigate a system problem with various alternative methods. The controllability and observability of linear systems are defined and their applications investigated. Some state-space controller design problems appear in the end. We also present MATLAB tools to solve most state-space problems.

4-1 Modern Control Theory

The modern trend in engineering systems is toward greater complexity, due mainly to the requirements of complex tasks and good accuracy. Complex systems may have multiple inputs and multiple outputs and may be time varying. Because of the necessity of meeting increasingly stringent requirements on the performance of control systems, the increase in system complexity, and easy access to large scale computers, modern control theory, which is a new approach to the analysis and design of complex control systems, has been developed since around 1960.This new approach is based on the concept of state. The concept of state by itself is not new, since it has been in existence for a long time in the field of classical dynamics and other fields.

Modern Control Theory Versus Conventional Control Theory. Modern control theory is contrasted with conventional control theory in that the former is applicable to multiple-input, multiple-output systems, which may be linear or nonlinear, time invariant or time varying, while the latter is applicable only to linear time invariant single-input, single-output systems. Also, modern control theory is essentially time-domain approach and frequency domain approach (in certain cases such as H-infinity control), while conventional control theory is a complex frequency-domain approach. Before we proceed further, we must define state, state variables, state vector, and state space.

State. The state of a dynamic system is the smallest set of variables (called state *variables*) such that knowledge of these variables at $= t_0$, together with knowledge of the input for $t \geq t_0$, completely determines the behavior of the system for any time $t \geq t_0$.

Note that the concept of state is by no means limited to physical systems. It is applicable to biological systems, economic systems, social systems, and others.

State Variables. The state variables of a dynamic system are the variables making up the smallest set of variables that determine the state of the dynamic system. If at least n variables $x_1, x_2, \ldots, x_n$ are needed to completely describe the behavior of a dynamic system (so that once the input is given for $t \geq t_0$ and the initial state at $t = t_0$ is specified, the future state of the system is completely determined), then such n variables are a set of state variables.

Note that state variables need not be physically measurable or observable quantities. Variables that do not represent physical quantities and those that are neither measurable nor observable can be chosen as state variables. Such freedom in choosing state variables is an advantage of the state-space methods. Practically, however, it is convenient to choose easily measurable quantities for the state variables, if this is possible at all, because optimal control laws will require the feedback of all state variables with suitable weighting.

State Vector. If n state variables are needed to completely describe the behavior of a given system, then these n state variables can be considered the n components of a vector x. Such a vector is called a *state vector*. A state vector is thus a vector that determines uniquely the system state $\boldsymbol{x}(t)$ for any time $\geq t_0$, once the state at $t = t_0$ is given and the input $u(t)$ for $t \geq t_0$ is specified.

State Space. The *n*-dimensional space whose coordinate axes consist of the x_1 axis, x_2 axis, ... , x_n axis, where x_1, x_2, ... , x_n are state variables, is called a *state space*. Any state can be represented by a point in the state space.

State-Space Equations. In state-space analysis we are concerned with three types of variables that are involved in the modeling of dynamic systems: input variables, output variables, and state variables. The state-space representation for a given system is not unique, except that the number of state variables is the same for any of the different state-space representations of the same system.

The dynamic system must involve elements that memorize the values of the input for $\geq t_1$. Since integrators in a continuous-time control system serve as memory devices, the outputs of such integrators can be considered as the variables that define the internal state of the dynamic system. Thus the outputs of integrators serve as state variables. The number of state variables to completely define the dynamics of the system is equal to the number of integrators involved in the system.

Assume that a multiple-input, multiple-output system involves n integrators. Assume also that there are r inputs $u_1(t),\ u_2(t), \dots, u_r(t)$ and m outputs $y_1(t), y_2(t), \dots, y_m(t)$. Define n outputs of the integrators as state variables: $x_1(t), x_2(t), \dots, x_n(t)$ Then the system may be described by

$$\begin{aligned}
\dot{x}_1(t) &= f_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
\dot{x}_2(t) &= f_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
&\vdots \\
\dot{x}_n(t) &= f_n(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t)
\end{aligned} \qquad (4-1)$$

The outputs $y_1(t), y_2(t), \dots, y_m(t)$ of the system may be given by

$$\begin{aligned}
y_1(t) &= g_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
y_2(t) &= g_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\
&\vdots \\
y_m(t) &= g_m(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t)
\end{aligned} \qquad (4-2)$$

If we define

$$\mathrm{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \cdot \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix}, \qquad \mathrm{f(x,u,t)} = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ f_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ \cdot \\ \cdot \\ \cdot \\ f_n(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \end{bmatrix}$$

$$\mathrm{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \cdot \\ \cdot \\ \cdot \\ y_m(t) \end{bmatrix}, \qquad \mathrm{g(x,u,t)} = \begin{bmatrix} g_1(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ g_2(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \\ \cdot \\ \cdot \\ \cdot \\ g_m(x_1, x_2, \dots, x_n; u_1, u_2, \dots, u_r; t) \end{bmatrix}, \qquad \mathrm{u(t)} = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ \cdot \\ \cdot \\ u_r(t) \end{bmatrix}$$

then Equations (4–1) and (4–2) become

$$\dot{\mathrm{x}}(\mathrm{t}) = \mathrm{f(x,u,t)} \qquad (4-3)$$

$$\mathrm{y(t)} = \mathrm{g(x,u,t)} \qquad (4-4)$$

where Equation (4–3) is the state equation and Equation (4–4) is the output equation. If vector functions f and/or g involve time t explicitly, then the system is called a time varying system.

If Equations (4–3) and (4–4) are linearized about the operating state, then we have the following linearized state equation and output equation:

$$\dot{\mathrm{x}}(t) = \mathrm{A(t)x(t) + B(t)u(t)} \qquad (4-5)$$

$$\mathrm{y}(t) = \mathrm{C(t)x(t) + D(t)u(t)} \qquad (4-6)$$

where $A(t)$ is called the state matrix, $B(t)$ the input matrix, $C(t)$ the output matrix, and $D(t)$ the direct transmission matrix. A block diagram representation of Equations (4–5) and (4–6) is shown in Figure 4–1.

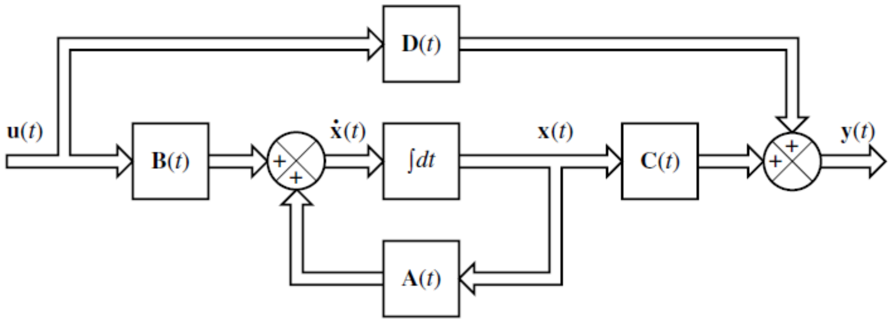


Figure 4–1 Block diagram of the linear, continuous time control system represented in state space.

If vector functions f and g do not involve time t explicitly then the system is called a time-invariant system. In this case, Equations (4–5) and (4–6) can be simplified to

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad (4-7)$$
$$\dot{y}(t) = Cx(t) + Du(t) \qquad (4-8)$$

Equation (4–7) is the state equation of the linear, time-invariant system and Equation (4–8) is the output equation for the same system. We shall be concerned mostly with systems described by Equations (4–7) and (4–8).

Example 4-1: [6] The mechanical system shown in Figure 4–2.We assume that the system is linear. The external force $u(t)$ is the input to the system, and the displacement $y(t)$ of the mass is the output. The displacement $y(t)$ is measured from the equilibrium position in the absence of the external force. This system is a single-input, single-output system.

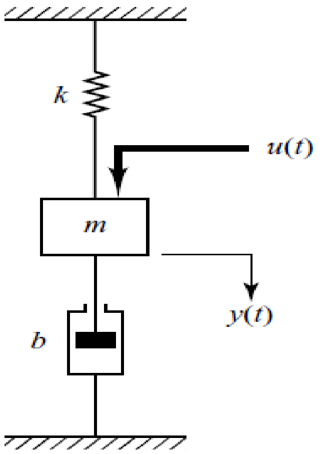


Figure 4–2 Mechanical system.

From the diagram, the system equation is

$$m\ddot{y} + b\dot{y} + ky = u \qquad (4-9)$$

This system is of second order. This means that the system involves two integrators. Let us define state variables $x_1(t)\ and\ x_2(t)$ as

$$x_1(t) = y(t)$$
$$x_2(t) = \dot{y}(t)$$

Then we obtain

$$\dot{x_1} = x_2$$

$$\dot{x}_2 = \frac{1}{m}(-ky - b\dot{y}) + \frac{1}{m}u$$

or

$$\dot{x}_1 = x_2 \qquad (4-10)$$

$$\dot{x}_2 = -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \frac{1}{m}u \qquad (4-11)$$

The output equation is

$$y = x_1 \qquad (4-12)$$

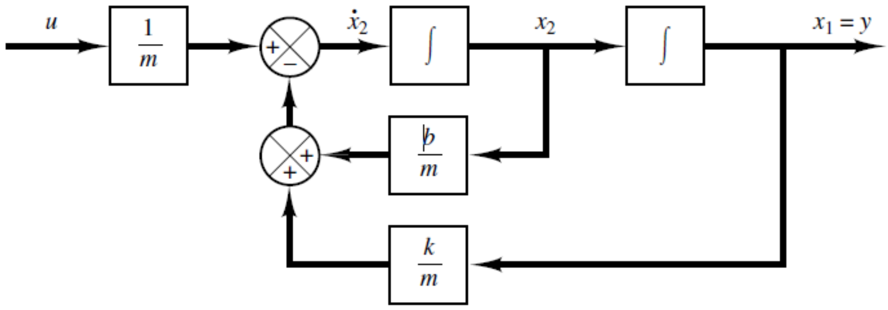


Figure 4–3 Block diagram of the mechanical system shown in Figure 3–24.

In a vector-matrix form, Equations (4–10) and (4–11) can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \qquad (4-13)$$

The output equation, Equation (4–12), can be written as

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (4-14)$$

Equation (4–13) is a state equation and Equation (4–14) is an output equation for the system.

They are in the standard form:

$$\dot{\mathrm{x}} = \mathrm{Ax} + \mathrm{Bu}$$

$$\mathrm{y} = \mathrm{Cx} + \mathrm{Du}$$

where

$$\mathrm{A} = \begin{bmatrix} 0 & 1 \\ -\frac{\mathrm{k}}{\mathrm{m}} & -\frac{\mathrm{b}}{\mathrm{m}} \end{bmatrix}, \qquad \mathrm{B} = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad \mathrm{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad D = 0$$

Figure 4–3 is a block diagram for the system. Notice that the outputs of the integrators are state variables.

4-1-1 Correlation Between Transfer Functions and State-Space Equations

We shall show how to derive the transfer function of a single-input, single-output system from the state-space equations.

Let us consider the system whose transfer function is given by

$$\frac{Y(s)}{U(s)} = G(s) \qquad (4-15)$$

This system may be represented in state space by the following equations:

$$\dot{x} = Ax + Bu \qquad (4-16)$$

$$y = Cx + Du \qquad (4-17)$$

where x is the state vector, u is the input, and y is the output. The Laplace transforms of Equations (4–16) and (4–17) are given by

$$sX(s) - x(0) = AX(s) + BU(s) \qquad (4-18)$$

$$Y(s) = CX(s) + DU(s) \qquad (4-19)$$

Since the transfer function was previously defined as the ratio of the Laplace transform of the output to the Laplace transform of the input when the initial conditions were zero, we set $x(0)$ in Equation (4–18) to be zero. Then we have

$$sX(s) - AX(s) = BU(s)$$

or

$$(sI - A)X(s) = BU(s)$$

By premultiplying $(sI - A)^{-1}$ to both sides of this last equation, we obtain

$$X(s) = (sI - A)^{-1}BU(s) \qquad (4-20)$$

By substituting Equation (4–20) into Equation (4–19), we get

$$Y(s) = [C(sI - A)^{-1}B + D]U(s) \qquad (4-21)$$

Upon comparing Equation (4–21) with Equation (4–15), we see that

$$G(s) = C(sI - A)^{-1}B + D \qquad (4-22)$$

This is the transfer-function expression of the system in terms of A, B, C, and D.

Note that the right-hand side of Equation (4–22) involves.$(sI - A)^{-1}$ Hence $G(s)$ can be written as

$$G(s) = \frac{Q(s)}{|sI - A|}$$

where $Q(s)$ is a polynomial in s. Notice that $|sI - A|$ is equal to the characteristic polynomial of $G(s)$. In other words, the eigenvalues of A are identical to the poles of $G(s)$.

Example 4–2: [6] Again the mechanical system shown in Figure 4–2. State-space equations for the system are given by Equations (4–13) and (4–14). We shall obtain the transfer function for the system from the state-space equations.

By substituting A, B, C, and D into Equation (4–22), we obtain

$$G(s) = C(sI - A)^{-1}B + D$$

$$= [1 \quad 0]\left\{\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}\right\}^{-1} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} + 0$$

$$= [1 \quad 0]\begin{bmatrix} s & -1 \\ \frac{k}{m} & s+\frac{b}{m} \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}$$

Note that

$$\begin{bmatrix} s & -1 \\ \frac{k}{m} & s+\frac{b}{m} \end{bmatrix}^{-1} = \frac{1}{s^2+\frac{b}{m}s+\frac{k}{m}}\begin{bmatrix} s+\frac{b}{m} & 1 \\ -\frac{k}{m} & s \end{bmatrix}$$

Thus, we have

$$G(s) = [1 \quad 0]\frac{1}{s^2+\frac{b}{m}s+\frac{k}{m}}\begin{bmatrix} s+\frac{b}{m} & 1 \\ -\frac{k}{m} & s \end{bmatrix}\begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}$$

$$= \frac{1}{ms^2+bs+k}$$

which is the transfer function of the system. The same transfer function can be obtained from Equation (4–9).

Transfer Matrix. Next, consider a multiple-input, multiple-output system. Assume that there are r inputs $u_1, u_2, \ldots, u_r$, and m outputs $y_1, y_2, \ldots, y_m$. Define

$$\mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix}, \qquad \mathrm{u} = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_r \end{bmatrix}$$

The transfer matrix $G(s)$ relates the output $Y(s)$ to the input $U(s)$, or

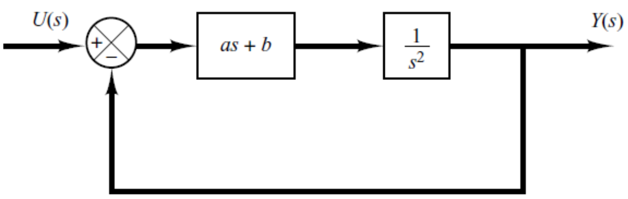
$$Y(s) = G(s)U(s)$$

where $G(s)$ is given by

$$G(s) = C(sI - A)^{-1}B + D$$

[The derivation for this equation is the same as that for Equation (4–22).] Since the input vector u is r dimensional and the output vector y is m dimensional, the transfer matrix G(s) is an $m \times r$ matrix.

Example 4-3: [6] Obtain a state-space model for the system shown in Figure 4–4(a).



(a)

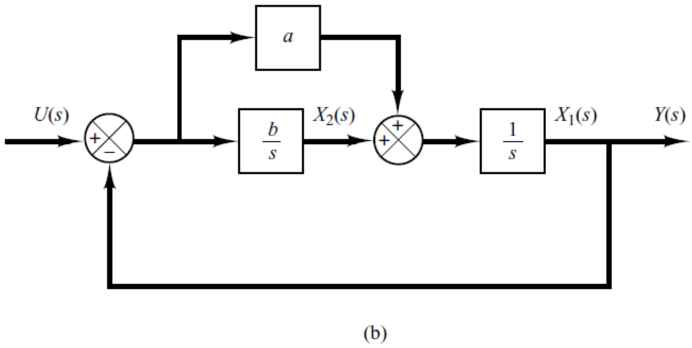


Figure 4–4 (a) Control system; (b) modified block diagram.

Solution. First, notice that $(as+b)/s^2$ involves a derivative term. Such a derivative term may be avoided if we modify $(as+b)/s^2$ as

$$\frac{as+b}{s^2}=\left(a+\frac{b}{s}\right)\frac{1}{s}$$

Using this modification, the block diagram of Figure 4–4(a) can be modified to that shown in Figure 4–4(b).

Define the outputs of the integrators as state variables, as shown in Figure 4–4(b).Then from Figure 4–4(b) we obtain

$$\frac{X_1(s)}{X_2(s)+a[U(s)-X_1(s)]}=\frac{1}{s}$$

$$\frac{X_2(s)}{U(s)-X_1(s)}=\frac{b}{s}$$

$$Y(s)=X_1(s)$$

which may be modified to

$$sX_1(s)=X_2(s)+a[U(s)-X_1(s)]$$

$$sX_2(s)=-bX_1(s)+bU(s)$$

$$Y(s)=X_1(s)$$

Taking the inverse Laplace transforms of the preceding three equations, we obtain

$$\dot{x}_1=-ax_1+x_2+au$$

$$\dot{x}_2=-bx_1+bu$$

$$y=x_1$$

Rewriting the state and output equations in the standard vector-matrix form, we obtain

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}-a & 1\\ -b & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}a\\ b\end{bmatrix}u$$

$$y=\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

4-2 Block Diagrams, Transfer Functions, and State Diagrams

4-2-1 Transfer Functions (Multivariable Systems)

The definition of a transfer function is easily extended to a system with multiple inputs and outputs. A system of this type is often referred to as a multivariable system. In a multivariable system, a differential equation of the form of Eq. (2-23) may be used to describe the relationship between a pair of input and output variables, when all other inputs are set to zero. This equation is restated as

$$y^n+a_{n-1}y^{n-1}+\cdots+a_1\dot{y}+a_0y=b_mu^m+b_{m-1}u^{m-1}+\cdots+b_1\dot{u}+b_0u \quad n\geq m \qquad (4-23)$$

The coefficients $a_0, a_1, \ldots, a_{n-1}$ and $b_0, b_1, \ldots, b_m$ are real constants. Because the principle of superposition is valid for linear systems, the total effect on any output due to all the inputs acting simultaneously is obtained by adding up the outputs due to each input acting alone.

In general, if a linear system has p inputs and q outputs, the transfer function between the jth input and the ith output is defined as

$$G_{ij}(t) = \frac{Y_i(s)}{R_j(s)} \qquad (4-24)$$

with $R_k(s) = 0, k = 1,2,\ldots,p;, k \neq j$. Note that Eq. (4-24) is defined with only the jth input in effect, whereas the other inputs are set to zero. When all the p inputs are in action, the ith output transform is written

$$Y_i(s) = G_{i1}(s)R_1(s) + G_{i2}(s)R_2(s) + \cdots + G_{ip}(s)R_p(s) \qquad (4-25)$$

It is convenient to express Eq. (4-25) in matrix-vector form:

$$Y(s) = G(s)R(s) \qquad (4-26)$$

where

$$Y(s) = \begin{bmatrix} Y_1(s) \\ Y_2(s) \\ \vdots \\ Y_q(s) \end{bmatrix} \qquad (4-27)$$

is the $q \times 1$ transformed output vector,

$$R(s) = \begin{bmatrix} R_1(s) \\ R_2(s) \\ \vdots \\ R_p(s) \end{bmatrix} \qquad (4-28)$$

is the px1 transformed input vector, and

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1p}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2p}(s) \\ \vdots & \vdots & \vdots & \vdots \\ G_{q1}(s) & G_{q2}(s) & \cdots & G_{qp}(s) \end{bmatrix} \qquad (4-29)$$

is the $q \times p$ transfer-function matrix.

4-2-2 Block Diagrams and Transfer Functions of Multivariable Systems

We shall illustrate the block diagram and matrix representations of multivariable systems. Two block-diagram representations of a multivariable system with p inputs and q outputs are shown in Fig. 4-5(a) and (b). In Fig. 4-5(a), the

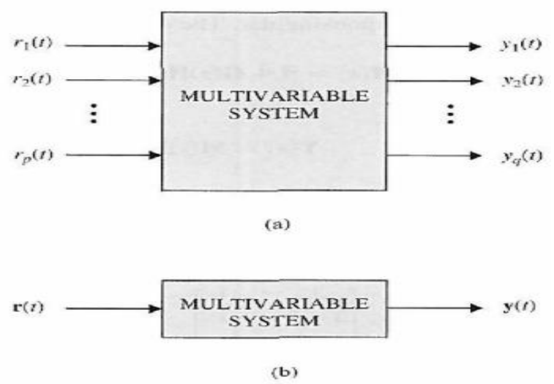


Figure 4-5 Block diagram representations of a multivariable system.

individual input and output signals are designated, whereas in the block diagram of Fig. 4-5(b), the multiplicity of the inputs and outputs is denoted by vectors. The case of Fig. 4-5(b) is preferable in practice because of its simplicity.

Fig. 4-6 shows the block diagram of a multivariable feedback control system. The transfer function relationships of the system are expressed in vector-matrix form:

$$Y(s) = G(s)U(s) \qquad (4-30)$$

$$U(s) = R(s) - B(s) \qquad (4-31)$$

$$B(s) = H(s)Y(s) \qquad (4-32)$$

where *Y(s)* is the $q \times 1$ output vector; U(s), R(s), and B(s) are all $p \times 1$ vectors; and G(s) and H(s) are $q \times p$ and $p \times q$ transfer-function matrices, respectively. Substituting Eq. (4-31) into Eq. (4-30) and then from Eq. (4-30) to Eq. (4-32), we get

$$Y(s) = G(s)R(s) - G(s)H(s)Y(s) \qquad (4-33)$$

Solving for $Y(s)$ from Eq. (4-33) gives

$$Y(s) = [I + G(s)H(s)]^{-1}G(s)R(s) \qquad (4-34)$$

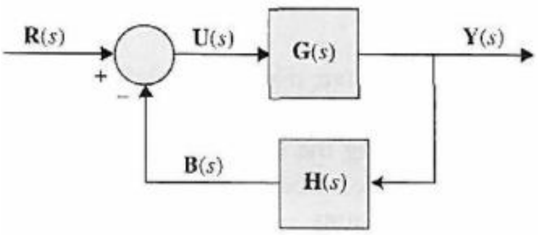


Figure 4-6 Block diagram of a multivariable feedback control system.

provided that $I + G(S)H(s)$ is nonsingular. The closed-loop transfer matrix is defined as

$$M(s) = [I + G(s)H(s)]^{-1}G(s) \qquad (4-35)$$

Then Eq. (4-34) is written

$$Y(s) = M(s)R(s) \qquad (4-36)$$

Example 4-4: [59] Consider that the forward-path transfer function matrix and the feedback-path transfer function matrix of the system shown in Fig. 4-6 are

$$G(s) = \begin{bmatrix} \frac{1}{s+1} & -\frac{1}{s} \\ 2 & \frac{1}{s+2} \end{bmatrix} \quad , H(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (4-37)$$

respectively. The closed-loop transfer function matrix of the system is given by Eq. (4-36) and is evaluated as follows:

$$I + G(s)H(s) = \begin{bmatrix} 1 + \frac{1}{s+1} & -\frac{1}{s} \\ 2 & 1 + \frac{1}{s+2} \end{bmatrix} = \begin{bmatrix} \frac{s+2}{s+1} & -\frac{1}{s} \\ 2 & \frac{s+3}{s+2} \end{bmatrix} \qquad (4-38)$$

The closed-loop transfer function matrix is

$$M(s) = [I + G(s)H(s)]^{-1}G(s) = \frac{1}{\Delta}\begin{bmatrix} \frac{s+3}{s+2} & \frac{1}{s} \\ -2 & \frac{s+2}{s+1} \end{bmatrix}\begin{bmatrix} \frac{1}{s+1} & -\frac{1}{s} \\ 2 & \frac{1}{s+2} \end{bmatrix} \qquad (4-39)$$

where

$$\Delta = \frac{s+2}{s+1}\frac{s+3}{s+2} + \frac{2}{s} = \frac{s^2 + 5s + 2}{s(s+1)} \qquad (4-40)$$

Thus,

$$M(s) = \frac{s(s+1)}{s^2+5s+2}\begin{bmatrix} \frac{3s^2+9s+4}{s(s+1)(s+2)} & -\frac{1}{s} \\ 2 & \frac{3s+2}{s(s+1)} \end{bmatrix} \qquad (4-41)$$

4-2-3 State Diagram

We introduce the state diagram, which is an extension of the SFG to portray state equations and differential equations. The significance of the state diagram is that it forms a close relationship among the state equations, computer simulation, and transfer functions. A state diagram is constructed following all the rules of the SFG using the Laplace-transformed state equations.

The basic elements of a state diagram are similar to the conventional SFG, except for the integration operation. Let the variables $x_1(t)$ and $x_2(t)$ be related by the first-order differentiation:

$$\dot{x}_1 = x_2 \qquad (4-42)$$

Integrating both sides of the last equation with respect to t from the initial time to, we get

$$x_1 = \int_{t_0}^{t} x_2(\tau)d\tau + x_1(t_0) \qquad (4-43)$$

Because the SFG algebra does not handle integration in the time domain, we must take the Laplace transform on both sides of Eq. (4-42). We have

$$X_1(s) = \mathcal{L}\left[\int_{t_0}^{t} x_2(\tau)d\tau\right] + \frac{x_1(t_0)}{s} = \mathcal{L}\left[\int_{0}^{t} x_2(\tau)d\tau - \int_{0}^{t_0} x_2(\tau)d\tau\right] + \frac{x_1(t_0)}{s}$$

$$= \frac{X_2(s)}{s} - \mathcal{L}\left[\int_{0}^{t_0} x_2(\tau)d\tau\right] + \frac{x_1(t_0)}{s} \qquad (4-44)$$

Because the past history of the integrator is represented by $x_1(t_0)$, and the state transition is assumed to start at $\tau = t_0, x_2(\tau) = 0$ for $0 < \tau < t_0$. Thus, Eq. (4-44) becomes

$$X_1(s) = \frac{X_2(s)}{s} + \frac{x_1(t_0)}{s} \quad \tau \geq t_0 \qquad (4-45)$$

Eq. (4-45) is now algebraic and can be represented by an SFG, as shown in Fig. 4-7. Fig. 4-7 shows that the output of the integrator is equal to s^{-1} times the input, plus the initial condition $x_1(t_0)/s$. An alternative SFG with fewer elements for Eq. (4-45) is shown in Fig. 4-8.

Before embarking on several illustrative examples on the construction of state diagrams, let us point out the important uses of the state diagram.

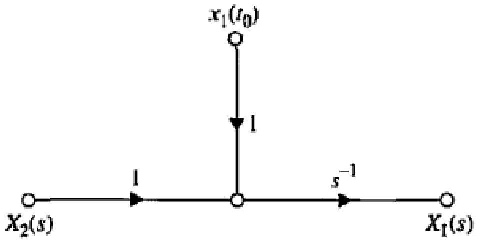


Figure 4-7 Signal-flow graph representation of $X_1(s) = [X_2(s)/s] + [x_1(t_0)/s]$.

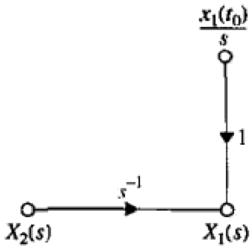


Figure 4-8 Signal-flow graph representation of $X_1(s) = [X_2(s)/s] + [\frac{x_1(t_0)}{s}]$

1. A state diagram can be constructed directly from the system's differential equation. This allows the determination of the state variables and the state equations.
2. A state diagram can be constructed from the system's transfer function. This step is defined as the decomposition of transfer functions (Chapter Five).
3. The state diagram can be used to program the system on an analog computer or for simulation on a digital computer.
4. The state-transition equation in the Laplace transform domain may be obtained from the state diagram by using the SFG gain formula.
5. The transfer functions of a system can be determined from the state diagram.
6. The state equations and the output equations can be determined from the state diagram.

The details of these techniques will follow.

From Differential Equations to State Diagrams. When a linear system is described by a high-order differential equation, a state diagram can be constructed from these equations, although a direct approach is not always the most convenient, Consider the following differential equation:

$$y^n + a_n y^{n-1} + \cdots + a_2 \dot{y} + a_1 y = r(t) \qquad (4-46)$$

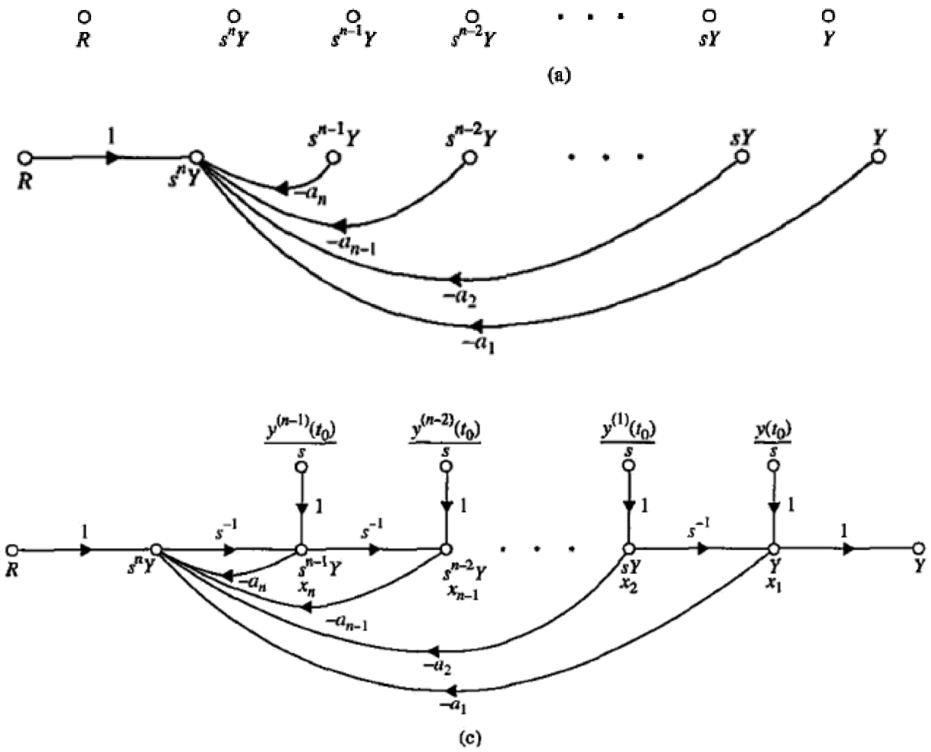


Figure 4-9 State-diagram representation of the differential equation of Eq. (4-46).

To construct a state diagram using this equation, we rearrange the equation as

$$y^n = -a_n y^{n-1} - \cdots - a_2 \dot{y} - a_1 y + r(t) \qquad (4-47)$$

As a first step, the nodes representing $R(s), s^n Y(s), s^{n-1} Y(s), \ldots, sY(s)$, and $Y(s)$ are arranged from left to right, as shown in Fig. 4-9(a). Because $s^i Y(s)$ corresponds to $d^i y(t)/\, dt^i, i = 0,1,2, \ldots, n,$ in the Laplace domain, as the next step, the nodes in Fig. 4-9(a) are connected by branches to portray Eq. (4-47), resulting in Fig. 4-9(b). Finally, the integrator branches with gains of s^{-1} are inserted, and the initial conditions are added to the outputs of the integrators, according to the basic scheme in Fig. 4-7. The complete state diagram is drawn as shown in Fig. 4-9(c). *The outputs of the integrators are defined as the state variables,* $x_1, x_2, \ldots, x_n$. This is usually the natural choice of state variables once the state diagram is drawn.

When the differential equation has derivatives of the input on the right side, the problem of drawing the state diagram directly is not as straightforward as just illustrated. We will show that, in general, it is more convenient to obtain the transfer function from the differential equation first and then arrive at the state diagram through decomposition.

Example 4-5: [59] Consider the differential equation

$$\frac{d^2 y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = r(t) \qquad (4-48)$$

Equating the highest-ordered term of the last equation to the rest of the terms, we have

$$\frac{d^2 y(t)}{dt^2} = -3\frac{dy(t)}{dt} - 2y(t) + r(t) \qquad (4-49)$$

Following the procedure just outlined, the state diagram of the system is drawn as shown in Fig. 4-10. The state variables x_1 and x_2 are assigned as shown.

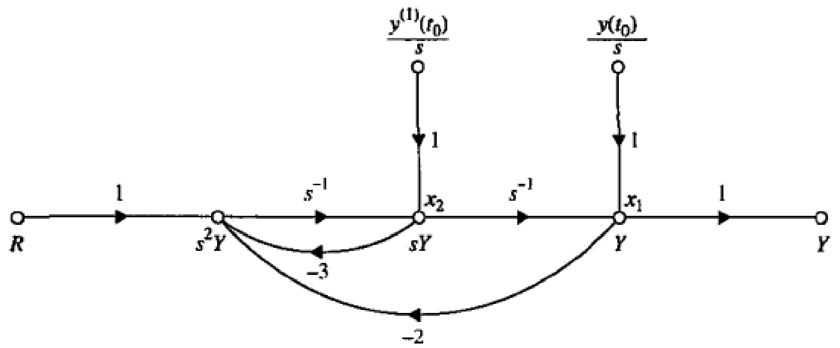


Figure 4-10 State diagram for Eq. (4-48).

From State Diagrams to Transfer Functions. The transfer function between an input and an output is obtained from the state diagram by using the gain formula and setting all other inputs and initial states to zero. The following example shows how the transfer function is obtained directly from a state diagram.

Example 4-6: [59] Consider the state diagram of Fig. 4-10. The transfer function between *R(s)* and *Y(s)* is obtained by applying the gain formula between these two nodes and setting the initial states to zero. We have

$$\frac{Y(s)}{R(s)} = \frac{1}{s^2 + 3s + 2} \qquad (4-50)$$

From State Diagrams to State and Output Equations. The state equations and the output equations can be obtained directly from the state diagram by using the SFG gain formula.

State equation:

$$\dot{x}(t) = Ax(t) + Br(t) \qquad (4-51)$$

Output equation:

$$y(t) = Cx(t) + Dr(t) \qquad (4-52)$$

where $x(t)$ is the state variable; $r(t)$ is the input; $y(t)$ is the output; and A, B, C, and D are constant coefficients. Based on the general form of the state and output equations, the following procedure of deriving the state and output equations from the state diagram are outlined:

1. Delete the initial states and the integrator branches with gains s^{-1} from the state diagram, since the state and output equations do not contain the Laplace operator *s* or the initial states.
2. For the state equations, regard the nodes that represent the derivatives of the state variables as output nodes, since these variables appear on the left-hand side of the state equations. The output $y(t)$ in the output equation is naturally an output node variable.
3. Regard the state variables and the inputs as input variables on the state diagram, since these variables are found on the right-hand side of the state and output equations.
4. Apply the SFG gain formula to the state diagram.

Example 4-7: [59] Fig. 4-11 shows the state diagram of Fig. 4-10 with the integrator branches and the initial states eliminated. Using $dx_1(t)/dt$ and $dx_2(t)/dt$ as the output nodes and $x_1(t), x_2(t)$ and $r(t)$ as input nodes, and applying the gain formula between these nodes, the state equations are obtained as

$$\dot{x}_1 = x_2 \qquad (4-53)$$

$$\dot{x}_2 = -2x_1(t) - 3x_2(t) + r(t) \qquad (4-54)$$

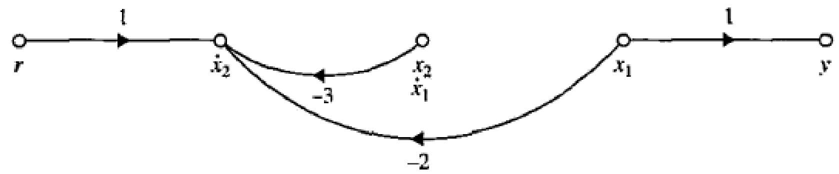


Figure 4-11 State diagram of Fig. 4-10 with the initial states and the integrator branches left out.

Applying the gain formula with $x_1(t), x_2(t)$, and $r(t)$ as input nodes and $y(t)$ as the output node, the output equation is written

$$y(t) = x_1(t) \qquad (4-55)$$

Example 4-8: [59] As another example on the determination of the state equations from the state diagram, consider the state diagram shown in Fig. 4-12(a). This example will also emphasize the importance of applying the gain formula. Fig. 4-12(b) shows the state diagram with the initial states and the integrator branches deleted.

Notice that, in this case, the state diagram in Fig. 4-12(b) still contains a loop. By applying the gain formula to the state diagram in Fig. 10-8(b) with $\dot{x}_1(t), \dot{x}_2(t)$, and $\dot{x}_3(t)$ as output-node variables and $r(t), x_1(t), x_2(t)$, and $x_3(t)$ as input nodes, the state equations are obtained as follows in vector-matrix form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{-(a_2+a_3)}{1+a_0a_3} & -a_1 & \frac{1-a_0a_2}{1+a_0a_3} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} r(t) \qquad (4-56)$$

The output equation is

$$y(t) = \frac{1}{1+a_0a_3} x_1(t) + \frac{a_0}{1+a_0a_3} x_3(t) \qquad (4-57)$$

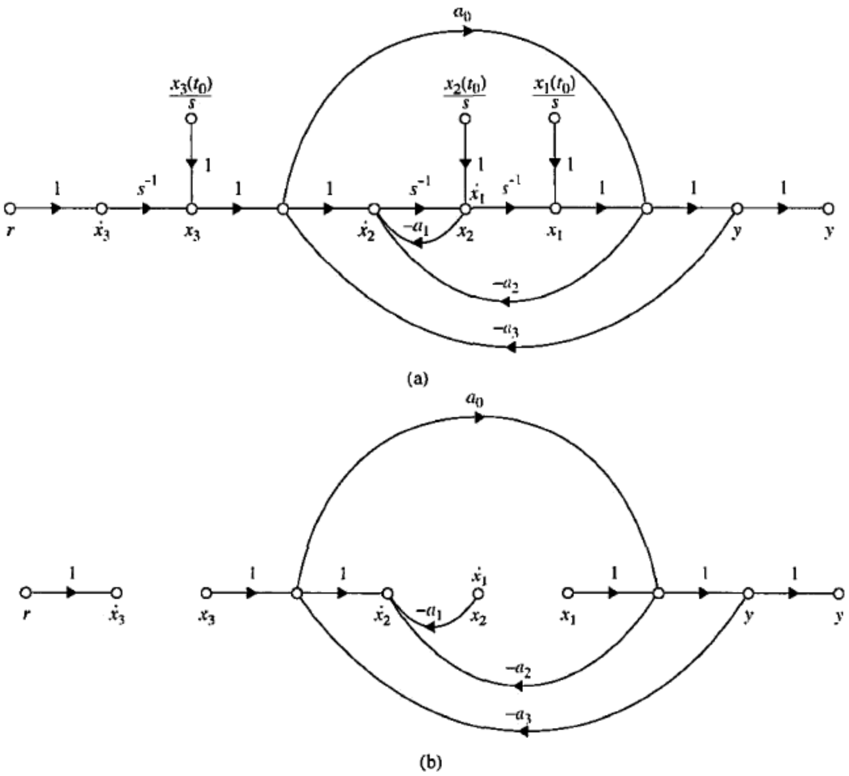


Figure 4-12 (a) State diagram, (b) State diagram in part (a) with all initial states and integrators left out.

4-3 Vector-Matrix Representation of State Equations

Let the *n* state equations of an nth-order dynamic system be represented as

$$\frac{dx_1(t)}{dt} = f_i[x_1(t), \dots, x_n(t), u_1(t), \dots, u_p(t), w_1(t), \dots, w_v(t)] \quad (4-58)$$

where $i = 1,2,\dots,n$. The *i*th state variable is represented by $x_i(t); u_j(t)$ denotes the *j*th input for $j = 1,2,\dots,p;$ and $w_k(t)$ denotes the *kth* disturbance input, with $k = 1,2,\dots,v$.

Let the variables $y_1(t), y_2(t), \dots, y_q(t)$ be the q output variables of the system. In general, the output variables are functions of the state variables and the input variables. The **output equations** can be expressed as

$$y_j(t) = g_j[x_1(t), \dots, x_n(t), u_1(t), \dots, u_p(t), w_1(t), \dots, w_v(t)] \quad (4-59)$$

where $j = 1,2,\dots,q$.

The set of *n* state equations in Eq. (4-58) and *q* output equations in Eq. (4-59) together form the **dynamic equations.** For ease of expression and manipulation, it is convenient to represent the dynamic equations in vector-matrix form. Let us define the following vectors:

State vector:

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}_{(n\times 1)} \tag{4-60}$$

Input vector:

$$u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_p(t) \end{bmatrix}_{(p\times 1)} \tag{4-61}$$

Output vector:

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_q(t) \end{bmatrix}_{(q\times 1)} \tag{4-62}$$

Disturbance vector:

$$w(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ w_v(t) \end{bmatrix}_{(v\times 1)} \tag{4-63}$$

By using these vectors, the *n* state equations of Eq. (4-58) can be written

$$\frac{dx(t)}{dt} = f[x(t), u(t), w(t)] \tag{4-64}$$

where *f* denotes an $n \times 1$ column matrix that contains the functions $f_1, f_2, \ldots, f_n$ as elements. Similarly, the *q* output equations in Eq. (4-59) become

$$y(t) = \mathrm{g}[x(t), u(t), w(t)] \tag{4-65}$$

where g denotes a $q \times 1$ column matrix that contains the functions $\mathrm{g}_1, \mathrm{g}_2, \ldots, \mathrm{g}_q$ as elements.

For a linear time-invariant system, the dynamic equations are written as

State equation:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + Ew(t) \tag{4-66}$$

Output equation:

$$y(t) = Cx(t) + Du(t) + Hw(t) \tag{4-67}$$

Where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}_{(n\times n)} \tag{4-68}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}_{(n\times p)} \qquad (4-69)$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \dots & c_{qn} \end{bmatrix}_{(q\times n)} \qquad (4-70)$$

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1p} \\ d_{21} & d_{22} & \cdots & d_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ d_{q1} & d_{q2} & \dots & d_{qp} \end{bmatrix}_{(q\times p)} \qquad (4-71)$$

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1v} \\ e_{21} & e_{22} & \cdots & e_{2v} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1} & e_{n2} & \dots & e_{nv} \end{bmatrix}_{(n\times v)} \qquad (4-72)$$

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1v} \\ h_{21} & h_{22} & \cdots & h_{2v} \\ \vdots & \vdots & \ddots & \vdots \\ h_{q1} & h_{q2} & \dots & h_{qv} \end{bmatrix}_{(q\times v)} \qquad (4-73)$$

4-4 State-Transition Matrix

Once the state equations of a linear time-invariant system are expressed in the form of Eq. (4-66), the next step often involves the solutions of these equations given the initial state vector $x(t_0)$, the input vector $u(t)$, and the disturbance vector $w(t)$, for $t \geq t_0$. The first term on the right-hand side of Eq. (4-66) is known as the homogeneous part of the state equation, and the last two terms represent the forcing functions $u(t)$ and $w(t)$**.**

The **state-transition matrix** is defined as a matrix that satisfies the linear homogeneous state equation:

$$\frac{dx(t)}{dt} = Ax(t) \qquad (4-74)$$

Let $\emptyset(t)$ be the $n \times n$ matrix that represents the state-transition matrix; then it must satisfy the equation

$$\frac{d\emptyset(t)}{dt} = A\emptyset(t) \qquad (4-75)$$

Furthermore, let $x(0)$ denote the initial state at $t = 0$; then $\emptyset(t)$ is also defined by the matrix equation

$$x(t) = \emptyset(t)x(0) \qquad (4-76)$$

which is the solution of the homogeneous state equation for $t \geq 0$.

One way of determining $\emptyset(t)$ is by taking the Laplace transform on both sides of Eq. (4-74); we have

$$sX(s) - X(0) = AX(s) \qquad (4-77)$$

Solving for $X(s)$ from Eq. (4-77), we get

$$X(s) = (sI - A)^{-1}X(0) \quad (4-78)$$

where it is assumed that the matrix $(sI - A)$ is nonsingular. Taking the inverse Laplace transform on both sides of Eq. (4-78) yields

$$x(t) = \mathcal{L}^{-1}[(sI - A)^{-1}]x(0) \quad t \geq 0 \quad (4-79)$$

By comparing Eq. (4-76) with Eq. (4-79), the state-transition matrix is identified to be

$$\emptyset(t) = \mathcal{L}^{-1}[(sI - A)^{-1}] \quad (4-80)$$

An alternative way of solving the homogeneous state equation is to assume a solution, as in the classical method of solving linear differential equations. We let the solution to Eq. (4-74) be

$$x(t) = e^{At}x(0) \quad (4-81)$$

for $t \geq 0$, where e^{At} represents the following power series of the matrix At, and

$$e^{At} = I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \quad (4-82)$$

It is easy to show that Eq. (4-81) is a solution of the homogeneous state equation, since, from Eq. (4-82),

$$\frac{de^{At}}{dt} = Ae^{At} \quad (4-83)$$

Therefore, in addition to Eq. (4-80), we have obtained another expression for the state-transition matrix:

$$\emptyset(t) = e^{At} = I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \quad (4-84)$$

Eq. (4-84) can also be obtained directly from Eq. (4-80).

Significance of the State-Transition Matrix. Because the state-transition matrix satisfies the homogeneous state equation, it represents the **free response** of the system. In other words, it governs the response that is excited by the initial conditions only. In view of Eqs. (4-80) and (4-84), the state-transition matrix is dependent only upon the matrix A and, therefore, is sometimes referred to as the **statetransition matrix of A.** As the name implies, the state-transition matrix $\emptyset(t)$ completely defines the transition of the states from the initial time $t = 0$ to any time *t* when the inputs are zero.

Properties of the State-Transition Matrix. The state-transition matrix $\emptyset(t)$ possesses the following properties:

1. $\emptyset(t) = I \quad$ (the identity matrix) $\quad (4-85)$

Proof: Eq. (4-85) follows directly from Eq. (4-84) by setting $t = 0$.

2. $\emptyset^{-1}(t) = \emptyset(-t) \quad (4-86)$

Proof: Post-multiplying both sides of Eq. (4-84) by e^{-At}, we get

$$\emptyset(t)e^{-At} = e^{At}e^{-At} = I \quad (4-87)$$

Then, pre-multiplying both sides of Eq. (4-87) by $\emptyset^{-1}(t)$, we get

$$e^{-At} = \emptyset^{-1}(t) \quad (4-88)$$

Thus,

$$\emptyset(-t) = \emptyset^{-1}(t) = e^{-At} \quad (4-89)$$

An interesting result from this property of $\emptyset(t)$ is that Eq. (4-81) can be rearranged to read

$$x(0) = \emptyset(-t)x(t) \quad (4-90)$$

which means that the state-transition process can be considered as bilateral in time. That is, the transition in time can take place in either direction.

3. $\emptyset(t_2 - t_1)\emptyset(t_1 - t_0) = \emptyset(t_2 - t_0)$ for any $t_0, t_1 t_2$ $\quad (4-91)$

Proof:

$$\emptyset(t_2 - t_1)\emptyset(t_1 - t_0) = e^{A(t_2-t_1)}e^{A(t_1-t_0)} = e^{A(t_2-t_0)}$$
$$= \emptyset(t_2 - t_0) \qquad (4-92)$$

This property of the state-transition matrix is important because it implies that a state-transition process can be divided into a number of sequential transitions. Fig. 4-13 illustrates that the transition from $t = t_0$ to $t = t_2$ is equal to the transition from t_0 to t_1 and then from t_1 to t_2. In general, the state-transition process can be divided into any number of parts.

4. $[\emptyset(t)]^k = \emptyset(kt)$ for $k =$ positive integer $\quad (4-93)$

Proof:

$$[\emptyset(t)]^k = e^{At}e^{At} \ldots e^{At} = e^{kAt} = \emptyset(kt) \qquad (4-94)$$

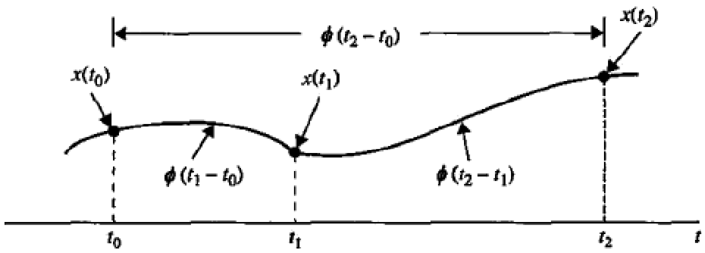


Figure 4-13 Property of the state-transition matrix.

4-5 State-Transition Equation

The state-transition equation is defined as the solution of the linear homogeneous state equation. The linear time-invariant state equation

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + Ew(t) \qquad (4-95)$$

can be solved using either the classical method of solving linear differential equations or the Laplace transform method. The Laplace transform solution is presented in the following equations.

Taking the Laplace transform on both sides of Eq. (4-95), we have

$$sX(s) - X(0) = AX(s) + BU(s) + EW(s) \qquad (4-96)$$

where x(0) denotes the initial-state vector evaluated at $t = 0$. Solving for $X(s)$ in Eq. (4-96) yields

$$X(s) = (sI - A)^{-1}X(0) + (sI - A)^{-1}[BU(s) + EW(s)] \qquad (4-97)$$

The state-transition equation of Eq. (4-95) is obtained by taking the inverse Laplace transform on both sides of Eq. (4-97):

$$x(t) = \mathcal{L}^{-1}[(sI - A)^{-1}]x(0) + \mathcal{L}^{-1}\{(sI - A)^{-1}[BU(s) + EW(s)]\}$$
$$= \emptyset(t)x(0) + \int_0^t \emptyset(t - \tau)[Bu(\tau) + Ew(\tau)]d\tau \quad t \geq 0 \ (4-98)$$

The state-transition equation in Eq. (4-98) is useful only when the initial time is denned to be at $t = 0$. In the study of control systems, especially discrete-data control systems, it is often desirable to break up a state-transition process into a sequence of transitions, so a more flexible initial time must be chosen. Let the initial time be represented by t_0 and the corresponding initial state by $x(t_0)$, and assume that the input $u(t)$ and the disturbance $w(t)$ are applied at $t \geq 0$. We start with Eq. (4-98) by setting $t = t_0$, and solving for $x(0)$, we get

$$x(0) = \emptyset(-t_0)x(t_0) - \emptyset(-t_0)\int_0^{t_0} \emptyset(t_0 - \tau)[Bu(\tau) + Ew(\tau)]d\tau \quad (4-99)$$

where the property on $\emptyset(t)$ of Eq. (4-86) has been applied.

Substituting Eq. (4-99) into Eq. (4-98) yields

$$x(t) = \emptyset(t)\emptyset(-t_0)x(t_0) - \emptyset(t)\emptyset(-t_0)\int_0^{t_0} \emptyset(t_0 - \tau)[Bu(\tau) + Ew(\tau)]d\tau$$

$$+\int_0^{t} \emptyset(t - \tau)[Bu(\tau) + Ew(\tau)]d\tau \quad (4-100)$$

Now by using the property of Eq. (4-91) and combining the last two integrals, Eq. (4-100) becomes

$$x(t) = \emptyset(t - t_0)x(t_0) + \int_0^{t} \emptyset(t - \tau)[Bu(\tau) + Ew(\tau)]d\tau \quad (4-101)$$

It is apparent that Eq. (4-101) reverts to Eq. (4-99) when $t_0 = 0$.

Once the state-transition equation is determined, the output vector can be expressed as a function of the initial state and the input vector simply by substituting $x(t)$ from Eq. (4-101) into Eq. (4-67). Thus, the output vector is

$$y(t) = C\emptyset(t - t_0)x(t_0) + \int_{t_0}^{t} C\emptyset(t - \tau)\,[Bu(\tau) + Ew(\tau)]d\tau$$

$$+Du(t) + Hw(t) \quad t \geq t_0 \quad (4-102)$$

The following example illustrates the determination of the state-transition matrix and equation.

Example 4-9: [59] Consider the state equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (4-103)$$

The problem is to determine the state-transition matrix $\emptyset(t)$ and the state vector $x(t)$ for $t \geq 0$ when the input is $u(t) = 1$ for $t \geq 0$. The coefficient matrices are identified to be

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \quad ,B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad ,E = 0 \quad (4-104)$$

Therefore,

$$sI - A = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix} \quad (4-105)$$

The inverse matrix of $(sI - A)$ is

$$(sI - A)^{-1} = \frac{1}{s^2 + 3s + 2}\begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix} \quad (4-106)$$

The state-transition matrix of A is found by taking the inverse Laplace transform of Eq. (4-106). Thus,

$$\emptyset(t) = \mathcal{L}^{-1}[(sI - A)^{-1}] = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \quad (4-107)$$

The state-transition equation for $t > 0$ is obtained by substituting Eq. (4-107), B, and $u(t)$ into Eq. (4-98). We have

$$x(t) = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} x(0)$$

$$+\int_0^t \begin{bmatrix} 2e^{-(t-\tau)} - e^{-2(t-\tau)} & e^{-(t-\tau)} - e^{-2(t-\tau)} \\ -2e^{-(t-\tau)} + 2e^{-2(t-\tau)} & -e^{-(t-\tau)} + 2e^{-2(t-\tau)} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} d\tau \quad (4-108)$$

or

$$x(t) = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} x(0) + \begin{bmatrix} 0.5 - e^{-t} + 0.5e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix} \quad t \geq 0 \quad (4-109)$$

As an alternative, the second term of the state-transition equation can be obtained by taking the inverse Laplace transform of $(sI - A)^{-1}BU(s)$. Thus, we have

$$\mathcal{L}^{-1}[(sI - A)^{-1}]BU(s) = \mathcal{L}^{-1}\left(\frac{1}{s^2 + 3s + 2}\begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\frac{1}{s}\right)$$

$$= \mathcal{L}^{-1}\left(\frac{1}{s^2 + 3s + 2}\begin{bmatrix} \frac{1}{s} \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0.5 - e^{-t} + 0.5e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix} \quad t \geq 0 \quad (4-110)$$

State-Transition Equation Determined from the State Diagram. Eqs. (4-97) and (4-98) show that the Laplace transform method of solving the state equations requires obtaining the inverse of matrix $(sI - A)$. We shall now show that the state diagram described earlier in Section 4-3-3 and the SFG gain formula (Chapter 3) can be used to solve for the state-transition equation in the Laplace domain of Eq. (4-97). Let the initial time be t_0; then Eq. (4-97) is rewritten as

$$X(s) = (sI - A)^{-1}X(t_0) + (sI - A)^{-1}[BU(s) + EW(s)] \quad t \geq t_0 \quad (4-111)$$

The last equation can be written directly from the state diagram using the gain formula, with $X_i(s), i = 1, 2, \ldots, n$ as the output nodes. The following example illustrates the state-diagram method of finding the state-transition equations for the system described in Example 4-4.

Example 4-10: [59] The state diagram for the system described by Eq. (4-103) is shown in Fig. 4-14 with t_0 as the initial time. The outputs of the integrators are assigned as state variables. Applying the gain formula to the state diagram in Fig. 4-14, with $X_1(s)$ and $X_2(s)$ as output nodes and $x_1(t_0), x_2(t_0)$, and $U(s)$ as input nodes, we have

$$X_1(s) = \frac{s^{-1}(1 + 3s^{-1})}{\Delta}x_1(t_0) + \frac{s^{-2}}{\Delta}x_2(t_0) + \frac{s^{-2}}{\Delta}U(s) \quad (4-112)$$

$$X_2(s) = \frac{-2s^{-2}}{\Delta}x_1(t_0) + \frac{s^{-1}}{\Delta}x_2(t_0) + \frac{s^{-1}}{\Delta}U(s) \quad (4-113)$$

where

$$\Delta = 1 + 3s^{-1} + 2s^{-2} \quad (4-114)$$

After simplification, Eqs. (4-112) and (4-113) are presented in vector-matrix form:

$$\begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix} = \frac{1}{(s+1)(s+2)}\begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix}\begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} + \frac{1}{(s+1)(s+2)}\begin{bmatrix} 1 \\ s \end{bmatrix}U(s) \quad (4-115)$$

The state-transition equation for $t \geq t_0$ is obtained by taking the inverse Laplace transform on both sides of Eq. (4-115).

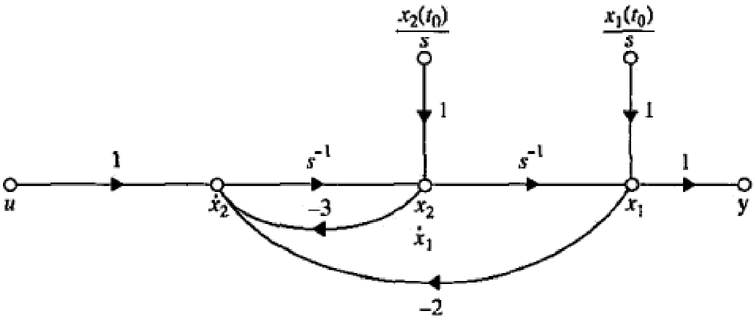


Figure 4-14 State diagram for Eq. (4-103).

Consider that the input $u(t)$ is a unit-step function applied at $t = t_0$ Then the following inverse Laplace transform relationships are identified:

$$\mathcal{L}^{-1}\left(\frac{1}{s}\right) = u_s(t - t_0) \quad t \geq t_0 \qquad (4-116)$$

$$\mathcal{L}^{-1}\left(\frac{1}{s+a}\right) = e^{-a(t-t_0)}u_s(t - t_0) \quad t \geq t_0 \qquad (4-117)$$

Because the initial time is defined to be t_0, the Laplace transform expressions here do not have the delay factor $e^{-t_0 s}$. The inverse Laplace transform of Eq. (10-93) is

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 2e^{-(t-t_0)} - e^{-2(t-t_0)} & e^{-(t-t_0)} - e^{-2(t-t_0)} \\ -2e^{-(t-t_0)} + 2e^{-2(t-t_0)} & -e^{-(t-t_0)} + 2e^{-2(t-t_0)} \end{bmatrix} \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} + \begin{bmatrix} 0.5u_s(t-t_0) - e^{-(t-t_0)} + 0.5e^{-2(t-t_0)} \\ e^{-(t-t_0)} - e^{-2(t-t_0)} \end{bmatrix} \quad t \geq 0 \qquad (4-118)$$

We should compare this result with that in Eq. (4-109), which is obtained for $t \geq 0$.

Example 4-11: [59] In this example, we illustrate the utilization of the state-transition method to a system with input discontinuity. An *RL* network is shown in Fig. 4-15. The history of the network is completely specified by the initial current of the inductance, $i(0)$ at $t = 0$. At time $t = 0$, the voltage $e_{in}(t)$ with the profile shown in Fig. 4-16 is applied to the network. The state equation of the network for $t \geq 0$ is

$$\frac{di(t)}{dt} = -\frac{R}{L}i(t) + \frac{1}{L}e_{in}(t) \qquad (4-119)$$

Comparing the last equation with Eq. (4-66), the scalar coefficients of the state equation are identified to be

$$A = -\frac{R}{L}, \qquad B = \frac{1}{L}, \qquad E = 0 \qquad (4-120)$$

The state-transition matrix is

$$\emptyset(t) = e^{-At} = e^{-Rt/L} \qquad (4-121)$$

The conventional approach of solving for $i(t)$ for $t \geq 0$ is to express the input voltage as

$$e(t) = E_{in}u_s(t) + E_{in}u_s(t - t_1) \qquad (4-122)$$

where $u_s(t)$ is the unit-step function. The Laplace transform of $e(t)$ is

$$E_{in}(s) = \frac{E_{in}}{s}(1 + e^{-t_1 s}) \qquad (4-123)$$

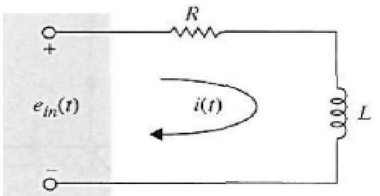


Figure 4-15 *RL* network.

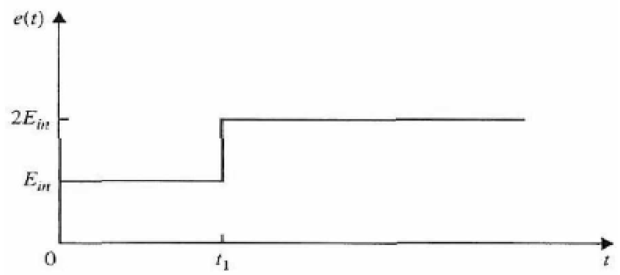


Figure 4-16 Input voltage waveform for the network in Fig. 10-3.

Then

$$(sI-A)^{-1}BU(s)=\frac{E_{in}}{Ls(s+R/L)}(1+e^{-t_1 s}) \qquad (4-124)$$

By substituting Eq. (4-124) into Eq. (4-98), the state-transition equation, the current for $t \geq 0$ is obtained:

$$i(t)=e^{-Rt/L}i(0)+u_s(t)+\frac{E_{in}}{R}\left(1-e^{-Rt/L}\right)u_s(t)$$
$$+\frac{E_{in}}{R}\left(1-e^{-R(t-t_1/L)}\right)u_s(t-t_1) \qquad (4-125)$$

Using the state-transition approach, we can divide the transition period into two parts: $t=0$ to $t=t_1$, and $t=t_1$ to $t=\infty$. First, for the time interval $0 \leq t \leq t_1$ the input is

$$e(t)=E_{in}u_s(t) \quad 0 \leq t \leq t_1 \qquad (4-126)$$

Then

$$(sI-A)^{-1}BU(s)=\frac{E_{in}}{Ls(s+R/L)}=\frac{E_{in}}{Rs(1+(L/R)s)} \qquad (4-127)$$

Thus, the state-transition equation for the time interval $0 \leq t \leq t_1$ is

$$i(t)=\left[e^{-Rt/L}i(0)+\tfrac{E_{in}}{R}(1-e^{-Rt/L})\right]u_s(t) \qquad (4-128)$$

Substituting $t=t_1$ into Eq. (4-128), we get

$$i(t_1)=e^{-Rt_1/L}i(0)+\tfrac{E_{in}}{R}(1-e^{-Rt_1/L}) \qquad (4-129)$$

The value of $i(t)$ at $t=t_1$ is now used as the initial state for the next transition period of $t_1 \leq t < \infty$. The amplitude of the input for the interval is $2E_{in}$. The state-transition equation for the second transition period is

$$i(t)=e^{-R(t-t_1)/L}i(t_1)+\tfrac{2E_{in}}{R}(1-e^{-R(t-t_1)/L}) \; t \geq t_1 \qquad (4-130)$$

where $i(t_1)$ is given by Eq. (4-129).

This example illustrates two possible ways of solving a state-transition problem. In the first approach, the transition is treated as one continuous process, whereas in the second, the transition period is divided into parts over which the input can be more easily presented. Although the first approach requires only one operation, the second method yields relatively simple results to the state-transition equation, and it often presents computational advantages. Notice that, in the second method, the state at $t=t_1$ is used as the initial state for the next transition period, which begins at t_1.

4-6 Relationship between State Equations and High-Order Differential Equation

We defined the state equations and their solutions for linear time-invariant systems. Although it is usually possible to write the state equations directly from the schematic diagram of a system, in practice the system may have been described by a high-order differential equation or transfer function. It becomes necessary to investigate how state equations can be written directly from the high-order differential equation or the transfer function.

The state equations are written in vector-matrix form:

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad (4-131)$$

Where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}_{(n\times n)} \qquad (4-132)$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{(n\times 1)} \qquad (4-133)$$

Notice that the last row of A contains the negative values of the coefficients of the homogeneous part of the differential equation in ascending order, except for the coefficient of the highest-order term, which is unity. B is a column matrix with the last row equal to one, and the rest of the elements are all zeros. The state equations in Eq. (4-131) with A and B given in Eqs. (4-132) and (4-133) are known as the phase-variable **canonical** form (PVCF), or the **controllability** canonical form (CCF).

The output equation of the system is written

$$y(t) = Cx(t) = x_1(t) \qquad (4-134)$$

where

$$C = [1 \quad 0 \quad 0 \quad \cdots \quad 0] \qquad (4-135)$$

We have shown earlier that the state variables of a given system are not unique. In general, we seek the most convenient way of assigning the state variables as long as the definition of state variables is satisfied.

Example 4-12: [59] Consider the differential equation

$$\frac{d^3y(t)}{dt^3} + 5\frac{d^2y(t)}{dt^2} + \frac{dy(t)}{dt} + 2y(t) = u(t) \qquad (4-136)$$

Rearranging the last equation so that the highest-order derivative term is set equal to the rest of the terms, we have

$$\frac{d^3y(t)}{dt^3} = -5\frac{d^2y(t)}{dt^2} - \frac{dy(t)}{dt} - 2y(t) + u(t) \qquad (4-137)$$

The state variables are defined as

$$x_1 = y \quad x_2 = \dot{y} \quad x_3 = \ddot{y} \qquad (4-138)$$

Then the state equations are represented by the vector-matrix equation

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad (4-139)$$

where $x(t)$ is the 2×1 state vector, $u(t)$ is the scalar input, and

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -1 & -5 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad (4-140)$$

The output equation is

$$y(t) = x_1(t) = [1 \quad 0 \quad 0]x(t) \qquad (4-141)$$

4-7 Relationship Between State Equations and Transfer Functions

We have presented the methods of modeling a linear time-invariant system by transfer functions and dynamic equations. We now investigate the relationship between these two representations.

Consider a linear time-invariant system described by the following dynamic equations:

$$\dot{x}(t) = Ax(t) + Bu(t) + Ew(t) \qquad (4-142)$$

$$y(t) = Cx(t) + Du(t) + Hw(t) \qquad (4-143)$$

where

$x(t) = n \times 1$ state vector

$u(t) = p \times 1$ input vector

$\mathrm{y}(t) = q \times 1$ output vector

$\mathrm{w}(t) = v \times 1$ disturbance vector

and A, B, C, D, E, and H are coefficient matrices of appropriate dimensions.

Taking the Laplace transform on both sides of Eq. (4-142) and solving for $X(s)$, we have

$$X(s) = (sI - A)^{-1}x(0) + (sI - A)^{-1}[BU(s) + EW(s)] \quad (4-144)$$

The Laplace transform of Eq. (4-143) is

$$Y(s) = CX(s) + DU(s) + HW(s) \qquad (4-145)$$

Substituting Eq. (4-144) into Eq. (4-145), we have

$$Y(s) = C(sI - A)^{-1}x(0) + C(sI - A)^{-1}[BU(s) + EW(s)] + DU(s) + HW(s) \qquad (4-146)$$

Because the definition of a transfer function requires that the initial conditions be set to zero, $x(0) = 0$; thus, Eq. (4-146) becomes

$$Y(s) = [C(sI - A)^{-1}B + D]U(s) + [C(sI - A)^{-1}E + H]W(s) \qquad (4-147)$$

Let us define

$$G_u(s) = C(sI - A)^{-1}B + D \qquad (4-148)$$

$$G_w(s) = C(sI - A)^{-1}E + H \qquad (4-149)$$

where $G_u(s)$ is a $q \times p$ transfer-function matrix between $u(t)$ and $y(t)$ when $w(t) = 0$, and $G_w(s)$ is a $q \times v$ transfer-function matrix between $w(t)$ and $y(t)$ when $u(t) = 0$.

Then, Eq. (4-147) becomes

$$Y(s) = G_u(s)U(s) + G_w(s)W(s) \qquad (4-150)$$

Example 4-13: [59] Consider that a multivariable system is described by the differential equations

$$\frac{d^2y_1(t)}{dt^2} + 4\frac{dy_1(t)}{dt} - 3y_2(t) = u_1(t) + 2w(t) \qquad (4-151)$$

$$\frac{dy_1(t)}{dt} + \frac{dy_2(t)}{dt} + y_1(t) + 2y_2(t) = u_2(t) \qquad (4-152)$$

The state variables of the system are assigned as:

$$x_1(t) = y_1(t)$$

$$x_2(t) = \dot{y}_1(t) \qquad (4-153)$$

$$x_3(t) = y_2(t)$$

These state variables are defined by mere inspection of the two differential equations, because no particular reasons for the definitions are given other than that these are the most convenient. Now equating the first term of each of the equations of Eqs. (4-151) and (4-152) to the rest of the terms and using the state-variable relations of Eq. (4-153), we arrive at the following state equations and output equations in vector-matrix form:

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \dot{x}_3(t)\end{bmatrix}=\begin{bmatrix}0 & 1 & 0\\ 0 & -4 & 3\\ -1 & -1 & -2\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix}+\begin{bmatrix}0 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}u_1(t)\\ u_2(t)\end{bmatrix}+\begin{bmatrix}0\\ 2\\ 0\end{bmatrix}w(t) \quad (4-154)$$

$$\begin{bmatrix}y_1(t)\\ y_2(t)\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix}=Cx(t) \quad (4-155)$$

To determine the transfer-function matrix of the system using the state-variable formulation, we substitute the A, B, C, D, and E matrices into Eq. (4-151). First, we form the matrix $(sI-A)$:

$$(sI-A)=\begin{bmatrix}0 & -1 & 0\\ 0 & s+4 & -3\\ 1 & 1 & s+2\end{bmatrix} \quad (4-156)$$

The determinant of $(sI-A)$ is

$$|sI-A|=s^3+6s^2+11s+3 \quad (4-157)$$

Thus,

$$(sI-A)^{-1}=\frac{1}{|sI-A|}\begin{bmatrix}s^2+6s+11 & s+2 & 3\\ -3 & s(s+2) & 3s\\ -(s+4) & -(s+1) & s(s+4)\end{bmatrix} \quad (4-158)$$

The transfer-function matrix between $u(t)$ and $y(t)$ is

$$G_u(s)=C(sI-A)^{-1}B$$
$$=\frac{1}{s^3+6s^2+11s+3}\begin{bmatrix}s+2 & 3\\ -(s+1) & s(s+4)\end{bmatrix} \quad (4-159)$$

and that between $w(t)$ and $y(t)$ is

$$G_w(s)=C(sI-A)^{-1}E=\frac{1}{s^3+6s^2+11s+3}\begin{bmatrix}2(s+2)\\ -2(s+1)\end{bmatrix} \quad (4-160)$$

Using the conventional approach, we take the Laplace transform on both sides of Eqs. (4-151) and (4-152) and assume zero initial conditions. The resulting transformed equations are written in vector-matrix form as

$$\begin{bmatrix}s(s+4) & -3\\ s+1 & s+2\end{bmatrix}\begin{bmatrix}Y_1(s)\\ Y_2(s)\end{bmatrix}=\begin{bmatrix}U_1(s)\\ U_2(s)\end{bmatrix}+\begin{bmatrix}2\\ 0\end{bmatrix}W(s) \quad (4-161)$$

Solving for $Y(s)$ from Eq. (4-161), we obtain

$$Y(s)=G_u(s)U(s)+G_w(s)W(s) \quad (4-162)$$

where

$$G_u(s)=\begin{bmatrix}s(s+4) & -3\\ s+1 & s+2\end{bmatrix}^{-1} \quad (4-163)$$

$$G_w(s)=\begin{bmatrix}s(s+4) & -3\\ s+1 & s+2\end{bmatrix}^{-1}\begin{bmatrix}2\\ 0\end{bmatrix} \quad (4-164)$$

which will give the same results as in Eqs. (4-159) and (4-160), respectively, when the matrix inverses are carried out.

CHAPTER FIVE

State Space representations of Dynamic Systems and Linearization of Nonlinear Mathematical Models

A dynamic system consisting of a finite number of lumped elements may be described by ordinary differential equations in which time is the independent variable. By use of vector-matrix notation, an *n*th-order differential equation may be expressed by a first-order vector-matrix differential equation. If *n* elements of the vector are a set of state variables, then the vector-matrix differential equation is a state equation. We shall present methods for obtaining state-space representations of continuous-time systems.

5-1 State Space representations of Dynamic Systems

State Space representations of *n*th- Order Systems of Linear Differential Equations in Which the Forcing Function Does Not Involve Derivative Terms. Consider the following *n*th-order system:

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}\dot{y} + a_n y = u \qquad (5-1)$$

Noting that the knowledge of $y(0), \dot{y}(0), \ldots, y^{(n-1)}(0)$, together with the input $u(t)$ for $t \geq 0$, determines completely the future behavior of the system, we may take $y(t), \dot{y}(t), \ldots, y^{(n-1)}(t)$ as a set of *n* state variables. (Mathematically, such a choice of state variables is quite convenient. Practically, however, because higher-order derivative terms are inaccurate, due to the noise effects inherent in any practical situations, such a choice of the state variables may not be desirable.)

Let us define

$$\begin{aligned} x_1 &= y \\ x_2 &= \dot{y} \\ &\vdots \\ x_n &= y^{(n-1)} \end{aligned} \qquad (5-2)$$

Then Equation (5–1) can be written as

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = x_3$$
$$\vdots \qquad (5-3)$$
$$\dot{x}_{n-1} = x_n$$
$$\dot{x}_n = -a_n x_1 - \cdots - a_1 x_n + u$$

or

$$\dot{\mathrm{x}} = A\mathrm{x} + \mathrm{Bu} \qquad (5-4)$$

Where

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}, \qquad \mathrm{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

The output can be given by

$$y = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}$$

or

$$y = C\mathrm{x} \qquad (5-5)$$

Where

$$C = [1 \quad 0 \quad \cdots \quad 0]$$

[Note that D in Equation (4–17) is zero.] The first-order differential equation, Equation (5–4), is the state equation, and the algebraic equation, Equation (5–5), is the output equation.

Note that the state-space representation for the transfer function system

$$\frac{Y(s)}{U(s)} = \frac{1}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n} \qquad (5-6)$$

is given also by Equations (5–4) and (5–5).

Example 5-1: Derive a state space model for the system shown. The input is f_a and the output is y.

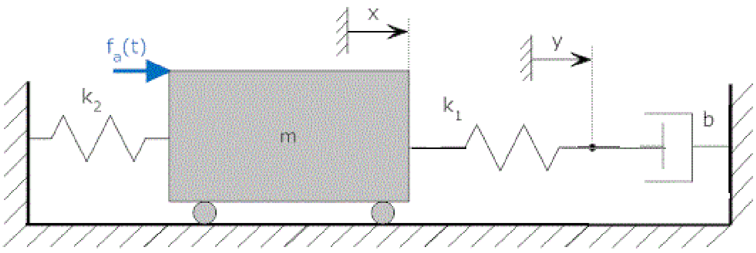
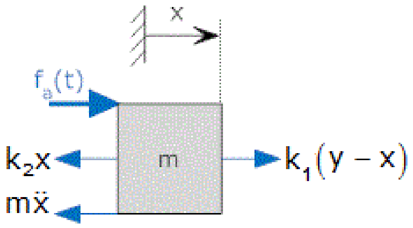


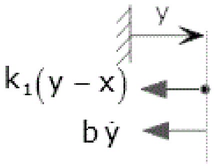
Figure 5-1 Direct Derivation of State Space Model (Mechanical Translating)

Solution. We can write free body equations for the system at x and at y.

Free body Diagram **Equation**



$$m\ddot{x} + k_1 x + k_2 x - k_1 y = \mathrm{f_a}$$



$$b\dot{y} + k_1 y - k_1 x = 0$$

There are three energy storage elements, so we expect three state equations. The energy storage elements are the spring, $\mathrm{k_2}$, the mass, m, and the spring, $\mathrm{k_1}$. Therefore we choose as our state variables x (the energy in spring k_2 is ½k_2x^2), the velocity at x (the energy in the mass m is ½mv^2, where v is the first derivative of x), and y (the energy in spring k_1 is ½$k_1(y-x)^2$, so we could pick $y-x$ as a state variable, but we'll just use y (since x is already a state variable; recall that the choice of state variables is not unique). Our state variables become:

$$q_1 = x$$

$$q_2 = \dot{x}$$

$$q_3 = y$$

Now we want equations for their derivatives. The equations of motion from the free body diagrams yield

$$\dot{q}_1 = \dot{x} = q_2$$

$$\dot{q}_2 = \ddot{x} = \frac{1}{m}(\mathrm{f_a} - k_1 x - k_2 x + k_1 y)$$

$$= \frac{1}{m}(\mathrm{f_a} - k_1 q_1 - k_2 q_1 + k_1 q_3)$$

$$\dot{q}_3 = \dot{y} = \frac{k_1}{b}(x - y) = \frac{k_1}{b}(q_1 - q_3)$$

or

$$\dot{q} = Aq + Bu \qquad A = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{k_1+k_2}{m} & 0 & \frac{k_1}{m} \\ \frac{k_1}{b} & 0 & -\frac{k_1}{b} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \\ 0 \end{bmatrix}$$

$$y = Cq + Du \qquad C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \qquad D = 0$$

with the input $u = \mathrm{f_a}$.

Example 5-2: Derive a state space model for the system shown. The input is i_a and the output is e_2.

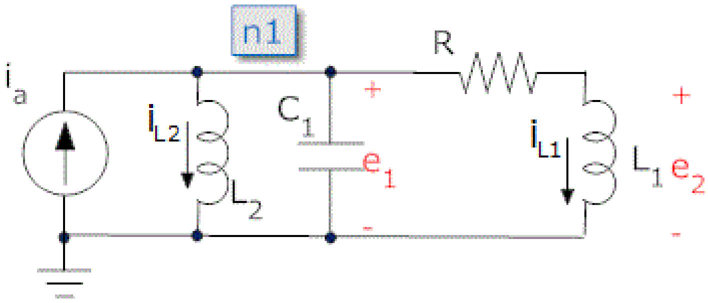


Figure 5-2 Direct Derivation of State Space Model (Electrical)

There are three energy storage elements, so we expect three state equations. Try choosing i_1, i_2 and e_1 as state variables. Now we want equations for their derivatives. The voltage across the inductor L_2 is e_1 (which is one of our state variables)

$$L_2 \frac{di_{L2}}{dt} = e_1$$

so our first state variable equation is

$$\frac{di_{L2}}{dt} = \frac{1}{L_2} e_1$$

If we sum currents into the node labeled n1 we get

$$i_a - i_{L2} - i_{C1} - i_{L1} = 0$$

This equation has our input (i_a) and two state variable (i_{L2} and i_{L1}) and the current through the capacitor. So from this we can get our second state equation

$$i_{C1} = C_1 \frac{de_1}{dt} = i_a - i_{L2} - i_{L1}$$

$$\frac{de_1}{dt} = \frac{1}{C_1}(i_a - i_{L2} - i_{L1})$$

Our third, and final, state equation we get by writing an equation for the voltage across L_1 (which is e_2) in terms of our other state variables

$$e_2 = L_1 \frac{di_{L1}}{dt} = e_1 - Ri_{L1}$$

$$\frac{di_{L1}}{dt} = \frac{1}{L_1}(e_1 - Ri_{L1})$$

We also need an output equation:

$$e_2 = e_1 - Ri_{L1}$$

So our state space representation becomes

$$q = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} i_{L2} \\ e_2 \\ i_{L1} \end{bmatrix}$$

$$\dot{q} = Aq + Bu$$

$$A = \begin{bmatrix} 0 & 1/L_2 & 0 \\ -1/C_1 & 0 & -1/C_1 \\ 0 & 1/L_1 & -R/L_1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1/C_1 \\ 0 \end{bmatrix}$$

$$y = Cq + Du \qquad\qquad C = \begin{bmatrix} 0 & 1 & -R \end{bmatrix}, \qquad D = 0$$

State Space representations of *n*th-Order Systems of Linear Differential Equations in which the Forcing Function Involves Derivative Terms. Consider the differential equation system that involves derivatives of the forcing function, such as

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}\dot{y} + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \cdots + b_{n-1}\dot{u} + b_n u \qquad (5-7)$$

The main problem in defining the state variables for this case lies in the derivative terms of the input *u*. The state variables must be such that they will eliminate the derivatives of *u* in the state equation.

One way to obtain a state equation and output equation for this case is to define the following *n* variables as a set of *n* state variables:

$$x_1 = y - \beta_0 u$$
$$x_2 = \dot{y} - \beta_0\dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u$$
$$x_3 = \ddot{y} - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u$$
$$\vdots \qquad (5-8)$$
$$x_n = y^{(n-1)} - \beta_0 u^{(n-1)} - \beta_1 u^{(n-2)} - \cdots - \beta_{n-2}\dot{u} - \beta_{n-1}u = \dot{x}_{n-1} - \beta_{n-1}u$$

where $\beta_0, \beta_1, \beta_2 \dots, \beta_{n-1}$ are determined from

$$\beta_0 = b_0$$
$$\beta_1 = b_1 - a_1\beta_0$$
$$\beta_2 = b_2 - a_1\beta_1 - a_2\beta_0$$
$$\beta_3 = b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0$$
$$\vdots \qquad (5-9)$$
$$\beta_{n-1} = b_{n-1} - a_1\beta_{n-2} - \cdots - a_{n-2}\beta_1 - a_{n-1}\beta_0$$

With this choice of state variables the existence and uniqueness of the solution of the state equation is guaranteed. (Note that this is not the only choice of a set of state variables.) With the present choice of state variables, we obtain

$$\dot{x}_1 = x_2 + \beta_1 u$$
$$\dot{x}_2 = x_3 + \beta_2 u$$
$$\vdots \qquad (5-10)$$
$$\dot{x}_{n-1} = x_n + \beta_{n-1} u$$
$$\dot{x}_n = -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + \beta_n u$$

Where β_n is given by

$$\beta_n = b_n - a_1\beta_{n-1} - \cdots - a_{n-1}\beta_1 - a_n\beta_0 \qquad (5-11)$$

[To derive Equation (5–10), see Example 5-3] In terms of vector-matrix equations, Equation (5–10) and the output equation can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{n-1} \\ \beta_n \end{bmatrix} u$$

$$y = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \beta_0 u$$

or

$$\dot{x} = Ax + Bu \qquad (5-12)$$
$$y = Cx + Du \qquad (5-13)$$

Where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

$$B = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{n-1} \\ \beta_n \end{bmatrix}, \qquad C = [[1 \quad 0 \quad \cdots \quad 0]], \qquad D = \beta_0 = b_0$$

In this state-space representation, matrices **A** and **C** are exactly the same as those for the system of Equation (5–1).The derivatives on the right-hand side of Equation (5– 7) affect only the elements of the **B** matrix.

Note that the state-space representation for the transfer function

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \cdots + b_n s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n} \tag{5 – 14}$$

is given also by Equations (5–12) and (5–13).

There are many ways to obtain state-space representations of systems. Methods for obtaining canonical representations of systems in state space (such as controllable canonical form, observable canonical form, diagonal canonical form, and Jordan canonical form) are presented.

MATLAB can also be used to obtain state-space representations of systems from transfer-function representations, and vice versa.

Example 5-3: [6] Show that for the differential equation system

$$\dddot{y} + a_1 \ddot{y} + a_2 \dot{y} + a_3 y = b_0 \dddot{u} + b_1 \ddot{u} + b_2 \dot{u} + b_3 u \tag{5 – 15}$$

state and output equations can be given, respectively, by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} u \tag{5 – 16}$$

and

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \beta_0 u \tag{5 – 17}$$

where state variables are defined by

$$x_1 = y - \beta_0 u$$

$$x_2 = \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u$$

$$x_3 = \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u$$

and

$$\beta_0 = b_0$$

$$\beta_1 = b_1 - a_1 \beta_0$$

$$\beta_2 = b_2 - a_1 \beta_1 - a_2 \beta_0$$

$$\beta_3 = b_3 - a_1 \beta_2 - a_2 \beta_1 - a_3 \beta_0$$

Solution. From the definition of state variables x_2 and x_3 , we have

$$\dot{x}_1 = x_2 + \beta_1 u \qquad (5-18)$$

$$\dot{x}_2 = x_3 + \beta_2 u \qquad (5-19)$$

To derive the equation for $\dot{x}_3$ we first note from Equation (5–15) that

$$\dddot{y} = -a_1 \ddot{y} - a_2 \dot{y} - a_3 y + b_0 \dddot{u} + b_1 \ddot{u} + b_2 \dot{u} + b_3 u$$

Since

$$x_3 = \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u$$

We have

$$\dot{x}_3 = \dddot{y} - \beta_0 \dddot{u} - \beta_1 \ddot{u} - \beta_2 \dot{u}$$

$$= (-a_1 \ddot{y} - a_2 \dot{y} - a_3 y) + b_0 \dddot{u} + b_1 \ddot{u} + b_2 \dot{u} + b_3 u - \beta_0 \dddot{u} - \beta_1 \ddot{u} - \beta_2 \dot{u}$$

$$= -a_1(\ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u) - a_1 \beta_0 \ddot{u} - a_1 \beta_1 \dot{u} - a_1 \beta_2 u$$

$$-a_2(\dot{y} - \beta_0 \dot{u} - \beta_1 u) - a_2 \beta_0 \dot{u} - a_2 \beta_1 u - a_3(y - \beta_0 u) - a_3 \beta_0 u$$

$$+b_0 \dddot{u} + b_1 \ddot{u} + b_2 \dot{u} + b_3 u - \beta_0 \dddot{u} - \beta_1 \ddot{u} - \beta_2 \dot{u}$$

$$= -a_1 x_3 - a_2 x_2 - a_3 x_1 + (b_0 - \beta_0) \dddot{u} + (b_1 - \beta_1 - a_1 \beta_0) \ddot{u}$$

$$+(b_2 - \beta_2 - a_1 \beta_1 - a_2 \beta_0) \dot{u} + (b_3 - a_1 \beta_2 - a_2 \beta_1 - a_3 \beta_0) u$$

$$= -a_1 x_3 - a_2 x_2 - a_3 x_1 + (b_3 - a_1 \beta_2 - a_2 \beta_1 - a_3 \beta_0) u$$

$$= -a_1 x_3 - a_2 x_2 - a_3 x_1 + \beta_3 u$$

Hence, we get

$$\dot{x}_3 = -a_3x_1 - a_2x_2 - a_1x_3 + \beta_3u \qquad (5-20)$$

Combining Equations (5–18), (5–19), and (5–20) into a vector-matrix equation, we obtain Equation (5–16). Also, from the definition of state variable x_1 , we get the output equation given by Equation (5–17).

5-2 State Space to Transfer Function

Consider the state space system in Equations (5-12) and (5-13) :

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

$$y(t) = Cx(t) + Du(t)$$

Now, take the Laplace Transform (with zero initial conditions since we are finding a transfer function):

$$sX(s) = AX(s) + BU(s) \qquad (5-21)$$

$$Y(s) = CX(s) + DU(s) \qquad (5-22)$$

We want to solve for the ratio of Y(s) to U(s), so we need so remove X(s) from the output equation (5-22). We start by solving the state equation for X(s)

$$sX(s) - AX(s) = BU(s)$$

$$(sI - A)X(s) = BU(s)$$

$$X(s) = (sI - A)^{-1}BU(s) = \Phi(s)BU(s) \qquad (5-23)$$

where $\quad \Phi(s) = (sI - A)^{-1}$

The matrix $\Phi(s)$ is called the state transition matrix. Now we put Equation (5-23) into the output equation (5-22)

$$Y(s) = [C(sI - A)^{-1}B + D]U(s) = [C\Phi(s)B + D]U(s)$$

Now we can solve for the transfer function:

$$G(s) = \frac{Y(s)}{U(s)} = C\Phi(s)B + D = C(sI - A)^{-1}B + D \qquad (5-24)$$

Note that although there are many state space representations of a given system, all of those representations will result in the same transfer function (i.e., the transfer function of a system is unique; the state space representation is not).

Example: 5-4: Find the transfer function of the system with state space representation

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

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t) \qquad (5-25)$$

$$y(t) = Cx(t) + Du(t)$$

$$= [b_1 \quad b_0 \quad 0]x(t) + 0.u \qquad (5-26)$$

First find $(sI - A)$ and the $\Phi = (sI - A)^{-1}$

$$(sI - A) = \begin{bmatrix} s & -1 & 0 \\ 0 & s & -1 \\ a_3 & a_2 & s + a_1 \end{bmatrix}$$

$$\Phi = (sI - A)^{-1} = \frac{1}{s^3 + a_1 s^2 + a_2 s + a_3} \begin{bmatrix} s^2 + a_1 s + a_2 & a_1 + s & 1 \\ -a_3 & s(a_1 + s) & s \\ -a_3 s & -a_3 - a_2 s & s^2 \end{bmatrix}$$

Now we can find the transfer function

$$G(s) = \frac{Y(s)}{U(s)} = C\Phi(s)B + D$$

$$= \frac{b_0 s + b_1}{s^3 + a_1 s^2 + a_2 s + a_3} \qquad (5-27)$$

5-3 Characteristic Equations, Eigenvalues, and Eigenvectors

Characteristic equations play an important role in the study of linear systems. They can be defined with respect to differential equations, transfer functions, or state equations.

5-3-1 Characteristic Equation from a Differential Equation

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 \dot{y} + a_0 y = b_m u^{(n)} + b_{m-1} u^{(n-1)} + \cdots + b_1 \dot{u} + b_0 u \qquad (5-28)$$

where $n > m$. By defining the operator s as

$$s^k = \frac{d^k}{dt^k} \qquad k = 1,2, \ldots, n \qquad (5-29)$$

Eq. (5-28) is written

$$(s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0) y(t)$$

$$= (b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0) u(t) \qquad (5-30)$$

The **characteristic equation** of the system is defined as

$$s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0 = 0 \qquad (5-31)$$

which is obtained by setting the homogeneous part of Eq. (5-30) to zero.

5-3-2 Characteristic Equation from a Transfer Function

The transfer function of the system described by Eq. (4-165) is

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0} \qquad (5-32)$$

The characteristic equation is obtained by equating the denominator polynomial of the transfer function to zero.

Example 5-5: [59] The transfer function of the system described by the differential equation in Eq. (4-136) is

$$\frac{Y(s)}{G(s)} = \frac{1}{s^3 + 5s^2 + s + 2} \qquad (5-33)$$

The same characteristic equation as in Eq. (4-169) is obtained by setting the denominator polynomial of Eq. (4-171) to zero.

5-3-3 Characteristic Equation from State Equations

From the state-variable approach, we can write Eq. (4-148) as

$$G_u(s) = c\frac{adj(sI - A)}{(sI - A)} B + D$$

$$= \frac{C[adj(sI - A)]B + |sI - A|D}{|sI - A|} \qquad (5-34)$$

Setting the denominator of the transfer-function matrix $G_u(s)$ to zero, we get the characteristic equation

$$|sI - A| = 0 \qquad (5-35)$$

which is an alternative form of the characteristic equation but should lead to the same equation as in Eq. (4-168). An important property of the characteristic equation is that, if the coefficients of A are real, then the coefficients of $|\mathrm{sI} - \mathrm{A}|$ are also real.

EXAMPLE 5-6: [59] The matrix A for the state equations of the differential equation in Eq. (4-136) is given in Eq. (4-150). The characteristic equation of A is

$$|sI - A| = \begin{vmatrix} s & -1 & 0 \\ 0 & s & -1 \\ 2 & 1 & s+5 \end{vmatrix} = s^3 + 5s^2 + s + 2 - 0 \qquad (5-36)$$

5-3-4 Eigenvalues

The roots of the characteristic equation are often referred to as the eigenvalues of the matrix A.

Some of the important properties of eigenvalues are given as follows.

1. If the coefficients of A are all real, then its eigenvalues are either real or in complex-conjugate pairs.
2. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A, then

$$tr(A) = \sum_{i=1}^{n} \lambda_i \qquad (5-37)$$

That is, the trace of A is the sum of all the eigenvalues of A.

3. If $\lambda_i, i = 1, 2, \ldots, n,$ is an eigenvalue of A, then it is an eigenvalue of A'.
4. If A is nonsingular, with eigenvalues $\lambda_i, i = 1, 2, \ldots, n,$ then $1/\lambda_i, i = 1, 2, \ldots, n,$ are the eigenvalues of A^{-1}.

Example 5-7: [59] The eigenvalues or the roots of the characteristic equation of the matrix A in Eq. (4-140) are obtained by solving for the roots of Eq. (5-36). The results are

$$s = -0.06047 + j0.63738 \qquad s = -0.06047 + j0.63738$$
$$s = -4.87906 \qquad (5-38)$$

Invariance of Eigenvalues.

To prove the invariance of the eigenvalues under a linear transformation, we must show that the characteristic polynomials $|\lambda I - A|$ and $|\lambda I - P^{-1}AP|$ are identical.

Since the determinant of a product is the product of the determinants, we obtain

$$|\lambda I - P^{-1}AP| = |\lambda P^{-1}P - P^{-1}AP|$$
$$= |P^{-1}(\lambda I - A)P|$$
$$= |P^{-1}||\lambda I - A||P|$$
$$= |P^{-1}||P||\lambda I - A|$$

Noting that the product of the determinants $|P^{-1}|$ and $|P|$ is the determinant of the product $|P^{-1}|$, we obtain

$$|\lambda I - P^{-1}AP| = |P^{-1}P||\lambda I - A|$$
$$= |\lambda I - A|$$

Thus, we have proved that the eigenvalues of **A** are invariant under a linear transformation.

5-3-5 Eigenvectors

Eigenvectors are useful in modern control methods, one of which is the similarity transformation, which will be discussed in a later section. Any nonzero vector p, that satisfies the matrix equation

$$(\lambda_i - A)P_i = 0 \qquad (5-39)$$

where $\lambda_i, i = 1, 2, \ldots, n$, denotes the ith eigenvalue of A, called the eigenvector of A associated with the eigenvalue λ_i. If A has distinct eigenvalues, the eigenvectors can be solved directly from Eq. (5-39).

Example 5-8: [59] Consider that the state equation of Eq. (4-66) has the coefficient matrices

$$A = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \qquad B = 0 \qquad (5-40)$$

The characteristic equation of A is

$$|sI - A| = s^2 - 1 \qquad (5-41)$$

The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$. Let the eigenvectors be written as

$$p_1 = \begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix}, \qquad p_2 = \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} \qquad (5-42)$$

Substituting $\lambda_1 = 1$ and p_1 into Eq. (4-39), we get

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (5-43)$$

Thus, $p_{21} = 0$, and pu is arbitrary, which in this case can be set equal to 1.

Similarly, for $\lambda_2 = -1.$, Eq. (5-39) becomes

$$\begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (5-44)$$

which leads to

$$-p_{12} + p_{22} = 0 \qquad (5-45)$$

The last equation has two unknowns, which means that one can be set arbitrarily. Let $p_{12} = 1$, then $p_{22} = 2$. The eigenvectors are

$$p_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad p_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad (5-46)$$

5-3-6 Generalized Eigenvectors

It should be pointed out that if A has multiple-order eigenvalues and is nonsymmetric, not all the **eigenvectors** can be found using Eq. (5-39). Let us assume that there are $q(< n)$ distinct eigenvalues among the n eigenvalues of A. The eigenvectors that correspond to the q distinct eigenvalues can be determined in the usual manner from

$$(\lambda_i I - A)p_i = 0 \qquad (5-47)$$

where λ_i, denotes the ith distinct eigenvalue, $i = 1, 2, \ldots, q$. Among the remaining highorder eigenvalues, let λ_i, be of the mth order $(m \leq n - q)$. The corresponding eigenvectors are called the generalized eigenvectors and can be determined from the following m vector equations:

$$\begin{aligned} (\lambda_j I - A)p_{n-q+1} &= 0 \\ (\lambda_j I - A)p_{n-q+2} &= -p_{n-q+1} \\ (\lambda_j I - A)p_{n-q+3} &= -p_{n-q+2} \qquad (5-48) \\ &\vdots \\ (\lambda_j I - A)p_{n-q+m} &= -p_{n-q+m-1} \end{aligned}$$

Example 5-9: [59] Given the matrix

$$A = \begin{bmatrix} 0 & 6 & -5 \\ 1 & 0 & 2 \\ 3 & 2 & 4 \end{bmatrix} \qquad (5-49)$$

The eigenvalues of A are $\lambda_1 = 2, \lambda_2 = \lambda_3 = 1$. Thus, A is a second-order eigenvalue at 1. The eigenvector that is associated with $\lambda_1 = 2$ is determined using Eq. (5-47). Thus,

$$(\lambda_1 I - A)p_1 = \begin{bmatrix} 2 & -6 & 5 \\ -1 & 2 & -2 \\ -3 & -2 & -2 \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \end{bmatrix} = 0 \qquad (5-50)$$

Because there are only two independent equations in Eq. (5-50), we arbitrarily set $p_{11} = 2$, and we have $p_{21} = -1$ and $p_{31} = 2$. Thus,

$$p_1 = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix} \qquad (5-51)$$

For the generalized eigenvectors that are associated with the second-order eigenvalues, we substitute $\lambda_2 = 1$ into the first equation of Eq. (5-48). We have

$$(\lambda_2 I - A)p_2 = \begin{bmatrix} 1 & -6 & 5 \\ -1 & 1 & -2 \\ -3 & -2 & -3 \end{bmatrix} \begin{bmatrix} p_{12} \\ p_{22} \\ p_{32} \end{bmatrix} = 0$$

Setting $p_{12} = 1$ arbitrarily, we have $p_{22} = -\frac{3}{7}$ and $p_{32} = -\frac{5}{7}$ Thus

$$p_2 = \begin{bmatrix} 1 \\ -\frac{3}{7} \\ -\frac{5}{7} \end{bmatrix}$$

Substituting $\lambda_3 = 1$ into the second equation of Eq. (5-48), we have

$$(\lambda_3 I - A)p_3 = \begin{bmatrix} 1 & -6 & 5 \\ -1 & 1 & -2 \\ -3 & -2 & -3 \end{bmatrix} \begin{bmatrix} p_{13} \\ p_{23} \\ p_{33} \end{bmatrix} = -p_2 = \begin{bmatrix} -1 \\ \frac{3}{7} \\ \frac{5}{7} \end{bmatrix}$$

Setting p_{13} arbitrarily to 1, we have the generalized eigenvector

$$p_3 = \begin{bmatrix} 1 \\ -\frac{22}{49} \\ -\frac{46}{49} \end{bmatrix}$$

5-4 Similarity Transformation

The dynamic equations of a single-input, single-output (SISO) system are

$$\dot{x} = Ax + Bu \qquad (5-52)$$

$$y = Cx + Du \qquad (5-53)$$

where $x(t)$ is the *n* x 1 state vector, and $u(t)$ and $y(t)$ are the scalar input and output, respectively. When carrying out analysis and design in the state domain, it is often advantageous to transform these equations into particular forms. For example, as we will show later, the controllability canonical form (CCF) has many interesting properties that make it convenient for controllability tests and state-feedback design.

Let us consider that the dynamic equations of Eqs. (5-52) and (5-53) are transformed into another set of equations of the same dimension by the following transformation:

$$x(t) = p\bar{x}(t) \qquad (5-54)$$

where P is an $n\ x\ n$ nonsingular matrix, so

$$\bar{x}(t) = p^{-1}x(t)$$

The transformed dynamic equations are written

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u \qquad (5-55)$$

$$\bar{y} = \bar{C}\bar{x} + \bar{D}u \qquad (5-56)$$

Taking the derivative on both sides of Eq. (5-69) with respect to *t*, we have

$$\frac{d\bar{x}(t)}{dt} = p^{-1}\frac{dx(t)}{dt} = p^{-1}Ap\bar{x}(t) + p^{-1}Bu(t)$$

$$= p^{-1}A\bar{x}(t) + p^{-1}Bu(t)$$

Comparing Eq. (5-72) with Eq. (5-55), we get

$$\bar{A} = p^{-1}Ap$$

and

$$\bar{B} = p^{-1}B$$

Using Eq. (5-54), Eq. (5-56) is written

$$\bar{y}(t) = Cpx(t) + \bar{D}u(t) \qquad (5-58)$$

Comparing Eq. (5-58) with Eq. (5-53), we see that

$$\bar{C} = Cp \qquad \bar{D} = D \qquad (5-59)$$

The transformation just described is called a similarity transformation, because in the transformed system such properties as the characteristic equation, eigenvectors, eigenvalues, and transfer function are all preserved by the transformation. We shall describe the controllability canonical form (CCF), the observability canonical form (OCF), and the diagonal canonical form (DCF) transformations in the following sections. The transformation equations are given without proofs.

5-4-1 Invariance Properties of the Similarity Transforma-tions

One of the important properties of the similarity transformations is that the characteristic equation, eigenvalues, eigenvectors, and transfer functions are invariant under the transformations.

Characteristic Equations, Eigenvalues, and Eigenvectors

The characteristic equation of the system described by Eq. (5-55) is $|sI — A| = 0$ and is written

$$|sI — \bar{A}| = |sI — P^{-1}AP| = |sP^{-1}P — P^{-1}AP|$$

Because the determinant of a product matrix is equal to the product of the determinants of the matrices, the last equation becomes

$$|sI — \bar{A}| = |P^{-1}||sI - A||P| = |sI - A|$$

Thus, the characteristic equation is preserved, which naturally leads to the same eigenvalues and eigenvectors.

Transfer-Function Matrix. From Eq. (5-56), the transfer-function matrix of the system of Eqs. (5-55) and (5-56) is

$$\bar{G}(s) = \bar{C}(sI - \bar{A})\bar{B} + \bar{D}$$
$$= CP(sI - P^{-1}AP)P^{-1}B + D$$

which is simplified to

$$G(s) = C(sI - A)B + D = G(s)$$

5-4-2 Controllability Canonical Form (CCF)

Consider the dynamic equations given in Eqs. (5-52) and (5-53). The characteristic equation of A is

$$|sI - A| = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 = 0$$

The dynamic equations in Eqs. (5-52) and (5-53) are transformed into CCFof the form of Eqs. (5-55) and (5-56) by the transformation of Eq. (5-54), with

$$P = SM \qquad (5-60)$$

Where

$$S = [B \quad AB \quad A^2B \quad \ldots \quad A^{n-1}B] \qquad (5-61)$$

and

$$M = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & 1 \\ a_2 & a_3 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1} & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \qquad (5-62)$$

Then,

$$\bar{A} = P^{-1}AP = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \qquad (5-63)$$

$$\bar{B} = P^{-1}B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \qquad (5-64)$$

The matrices $\bar{C}$ and $\bar{D}$ are given by Eq. (5-59) and do not follow any particular pattern. The CCF transformation requires that P^{-1} exists, which implies that the matrix S must have an inverse, because the inverse of M always exists because its determinant is $(-1)^{n-1}$, which is nonzero. The n x n matrix S in Eq. (5-51) is later defined as the controllability matrix.

Example 5-10: [59] Consider the coefficient matrices of the state equations in Eq. (5-52):

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

The state equations are to be transformed to CCF.

The characteristic equation of A is

$$|sI - A| = \begin{vmatrix} s-1 & -2 & -1 \\ 0 & s-1 & -3 \\ -1 & -1 & s-1 \end{vmatrix} = s^3 - 3s^2 - s - 3 = 0$$

Thus, the coefficients of the characteristic equation are identified as $a_0 = -3, a_1 = -1$, and $a_2 = -3$. From Eq. (5-62),

$$M = \begin{bmatrix} a_1 & a_2 & 1 \\ a_2 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -3 & 1 \\ -3 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad (5-65)$$

The controllability matrix is

$$S = [B \quad AB \quad A^2B] = \begin{bmatrix} 1 & 2 & 10 \\ 0 & 3 & 9 \\ 1 & 2 & 7 \end{bmatrix}$$

We can show that S is nonsingular, so the system can be transformed into the CCF. Substituting S and M into Eq. (5-60), we get

$$P = SM = \begin{bmatrix} 3 & -1 & 1 \\ 0 & 3 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Thus, from Eqs. (5-63) and (5-64), the CCF model is given by

$$\bar{A} = P^{-1}AP = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 1 & 3 \end{bmatrix}, \qquad \bar{B} = p^{-1}B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

which could have been determined once the coefficients of the characteristic equation are known; however, the exercise is to show how the CCF transformation matrix P is obtained.

5-4-3 Observability Canonical Form (OCF)

A dual form of transformation of the CCF is the observability canonical form (OCF). The system described by Eqs. (5-52) and (5- 53) is transformed to the OCF by the transformation

$$x(t) = Q\bar{x}(t)$$

The transformed equations are as given in Eqs. (5-55) and (5-56). Thus,

$$\bar{A} = Q^{-1}AQ, \qquad \bar{B} = Q^{-1}B, \qquad \bar{C} = CQ, \qquad \bar{D} = D$$

Where

$$\bar{A} = Q^{-1}AQ = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix} \qquad (5-66)$$

$$\bar{C} = CQ = [0 \quad 0 \quad \cdots \quad 0 \quad 1] \qquad (5-67)$$

The elements of the matrices B and D are not restricted to any form. Notice that A and C are the transpose of the A and B in Eqs. (5-63) and (5-64), respectively.

The OCF transformation matrix Q is given by

$$Q = (MV)^{-1} \qquad (5-68)$$

where M is as given in Eq. (5-84), and

$$V = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}_{(n\times n)}$$

The matrix V is often defined as the **observability matrix,** and V^{-1} must exist in order for the OCF transformation to be possible.

Example 5-11: [59] Consider that the coefficient matrices of the system described by Eqs. (5-52) and (4-160) are

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad C = [1 \quad 1 \quad 0]$$

Because the matrix A is identical to that of the system in Example 5-10, the matrix M is the same as that in Eq. (5-65). The observability matrix is

$$V = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix}_{(3\times 3)} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 3 & 4 \\ 5 & 9 & 14 \end{bmatrix}$$

We can show that V is nonsingular, so the system can be transformed into the OCF. Substituting V and M into Eq. (5-68), we have the OCF transformation matrix,

$$Q = (MV)^{-1} = \begin{bmatrix} 0.3333 & -0.1667 & 0.3333 \\ -0.3333 & 0.1667 & 0.6667 \\ 0.1667 & 0.1667 & 0.1667 \end{bmatrix}$$

From Eq. (5-63), the OCF model of the system is described by

$$\bar{A} = Q^{-1}AQ = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 3 \end{bmatrix}, \qquad \bar{C} = CQ = [0 \quad 0 \quad 1],$$

$$\bar{B} = Q^{-1}B = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Thus, $\bar{A}$ and $\bar{C}$ are of the OCF form given in Eqs. (5-66) and (5-67), respectively, and $\bar{B}$ does no conform to any particular form.

5-4-4 Diagonal Canonical Form (DCF)

Given the dynamic equations in Eqs. (5-52) and (5-53), if A has distinct eigenvalues, there is a nonsingular transformation

$$x(t) = T\bar{x}(t)$$

which transforms these equations to the dynamic equations of Eqs. (5-55) and (5-56), where

$$\bar{A} = T^{-1}AT, \qquad \bar{B} = T^{-1}B, \qquad \bar{C} = CT, \qquad \bar{D} = 0 \qquad (5-69)$$

The matrix $\bar{A}$ is a diagonal matrix,

$$\bar{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}_{(n\times n)} \qquad (5-70)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the n distinct eigenvalues of A. The coefficient matrices $\bar{B}$, $\bar{C}$, and $\bar{D}$ are given in Eq. (5-69) and do not follow any particular form.

It is apparent that one of the advantages of the DCF is that the transformed state equations are decoupled from each other and, therefore, can be solved individually.

We show in the following that the DCF transformation matrix T can be formed by use of the eigenvectors of A as its columns; that is,

$$T = [p_1 \quad p_2 \quad p_3 \quad \cdots \quad p_n]$$

where $p_i, i = 1,2,\ldots,n,$ denotes the eigenvector associated with the eigenvalue λ_i. This is proved by use of Eq. (5-39), which is written as

$$\lambda_i p_i = Ap_i, \qquad i = 1,2,\ldots,n$$

Now, forming the $n \ x \ n$ matrix,

$$[\lambda_1 p_1 \quad \lambda_2 p_2 \quad \cdots \quad \lambda_n p_n] = [Ap_1 \quad Ap_2 \quad \cdots \quad Ap_n]$$
$$= A[p_1 \quad p_2 \quad \cdots \quad p_n]$$

The last equation is written

$$[p_1 \quad p_2 \quad \cdots \quad p_n]\bar{A} = A[p_1 \quad p_2 \quad \cdots \quad p_n] \qquad (5-71)$$

where $\bar{A}$ is as given in Eq. (5-70). Thus, if we let

$$T = [p_1 \quad p_2 \quad \cdots \quad p_n]$$

Eq. (5-71) is written

$$\bar{A} = T^{-1}AT \qquad (5-72)$$

If the matrix A is of the CCF and A has distinct eigenvalues, then the DCF transformation matrix is the Vandermonde matrix,

$$T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_3^{n-1} & \cdots & \lambda_1^{n-1} \end{bmatrix} \qquad (5-73)$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A. This can be proven by substituting the CCF of A in Eq. (4-132) into Eq. (5-39). The result is that the ith eigenvector p_i is equal to the ith column of T in Eq.(5-73).

Example 5-12: [59] Consider the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$$

which has eigenvalues $\lambda_1 = -1$, $\lambda_2 = -2$, and $\lambda_3 = -3$. Because A is CCF, to transform it into DCF, the transformation matrix can be the Vandermonde matrix in Eq. (5-73). Thus,

$$T = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix}$$

Thus, the DCF of A is written

$$\bar{A} = T^{-1}AT = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

5-4-5 Jordan Canonical Form (JCF)

In general, when the matrix A has multiple-order eigenvalues, unless the matrix is symmetric with real elements, it cannot be transformed into a diagonal matrix. However, there exists a similarity transformation in the form of Eq. (5-72) such that the matrix $\bar{A}$ is almost diagonal. The matrix $\bar{A}$ is called the Jordan canonical form (JCF). A typical JCF is shown below.

$$\bar{A} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 \end{bmatrix} \qquad (5-74)$$

where it is assumed that A has a third-order eigenvalue λ_1 and distinct eigenvalues λ_2 and λ_3.

The JCF generally has the following properties:

1. The elements on the main diagonal are the eigenvalues.
2. All the elements below the main diagonal are zero.
3. Some of the elements immediately above the multiple-order eigenvalues on the main diagonal are 1s, as shown in Eq. (5-74).
4. The 1s together with the eigenvalues form typical blocks called the Jordan blocks. As shown in Eq. (5-74), the Jordan blocks are enclosed by dashed lines.
5. When the nonsymmetrical matrix A has multiple-order eigenvalues, its eigenvectors are not linearly independent. For an A that is $n \times n$, there are only r (where r is an integer that is less than n and is dependent on the number of multiple-order eigenvalues) linearly independent eigenvectors.

6. The number of Jordan blocks is equal to the number of independent eigenvectors r. There is one and only one linearly independent eigenvector associated with each Jordan block.
7. The number of 1s above the main diagonal is equal to $n - r$.

To perform the JCF transformation, the transformation matrix T is again formed by using the eigenvectors and generalized eigenvectors as its columns.

Example 5-13: [59] Consider the matrix given in Eq. (5-49). We have shown that the matrix has eigenvalues 2,1, and 1. Thus, the DCF transformation matrix can be formed by using the eigenvector and generalized eigenvector. That is,

$$A = [p_1 \quad p_2 \quad p_3] = \begin{bmatrix} 2 & 1 & 1 \\ -1 & -\frac{3}{7} & -\frac{22}{49} \\ -2 & -\frac{5}{7} & -\frac{46}{49} \end{bmatrix}$$

Thus, the DCF is

$$\bar{A} = T^{-1}AT = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that in this case there are two Jordan blocks, and there is one element of 1 above the main diagonal.

5-5 Decompositions of Transfer Function

Up to this point, various methods of characterizing linear systems have been presented. To summarize, it has been shown that the starting point of modeling a linear system may be the system's differential equation, transfer function, or dynamic equations; all these methods are closely related. Furthermore, the state diagram is also a useful tool that can not only lead to the solutions of state equations but also serve as a vehicle of transformation from one form of description to the others. The block diagram of Fig. 5-3 shows the relationships among the various ways of describing a linear system. For example, the block diagram shows that, starting with the differential equation of a system, one can find the solution by the transfer-function or state-equation method. The block diagram also shows that the majority of the relationships are bilateral, so a great deal of flexibility exists between the methods.

One subject remains to be discussed, which involves the construction of the state diagram from the transfer function between the input and the output. The process of going from the transfer function to the state diagram is called **decomposition**. In general, there are three basic ways to decompose transfer functions. These are **direct decomposition, cascade decomposition**, and **parallel decomposition**. Each of these three schemes of decomposition has its own merits and is best suited for a particular purpose.

5-5-1 Direct Decomposition

Direct decomposition is applied to an input-output transfer function that is not in factored form. Consider the transfer function of an *n*th-order SISO system between the input $U(s)$ and output $Y(s)$:

$$\frac{Y(s)}{U(s)} = \frac{b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} \qquad (5-75)$$

where we have assumed that the order of the denominator is at least one degree higher than that of the numerator.

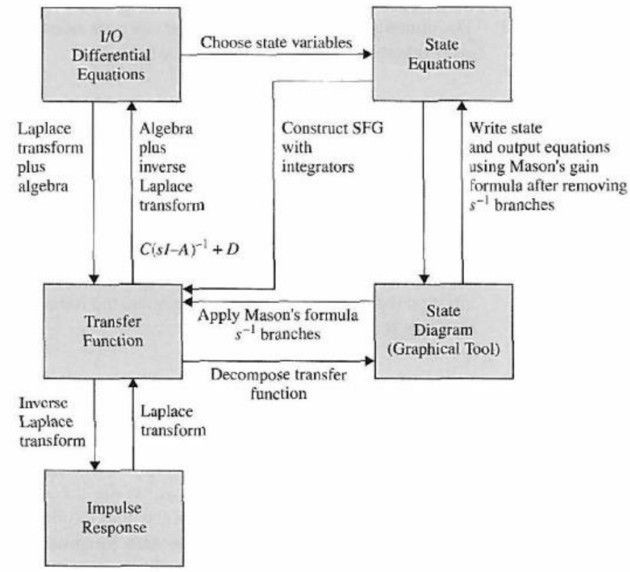


Figure 5-3 Block diagram showing the relationships among various methods of describing linear systems.

We next show that the direct decomposition can be conducted in at least two ways, one leading to a state diagram that corresponds to the CCF and the other to the OCF.

Direct Decomposition to CCF. The objective is to construct a state diagram from the transfer function of Eq. (5-75). The following steps are outlined:

1. Express the transfer function in negative powers of s. This is done by multiplying the numerator and the denominator of the transfer function by s^{-n}.
2. Multiply the numerator and the denominator of the transfer function by a dummy variable $X(s)$. By implementing the last two steps, Eq. (5-75) becomes
$$\frac{Y(s)}{U(s)}=\frac{b_{n-1}s^{-1}+b_{n-2}s^{-2}+\cdots+b_1s^{-n+1}+b_0s^{-n}}{1+a_{n-1}s^{-1}+\cdots+a_1s^{-n+1}+a_0s^{-n}}\frac{X(s)}{X(s)} \quad (5-76)$$
3. The numerators and the denominators on both sides of Eq. (5-76) are equated to each other, respectively. The results are:
$$Y(s)=(b_{n-1}s^{-1}+b_{n-2}s^{-2}+\cdots+b_1s^{-n+1}+b_0s^{-n})X(s) \quad (5-77)$$
$$U(s)=(1+a_{n-1}s^{-1}+\cdots+a_1s^{-n+1}+a_0s^{-n})X(s) \quad (5-78)$$
4. To construct a state diagram using the two equations in Eqs. (5-77) and (5-78), they must first be in the proper cause-and-effect relation. It is apparent that Eq. (5-77) already satisfies this prerequisite. However, Eq. (5-78) has the input on the left-hand side of the equation and must be rearranged. Eq. (5-78) is rearranged as
$$X(s)=U(s)-(a_{n-1}s^{-1}+a_{n-2}s^{-2}\ldots+a_1s^{-n+1}+a_0s^{-n})X(s)$$

The state diagram is drawn as shown in Fig. 5-4 using Eqs. (5-77) and (5-78). For simplicity, the initial states are not drawn on the diagram. The state variables $x_1(t), x_2(t), \ldots, x_n(t)$ are defined as the outputs of the integrators and are arranged in order from the right to the left on the state diagram. The state equations are obtained by applying the SFG gain formula to Fig. 5-4 with the derivatives of the state variables as the outputs and the state variables and $u(t)$ as the inputs, and overlooking the integrator branches. The output equation is determined by applying the gain formula among the state variables, the input, and the output $y(t)$. The dynamic equations are written

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \qquad (5-79)$$

$$y(t) = Cx(t) + Du(t) \qquad (5-80)$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \qquad (5-81)$$

$$C = [b_0 \quad b_1 \quad \cdots \quad b_{n-2} \quad b_{n-1}], \qquad D = 0$$

Apparently, A and B in Eq. (5-81) are of the CCF.

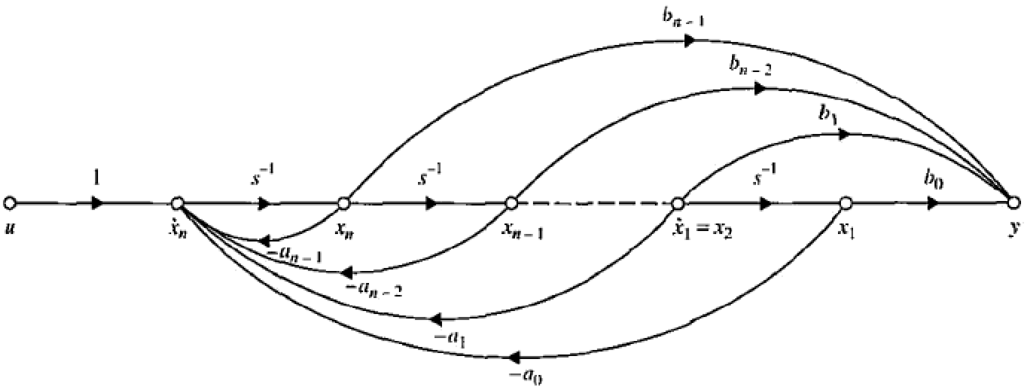


Figure 5-4 CCF state diagram of the transfer function in Eq. (5-75) by direct decomposition.

Direct Decomposition to OCF. Multiplying the numerator and the denominator of Eq. (5-75) by s^{-n}, the equation is expanded as

$$(1 + a_{n-1}s^{-1} + \cdots + a_1 s^{-n+1} + a_0 s^{-n})Y(s) = (b_{n-1}s^{-1} + b_{n-2}s^{-2} + \cdots + b_1 s^{-n+1} + b_0 s^{-n})U(s)$$

or

$$Y(s) = -(a_{n-1}s^{-1} + \cdots + a_1 s^{-n+1} + a_0 s^{-n})Y(s) + (b_{n-1}s^{-1} + b_{n-2}s^{-2} + \cdots + b_1 s^{-n+1} + b_0 s^{-n})U(s) \quad (5-82)$$

Fig. 5-5 shows the state diagram that results from using Eq. (5-82). The outputs of the integrators are designated as the state variables. However, unlike the usual convention, the state variables are assigned in descending order from right to left. Applying the SFG gain formula to the state diagram, the dynamic equations are written as in Eqs. (5-79) and (5-80), with

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}, \qquad B = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{n-1} \end{bmatrix}$$

and

$$C = [0 \quad 0 \quad \cdots \quad 0 \quad 1], \qquad D = 0$$

The matrices A and C are in OCF.

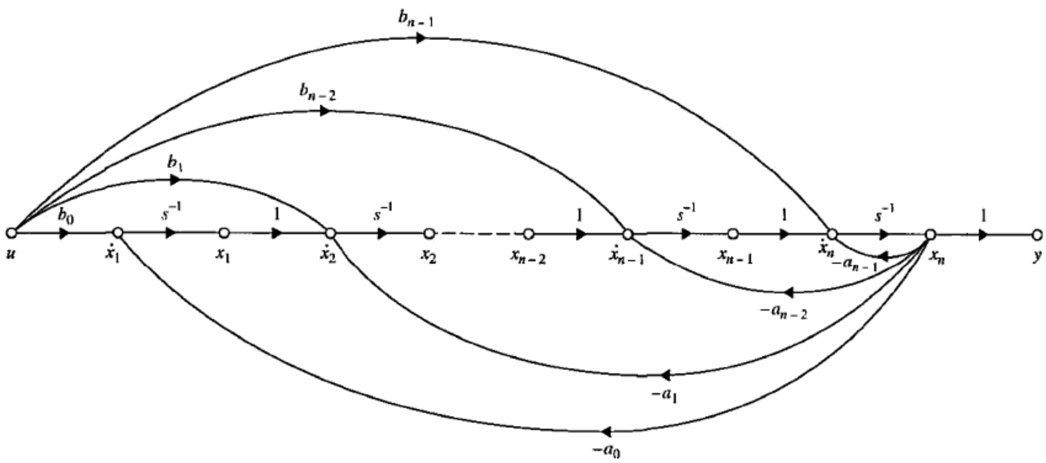


Figure 5-5 CCF state diagram of the transfer function in Eq. (5-75) by direct decomposition.

It should be pointed out that, given the dynamic equations of a system, the input-output transfer function is unique. However, given the transfer function, the state model is not unique, as shown by the CCF, OCF, and DCF, and many other possibilities. In fact, even for any one of these canonical forms (for example, CCF), while matrices A and B are defined, the elements of C and D could still be different depending on how the state diagram is drawn, that is, how the transfer function is decomposed. In other words, referring to Fig. 5-4, whereas the feedback branches are fixed, the feedforward branches that contain the coefficients of the numerator of the transfer function can still be manipulated to change the contents of C.

Example 5-14: [59] Consider the following input-output transfer function:

$$\frac{Y(s)}{U(s)} = \frac{2s^2 + s + 5}{s^3 + 6s^2 + 11s + 4} \qquad (5-83)$$

The CCF state diagram of the system is shown in Fig. 5-6, which is drawn from the following equations:

$$Y(s) = (2s^{-1} + s^{-2} + 5s^{-3})X(s)$$
$$X(s) = U(s) - (6s^{-1} + 11s^{-2} + 4s^{-3})X(s)$$

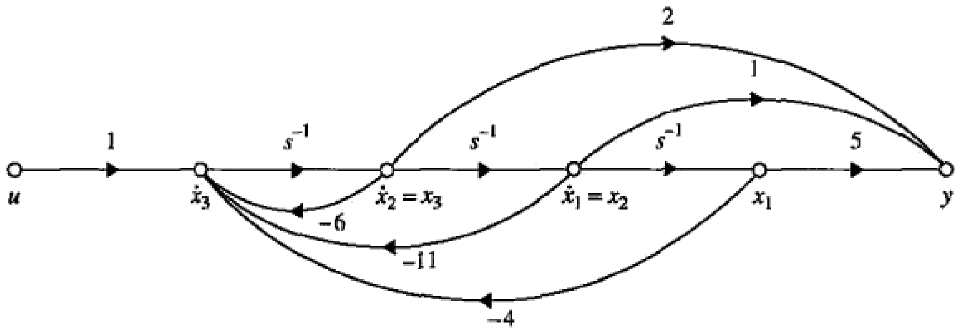


Figure 5-6 CCF state diagram of the transfer function in Eq. (5-83).

The dynamic equations of the system in CCF are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -11 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [5 \quad 1 \quad 2]x$$

For the OCF, Eq. (5-83) is expanded to

$$Y(s) = (2s^{-1} + s^{-2} + 5s^{-3})U(s) - (6s^{-1} + 11s^{-2} + 4s^{-3})Y(s)$$

which leads to the OCF state diagram shown in Fig. 5-7. The OCF dynamic equations are written

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -4 \\ 1 & 0 & -11 \\ 0 & 1 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} u$$

$$y = [0 \quad 0 \quad 1]x$$

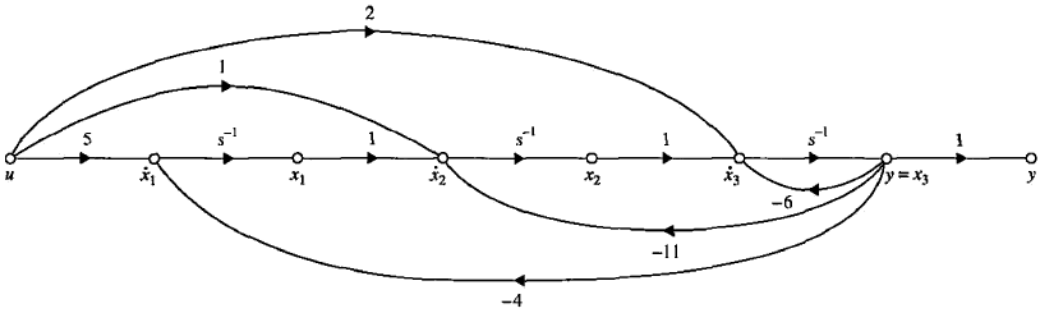


Figure 5-7 OCF state diagram of the transfer function in Eq. (5-83).

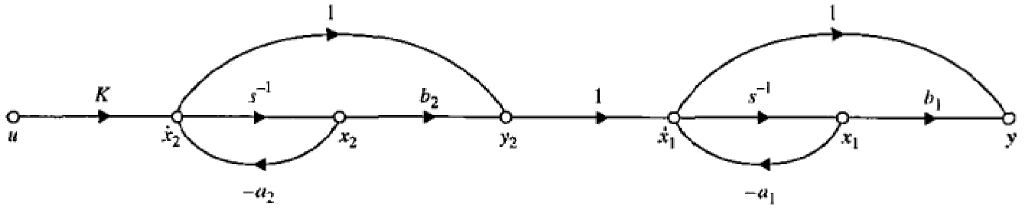


Figure 5-8 State diagram of the transfer function in Eq. (5-84) by cascade decomposition.

5-5-2 Cascade Decomposition

Cascade compensation refers to transfer functions that are written as products of simple first-order or second-order components. Consider the following transfer function, which is the product of two first-order transfer functions.

$$\frac{Y(s)}{U(s)} = K\left(\frac{s+b_1}{s+a_1}\right)\left(\frac{s+b_2}{s+a_2}\right) \qquad (5-84)$$

where a_1, a_2, b_1, and b_2 are real constants. Each of the first-order transfer functions is decomposed by the direct decomposition, and the two state diagrams are connected in cascade, as shown in Fig. 5-8. The state equations are obtained by regarding the derivatives of the state variables as outputs and the state variables and $u(t)$ as inputs and then applying the SFG gain formula to the state diagram in Fig. 5-8. The integrator branches are neglected when applying the gain formula. The results are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -a_1 & b_2 - a_2 \\ 0 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} K \\ K \end{bmatrix} u$$

The output equation is obtained by regarding the state variables and $u(t)$ as inputs and $y(t)$ as the output and applying the gain formula to Fig. 5-8. Thus,

$$y(t) = [b_1 - a_1 \quad b_2 - a_2]x(t) + Ku(t)$$

When the overall transfer function has complex poles or zeros, the individual factors related to these poles or zeros should be in second-order form. As an example, consider the following transfer function:

$$\frac{Y(s)}{U(s)} = K\left(\frac{s+5}{s+2}\right)\left(\frac{s+1.5}{s^2+3s+4}\right) \qquad (5-85)$$

where the poles of the second term are complex. The state diagram of the system with the two subsystems connected in cascade is shown in Fig. 5-9. The dynamic equations of the system are

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}0 & 1 & 0\\ -4 & -3 & 3\\ 0 & 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}0\\ 1\\ 1\end{bmatrix}u$$

$$y(t)=[1.5 \quad 1 \quad 0]x(t)$$

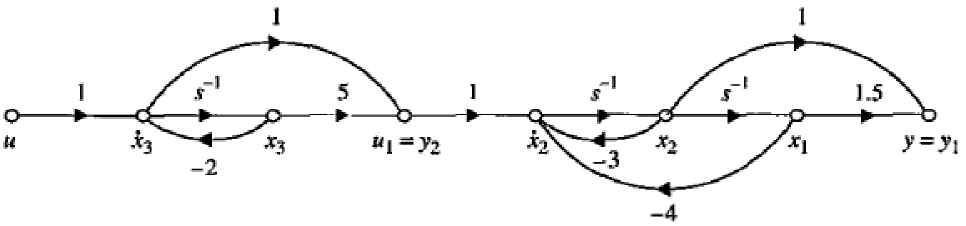


Figure 5-9 State diagram of the transfer function in Eq. (5-85) by cascade decomposition.

5-5-3 Parallel Decomposition

When the denominator of the transfer function is in factored form, the transfer function may be expanded by partial-fraction expansion. The resulting state diagram will consist of simple first- or second-order systems connected in parallel, which leads to the state equations in DCF or JCF, the latter in the case of multiple-order eigenvalues.

Consider that a second-order system is represented by the transfer function

$$\frac{Y(s)}{U(s)}=\frac{Q(s)}{(s+a_1)(s+a_2)} \qquad (5-86)$$

where $Q(s)$ is a polynomial of order less than 2, and a_1 and a_2 are real and distinct. Although, analytically, a_1 and a_2 may be complex, in practice, complex numbers are difficult to implement on a computer. Eq. (5-87) is expansion by partial fractions:

$$\frac{Y(s)}{U(s)}=\frac{K_1}{s+a_1}+\frac{K_2}{s+a_2} \qquad (5-87)$$

where K_1 and K_2 are real constants.

The state diagram of the system is drawn by the parallel combination of the state diagrams of each of the first-order terms in Eq. (5-87), as shown in Fig. 5-10. The dynamic equations of the system are

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}-a_1 & 0\\ 0 & -a_2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}1\\ 1\end{bmatrix}u$$

$$y(t)=[K_1 \quad K_2]x(t)$$

Thus, the state equations are of the DCF.

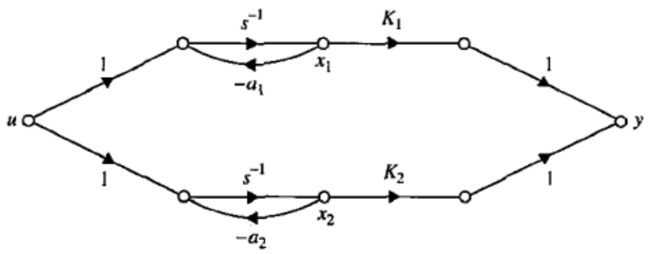


Figure 5-10 State diagram of the transfer function of Eq. (5-86) by parallel decomposition.

The conclusion is that, for transfer functions with distinct poles, parallel decomposition will lead to the **DCF** for the state equations. For transfer functions with multiple-order eigenvalues, parallel decomposition to a state diagram with a minimum number of integrators will lead to the **JCF** state equations. The following example will clarify this point.

Example 5-15: [59] Consider the following transfer function and its partial-fraction expansion:

$$\frac{Y(s)}{U(s)}=\frac{2s^2+6s+5}{(s+1)^2(s+2)}=\frac{1}{(s+1)^2}+\frac{1}{s+1}+\frac{1}{s+2} \qquad (5-88)$$

Note that the transfer function is of the third order, and, although the total order of the terms on the right-hand side of Eq. (5-88) is four, only three integrators should be used in the state diagram, which is drawn as shown in Fig. 5-11. The minimum number of three integrators is used, with one integrator being shared by two channels. The state equations of the system are written directly from Fig. 5-11.

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}0\\ 1\\ 1\end{bmatrix}u$$

which is recognized to be the JCF.

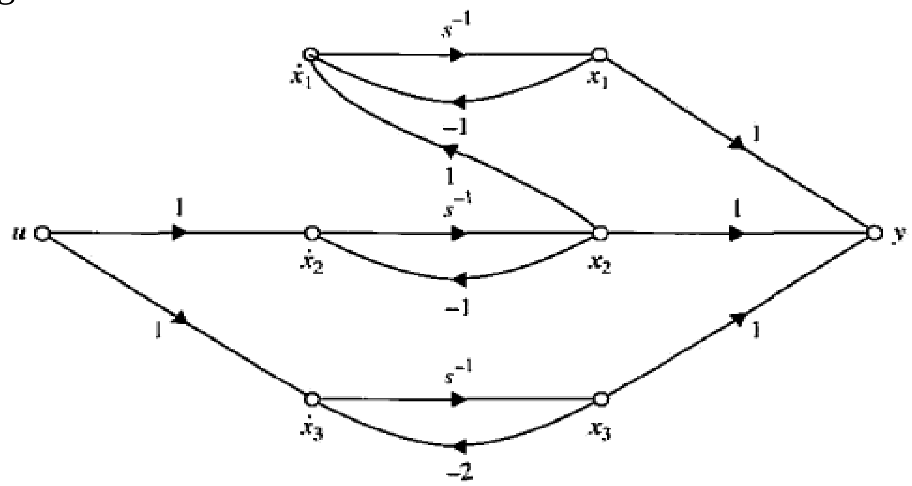


Figure 5-11 State diagram of the transfer function of Eq. (5-88) by parallel decomposition.

5-6 Nonuniqueness of a Set of State Variables

It has been stated that a set of state variables is not unique for a given system. Suppose that $x_1, x_2, \dots, x_n$ are a set of state variables.

Then we may take as another set of state variables any set of functions

$$\hat{x}_1 = X_1(x_1, x_1, \dots, x_n)$$
$$\hat{x}_2 = X_2(x_1, x_2, \dots, x_n)$$
$$\vdots$$
$$\hat{x}_n = X_n(x_1, x_2, \dots, x_n)$$

Provided that, for every set of values $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$, there corresponds a unique set of values $x_1, x_2, \dots, x_n$, and vice versa. Thus, if x is a state vector then x̂, where

$$\hat{\mathrm{x}} = \mathrm{Px}$$

is also a state vector, provided the matrix P is nonsingular. Different state vectors convey the same information about the system behavior.

5-7 Linearizations of Nonlinear Mathematical Models

5-7-1 Nonlinear Systems

A system is nonlinear if the principle of superposition does not apply. Thus, for a nonlinear system the response to two inputs cannot be calculated by treating one input at a time and adding the results.

Although many physical relationships are often represented by linear equations, in most cases actual relationships are not quite linear. In fact, a careful study of physical systems reveals that even so-called "linear systems" are really linear only in limited operating ranges. In practice, many electromechanical systems, hydraulic systems, pneumatic systems, and so on, involve nonlinear relationships among the variables.

For example, the output of a component may saturate for large input signals. There may be a dead space that affects small signals. (The dead space of a component is a small range of input variations to which the component is insensitive.) Square-law nonlinearity may occur in some components. For instance, dampers used in physical systems may be linear for low-velocity operations but may become nonlinear at high velocities and the damping force may become proportional to the square of the operating velocity.

5-7-2 Linearization of Nonlinear Systems

In control engineering a normal operation of the system may be around an equilibrium point, and the signals may be considered small signals around the equilibrium. (It should be pointed out that there are many exceptions to such a case.) However, if the system operates around an equilibrium point and if the signals involved are small signals, then it is possible to approximate the nonlinear system by a linear system. Such a linear system is equivalent to the nonlinear system considered within a limited operating range. Such a linearized model (linear, time-invariant model) is very important in control engineering.

The linearization procedure to be presented in the following is based on the expansion of nonlinear function into a Taylor series about the operating point and the retention of only the linear term. Because we neglect higher-order terms of the Taylor series expansion, these neglected terms must be small enough; that is, the variables deviate only slightly from the operating condition. (Otherwise, the result will be inaccurate.).

5-7-3 Linearization Using Taylor Series: Classical Representation

In general, Taylor series may be used to expand a nonlinear function $f(x(t))$ about a reference or operating value $x_0(t)$. An operating value could be the equilibrium position in a spring-mass-damper, a fixed voltage in an electrical system, steady state pressure in a fluid system, and so on. A function $f(x(t))$ can therefore be represented in a form

$$f\big(x(t)\big) = \sum_{i=1}^{n} c_i(x(t) - x_0(t))^i$$

where the constant c_i represents the derivatives of $f(t)$ with respect to $x(t)$ and evaluated at the operating point $x_0(t)$. That is

$$c_i = \frac{1}{i!}\frac{d^i f(x_0)}{dx^i}$$

or

$$f(x) = f(x_0) + \frac{df(x_0)}{dt}(x - x_0) + \frac{1}{2}\frac{d^2 f(x_0)}{dt^2}(x - x_0)^2 + \frac{1}{6}\frac{d^3 f(x_0)}{dt^3}(x - x_0)^3 + \cdots + \frac{1}{n!}\frac{d^n f(x_0)}{dt^n}(x - x_0)^n \qquad (5-89)$$

If $\Delta(x) = x(t) - x_0(t)$ is small, the series Eq. (5-156) converges, and a linearization scheme may be used by replacing $f(x(t))$ with the first two terms in Eq. (5-89). That is,

$$f\big(x(t)\big) \approx f\big(x_0(t)\big) + \frac{df\big(x_0(t)\big)}{dt}\big(x(t) - x_0(t)\big)$$
$$= c_0 + c_1 \Delta x$$

5-7-4 Linearization Using the State Space Approach

Alternatively, let us represent a nonlinear system by the following vector-matrix state equations:

$$\frac{dx(t)}{dt} = f[x(t),\ r(t)] \qquad (5-90)$$

where x(t) represents the n x l state vector; r(t), the p x 1 input vector; and f[x(r), r (r)], an n x I function vector. In general, f is a function of the state vector and the input vector.

Being able to represent a nonlinear and/or time-varying system by state equations is a distinct advantage of the state-variable approach over the transfer-function method, since the latter is strictly denned only for linear time-invariant systems.

As a simple example, the following nonlinear state equations are given:

$$\frac{dx_1(t)}{dt} = x_1(t) + x_2^2(t)$$
$$\frac{dx_2(t)}{dt} = x_1(t) + r(t)$$

Because nonlinear systems are usually difficult to analyze and design, it is desirable to perform a linearization whenever the situation justifies it.

A linearization process that depends on expanding the nonlinear state equations into a Taylor series about a nominal operating point or trajectory is now described. All the terms of the Taylor series of order higher than the first are discarded, and the linear approximation of the nonlinear state equations at the nominal point results.

Let the nominal operating trajectory be denoted by $x_0(t)$, which corresponds to the nominal input $r_0(t)$ and some fixed initial states. Expanding the nonlinear state equation of Eq. (5-90) into a Taylor series about $x(t) = x_0(t)$ and neglecting all the higher-order terms yields

$$\dot{x}_i(t) = f_i(x_0, r_0) + \sum_{j=1}^{n} \left.\frac{\partial f_i(x,r)}{\partial x_j}\right|_{x_0, r_0} \left(x_j - x_{0j}\right) + \sum_{j=1}^{p} \left.\frac{\partial f_i(x,r)}{\partial r_j}\right|_{x_0, r_0} \qquad (5-91)$$

where $i = 1, 2, \ldots, n$.

Let

$$\Delta x_i = x_i - x_{0i} \quad (5-92)$$

and

$$\Delta r_j = r_j - r_{0j}$$

Then

$$\Delta \dot{x}_i = \dot{x}_i - \dot{x}_{0i}$$

Since

$$\dot{x}_{0i} = f_i(x_0, r_0)$$

Eq. (5-91) is written

$$\Delta \dot{x}_i = \sum_{j=1}^{n} \left. \frac{\partial f_i(x,r)}{\partial x_j} \right|_{x_0, r_0} \Delta x_j + \sum_{j=1}^{p} \left. \frac{\partial f_i(x,r)}{\partial r_j} \right|_{x_0, r_0} \Delta r_j \quad (5-93)$$

Eq. (5-93) may be written in vector-matrix form:

$$\Delta \dot{x} = A^* \Delta x + B^* \Delta r$$

where

$$A^* = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

$$B^* = \begin{bmatrix} \frac{\partial f_1}{\partial r_1} & \frac{\partial f_1}{\partial r_2} & \cdots & \frac{\partial f_1}{\partial r_p} \\ \frac{\partial f_2}{\partial r_1} & \frac{\partial f_2}{\partial r_2} & \cdots & \frac{\partial f_2}{\partial r_p} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial r_1} & \frac{\partial f_n}{\partial r_2} & \cdots & \frac{\partial f_n}{\partial r_p} \end{bmatrix}$$

5-7-5 Linear Approximation of Nonlinear Mathematical Models

To obtain a linear mathematical model for a nonlinear system, we assume that the variables deviate only slightly from some operating condition. Consider a system whose input is x(t) and output is y(t). The relationship between y(t) and x(t) is given by

$$y = f(t) \quad (5-94)$$

If the normal operating condition corresponds to $\bar{x}, \bar{y}$, then Eq. (5–94) may be expanded into a Taylor series about this point as follows:

$$y = f(t)$$

$$= f(\bar{x}) + \frac{df}{dx}(x - \bar{x}) + \frac{1}{2!}\frac{d^2 f}{dx^2}(x - \bar{x})^2 + \cdots \quad (5-95)$$

where the derivatives $df/dx, d^2f/dx^2, ...$ are evaluated at $x = \bar{x}$ If the variation $x - \bar{x}$ is small, we may neglect the higher-order terms in $x - \bar{x}$. Then Equation (5–95) may be written as

$$y = \bar{y} + K(x - \bar{x}) \qquad (5-96)$$

Where

$$\bar{y} = f(\bar{x})$$

$$K = \left.\frac{df}{dx}\right|_{x=\bar{x}}$$

Equation (5–96) may be rewritten as

$$y - \bar{y} = K(x - \bar{x}) \qquad (5-97)$$

which indicates that $y - \bar{y}$ is proportional to $x - \bar{x}$. Equation (5–97) gives a linear mathematical model for the nonlinear system given by Equation (5–94) near the operating point $x = \bar{x}, \quad y = \bar{y}$.

Next, consider a nonlinear system whose output y is a function of two inputs x_1 and x_1 , so that

$$y = f(x_1, x_2) \qquad (5-98)$$

To obtain a linear approximation to this nonlinear system, we may expand Equation (5–98) into a Taylor series about the normal operating point $\bar{x}_1, \bar{x}_2$.Then Equation (5–98) becomes

$$y = f(\bar{x}_1, \bar{x}_2) + \left[\frac{\partial f}{\partial x_1}(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_2 - \bar{x}_2)\right] + \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x_1^2}(x_1 - \bar{x}_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \frac{\partial^2 f}{\partial x_2^2}(x_2 - \bar{x}_2)^2\right] + \cdots$$

where the partial derivatives are evaluated at Near the normal operating point, the higher-order terms may be neglected. The linear mathematical model of this nonlinear system in the neighborhood of the normal operating condition is then given by

$$y - \bar{y} = K_1(x_1 - \bar{x}_1) + K_2(x_2 - \bar{x}_2)$$

Where

$$\bar{y} = f(\bar{x}_1, \bar{x}_2)$$

$$K_1 = \left.\frac{\partial f}{\partial x_1}\right|_{x_1=\bar{x}_1, x_2=\bar{x}_2}$$

$$K_2 = \left.\frac{\partial f}{\partial x_2}\right|_{x_1=\bar{x}_1, x_2=\bar{x}_2}$$

Example 5-16: [6] Linearize the nonlinear equation

$$z = xy$$

in the region $5 \le x \le 7$, $10 \le y \le 12$. Find the error if the linearized equation is used to calculate the value of z when $x = 5, y = 10$.

Solution. Since the region considered is given by $5 \le x \le 7, 10 \le y \le 12$, choose $\bar{x} = 6, \bar{y} = 11$. Then $\bar{z} = \bar{x}\bar{y} = 66$. Let us obtain a linearized equation for the nonlinear equation near a point $\bar{x} = 6, \bar{y} = 11$.

Expanding the nonlinear equation into a Taylor series about point $x = \bar{x}, y = \bar{y}$ and neglecting the higher-order terms, we have

$$z - \bar{z} = a(x - \bar{x}) + b(y - \bar{y})$$

Where

$$a = \left.\frac{\partial(xy)}{\partial x}\right|_{x=\bar{x}, y=\bar{y}} = \bar{y} = 11$$

$$b = \left.\frac{\partial(xy)}{\partial y}\right|_{x=\bar{x}, y=\bar{y}} = \bar{x} = 6$$

Hence the linearized equation is

$$z - 66 = 11(x - 6) + 6(y - 11)$$

Or

$$z = 11x + 6y - 66$$

When $x = 5, y = 10$, the value of z given by the linearized equation is

$$z = 11x + 6y - 66 = 55 + 60 - 66 = 49$$

The exact value of z is $z = xy = 50$. The error is thus $50 - 49 = 1$. In terms of percentage, the error is 2%.

Example 5-17: [6] Linearize the nonlinear equation

$$z = x^2 + 4xy + 6y^2$$

in the region defined by $8 \leq x \leq 10, 2 \leq y \leq 4$.

Solution. Define $f(x, y) = z = x^2 + 4xy + 6y^2$

Then

$$z = f(x, y) = f(\bar{x}, \bar{y}) + \left[\frac{\partial f}{\partial x}(x - \bar{x}) + \frac{\partial f}{\partial y}(y - \bar{y})\right]_{x=\bar{x}, y=\bar{y}} + \cdots$$

where we choose $\bar{x} = 9, \bar{y} = 3$.

Since the higher-order terms in the expanded equation are small, neglecting these higher-order terms, we obtain

$$z - \bar{z} = K_1(x - \bar{x}) + K_2(y - \bar{y})$$

Where

$$K_1 = \left.\frac{\partial f}{\partial x}\right|_{x=\bar{x}, y=\bar{y}} = 2\bar{x} + 4\bar{y} = 2 \times 9 + 4 \times 3 = 30$$

$$K_2 = \left.\frac{\partial f}{\partial y}\right|_{x=\bar{x}, y=\bar{y}} = 4\bar{x} + 12\bar{y} = 4 \times 9 + 12 \times 3 = 72$$

$$\bar{z} = \bar{x}^2 + 4\bar{x}\bar{y} + 6\bar{y}^2 = 9^2 + 4 \times 9 \times 3 + 6 \times 3^2 = 243$$

Thus

$$z - 243 = 30(x - 9) + 72(y - 3)$$

Hence a linear approximation of the given nonlinear equation near the operating point is

$$z - 30x - 72y + 243 = 0$$

CHAPTER SIX

Transformation of Mathematical Models with MATLAB

MATLAB is quite useful to transform the system model from transfer function to state space, and vice versa. We shall begin our discussion with transformation from transfer function to state space.

Let us write the closed-loop transfer function as

$$\frac{Y(s)}{U(s)} = \frac{\text{numerator polynomial in s}}{\text{denominator polynomial in s}} = \frac{\text{num}}{\text{den}} \qquad (6-1)$$

Once we have this transfer-function expression, the MATLAB command

$$[\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}] = \mathrm{tf2ss}(\mathrm{num}, \mathrm{den}) \qquad (6-2)$$

will give a state-space representation. It is important to note that the state-space representation for any system is not unique. There are many (infinitely many) state-space representations for the same system. The MATLAB command gives one possible such state-space representation.

6-1 Transformation from Transfer Function to State Space Representations

Consider the transfer function system

$$\frac{Y(s)}{U(s)} = \frac{s}{(s+10)(s^2+4s+16)}$$

$$= \frac{s}{s^3+14s^2+56s+160} \qquad (6-3)$$

There are many (infinitely many) possible state-space representations for this system. One possible state-space representation is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -160 & -56 & -14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -14 \end{bmatrix} u \qquad (6-4)$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [0]u \qquad (6-5)$$

Another possible state-space representation (among infinitely many alternatives) is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -14 & -56 & -160 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u \qquad (6-6)$$

$$y = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [0]u \qquad (6-7)$$

MATLAB transforms the transfer function given by Equation (6–3) into the state-space representation given by Equations (6–6) and (6–7). For the example system considered here, MATLAB Program 6–1 will produce matrices **A**, **B**, **C**, and D.

MATLAB Program 6–1

```
>> num = [1 0]; den = [1 14 56 160]; [A,B,C,D] = tf2ss(num,den)
A =
  -14   -56  -160
    1     0     0
    0     1     0
B =
    1
    0
    0
C =   0     1     0
D =   0
```

6-2 Transformation from State Space Representations to Transfer Function

To obtain the transfer function from state-space equations, use the following command:

$$[\text{num}, \text{den}] = ss2\text{tf}(\text{A}, \text{B}, \text{C}, \text{D}, \text{iu}) \qquad (6-8)$$

iu must be specified for systems with more than one input. For example, if the system has three inputs (u_1, u_2, u_3), then iu must be either 1, 2, or 3, where 1 implies u_1, 2 implies u_2, and 3 implies u_3.

If the system has only one input, then either

$$[\text{num}, \text{den}] = ss2\text{tf}(\text{A}, \text{B}, \text{C}, \text{D}) \qquad (6-9)$$

or

$$[\text{num}, \text{den}] = \text{ss2tf}(\text{A}, \text{B}, \text{C}, \text{D}, 1) \qquad (6-10)$$

may be used. For the case where the system has multiple inputs and multiple outputs.

Example 6-1: [6] We Obtain the transfer function of the system defined by the following state-space equations:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -5 & -25 & -5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}1\\ 25\\ -120\end{bmatrix}u \qquad (6-11)$$

$$y = [1 \quad 0 \quad 0]\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} \qquad (6-12)$$

MATLAB Program 6-2 will produce the transfer function for the given system. The transfer function obtained is given by

$$\frac{Y(s)}{U(s)} = \frac{25s + 5}{s^3 + 5s^2 + 25s + 5} \qquad (6-13)$$

MATLAB Program 6–2

```
>> A = [0 1 0; 0 0 1; -5 -25 -5]; B = [0; 25; -120]; C = [1 0 0]; D = [0];
>> [num,den] = ss2tf(A,B,C,D)
num =
       0    0.0000   25.0000    5.0000
den =
   1.0000    5.0000   25.0000    5.0000
% ***** The same result can be obtained by entering the following command: *****
>> [num,den] = ss2tf(A,B,C,D,1)
num =
       0    0.0000   25.0000    5.0000
den =
   1.0000    5.0000   25.0000    5.0000
```

Example 6–2: [6] Obtain a state-space equation and output equation for the system defined by

$$\frac{Y(s)}{U(s)} = \frac{2s^3 + s^2 + s + 2}{s^3 + 4s^2 + 5s + 2} \qquad (6-14)$$

Solution: From the given transfer function, the differential equation for the system is

$$\dddot{y} + 4\ddot{y} + 5\dot{y} + 2y = 2\dddot{u} + \ddot{u} + \dot{u} + 2u \qquad (6-15)$$

Comparing this equation with the standard equation given by Equation (5–7), rewritten

$$\dddot{y} + a_1\ddot{y} + a_2\dot{y} + a_3y = b_0\dddot{u} + b_1\ddot{u} + b_2\dot{u} + b_3u \qquad (6-16)$$

we find

$$a_1 = 4, \qquad a_2 = 5, \qquad a_3 = 2$$
$$b_0 = 2, \qquad b_1 = 1, \qquad b_3 = 2$$

Referring to Equation (5–9), we have

$$\beta_0 = b_0 = 2$$
$$\beta_1 = b_1 - a_1\beta_0 = 1 - 4 \times 2 = -7$$
$$\beta_2 = b_2 - a_1\beta_1 - a_2\beta_0 = 1 - 4 \times (-7) - 5 \times 2 = 19$$

$$\beta_3 = b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0$$
$$= 2 - 4 \times 19 - 5 \times (-7) - 2 \times 2 = 43$$

Referring to Equation (5–8), we define

$$x_1 = y - \beta_0 u = y - 2u$$
$$x_2 = \dot{x}_1 - \beta_1 u = \dot{x}_1 + 7u$$
$$x_2 = \dot{x}_2 - \beta_2 u = \dot{x}_2 - 19u$$

Then referring to Equation (5–10),

$$\dot{x}_1 = x_2 - 7u$$
$$\dot{x}_2 = x_3 + 19u$$
$$\dot{x}_3 = -a_3 x_1 - a_2 x_2 - a_1 x_3 + \beta_3 u$$
$$= -2x_1 - 5x_2 - 4x_3 - 43u$$

Hence, the state-space representation of the system is

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -2 & -5 & -4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}-7\\ 19\\ -43\end{bmatrix}u \qquad (6-17)$$

$$y = \begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + 2u \qquad (6-18)$$

This is one possible state-space representation of the system. There are many (infinitely many) others. If we use MATLAB, it produces the following state-space representation:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-4 & -5 & -2\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}u \qquad (6-19)$$

$$y = \begin{bmatrix}-7 & -9 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + 2\,u \qquad (6-20)$$

MATLAB Program 6 – 3

```
>> num = [2 1 1 2]; den = [1 4 5 2];  [A,B,C,D] = tf2ss(num,den)
A =
   -4    -5    -2
    1     0     0
    0     1     0
B =
    1
    0
    0
C =  -7    -9    -2
D =   2
```

Example 6-3: [6] Consider a system with multiple inputs and multiple outputs. When the system has more than one output, the MATLAB command

$$[num, den] = ss2tf(A, B, C, D, iu) \qquad (6-21)$$

produces transfer functions for all outputs to each input. (The numerator coefficients are returned to matrix num with as many rows as there are outputs.)

Consider the system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -25 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} u \qquad (6-22)$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad (6-23)$$

This system involves two inputs and two outputs. Four transfer functions are involved: $Y_1(s)/U_1(s)$, $Y_2(s)/U_1(s)$, $Y_1(s)/U_2(s)$, and $Y_2(s)/U_2(s)$.(When considering input u_1 , we assume that input u_2 is zero and vice versa.)

Solution. MATLAB Program 6-4 produces four transfer functions.

This is the MATLAB representation of the following four transfer functions:

$$\frac{Y_1(s)}{U_1(s)} = \frac{s+4}{s^2+4s+25}, \qquad \frac{Y_2(s)}{U_1(s)} = \frac{-25}{s^2+4s+25} \qquad (6-24)$$

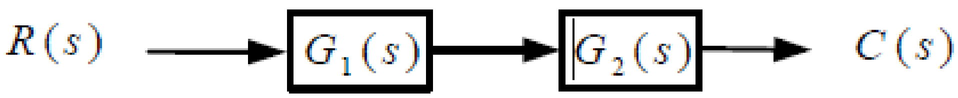
$$\frac{Y_1(s)}{U_2(s)} = \frac{s+5}{s^2+4s+25}, \qquad \frac{Y_2(s)}{U_2(s)} = \frac{s-25}{s^2+4s+25} \qquad (6-25)$$

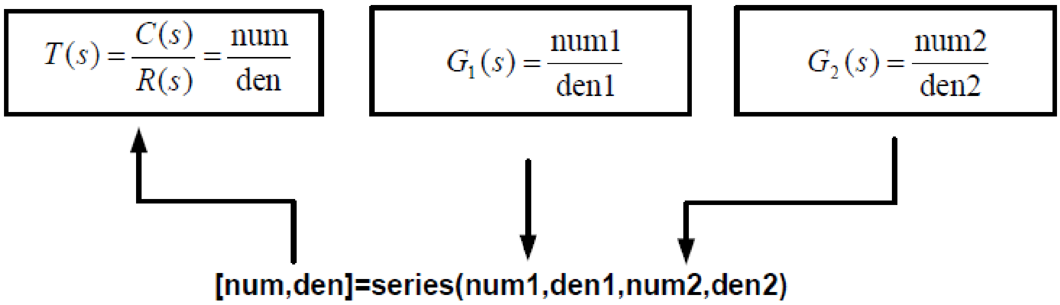
MATLAB Program 6–4

```
>> A = [0 1;-25 -4]; B = [1 1;0 1]; C = [1 0;0 1]; D = [0 0;0 0];
>> [num,den] = ss2tf(A,B,C,D,1)
num =
        0    1.0000    4.0000
        0         0  -25.0000
den =
    1.0000    4.0000   25.0000
>> [num,den] = ss2tf(A,B,C,D,2)
num =
        0    1.0000    5.0000
        0    1.0000  -25.0000
den =
    1.0000    4.0000   25.0000
```

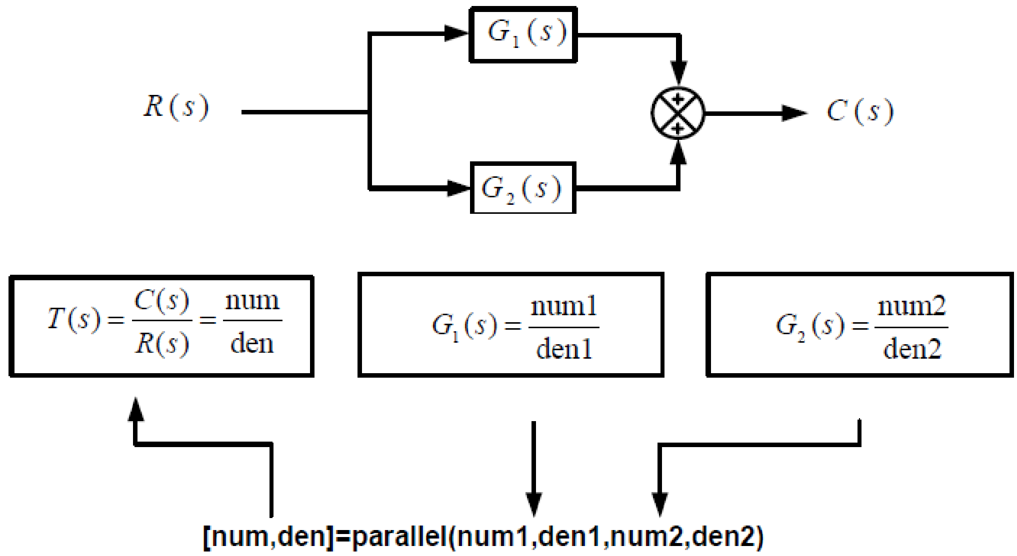
6-3 MATLAB Implementation

6-3-1 Series Connection

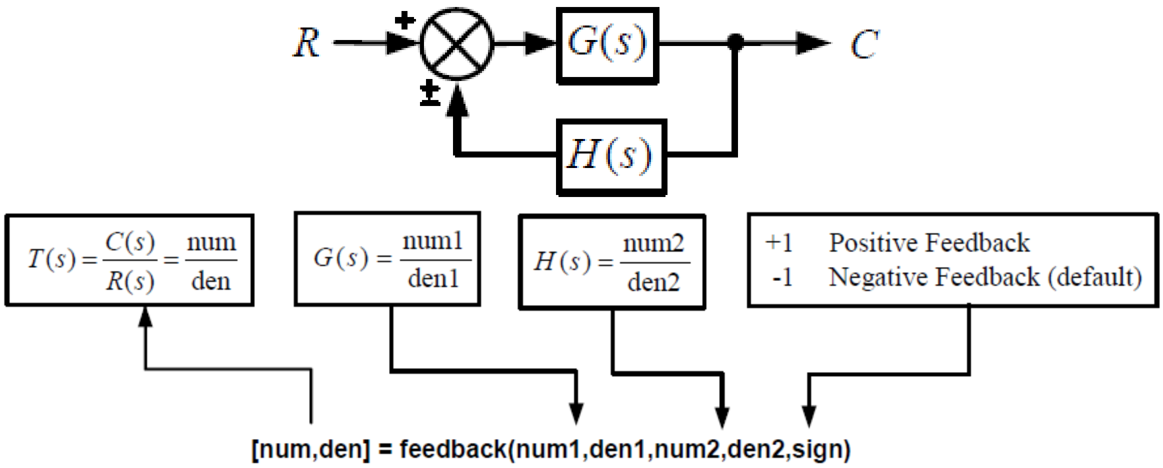




6-3-2 Parallel Connection



6-3-3 Feedback Connection



6-4 Partial-Fraction Expansion with MATLAB

MATLAB representation of Transfer Functions (G). The transfer function of a system is represented by two arrays of numbers. For example, consider a system defined by

$$\frac{Y(s)}{U(s)}=\frac{25}{s^2+4s+25} \qquad (6-26)$$

This system is represented as two arrays, each containing the coefficients of the polynomials in descending powers of s as follows

MATLAB Program 6–5
>> num = 25; den = [1 4 25]; sys = tf(num,den) Transfer function: 25 -------------- s^2 + 4 s + 25

Partial-Fraction Expansion with MATLAB. MATLAB allows us to obtain the partial-fraction expansion of the ratio of two polynomials,

$$\frac{B(s)}{A(s)}=\frac{num}{den}=\frac{b(1)s^{h}+b(2)s^{h-1}+\cdots+b(h)}{a(1)s^{1}+a(2)s^{n-1}+\cdots+a(n)} \qquad (6-27)$$

Where $a(1)\neq 0$, some of $a(i)$ and $b(j)$ may be zero, and num and den are row vectors that specify the numerator and denominator of $B(s)/A(s)$. That is,

```
>> num = [b(1) b(2) ... b(h)];
>> den = [a(1) a(2) ... a(h)];
```

The command

```
>> [r,p,k] = residu(num,den);
```

Finds the residues, poles and direct terms of a partial fraction expansion of the ratio of the two polynomials *B*(*s*) and *A*(*s*). The partial fraction expansion of $B(s)/A(s)$ is given by

$$\frac{B(s)}{A(s)}=k(s)+\frac{r(1)}{s-p(1)}+\frac{r(2)}{s-p(2)}+\cdots+\frac{r(n)}{s-p(n)} \qquad (6-28)$$

As an example, consider the function

$$\frac{B(s)}{A(s)}=\frac{num}{den}=\frac{s^4+8s^3+16s^2+9s+6}{s^3+6s^2+11s+6} \qquad (6-29)$$

MATLAB Program 6–6
>> num = [1 8 16 9 6]; den = [1 6 11 6]; [r,p,k] = residue(num,den) r = -6.0000 -4.0000 3.0000 p = -3.0000 -2.0000 -1.0000 k = 1 2

Therefore , the partial-fraction expansion of $B(s)/A(s)$ is:

$$\frac{B(s)}{A(s)}=\frac{num}{den}=\frac{s^4+8s^3+16s^2+9s+6}{s^3+6s^2+11s+6}$$

$$=s+2-\frac{6}{s+3}-\frac{4}{s+2}+\frac{3}{s+1} \qquad (6-30)$$

The command

[num,den] = residue(r,p,k)

where r,p and k are outputs , converts the partial-fraction expansion back to the polynomial ratio $B(s)/A(s)$ as shown below

MATLAB Program 6–7
>> r=[-6 -4 3]; p=[-3 -2 -1]; k=[1 2]; [num,den]=residue(r,p,k) num = 1 8 16 9 6 den = 1 6 11 6

6-5 Transient Response Analysis with MATLAB

MATLAB Representation of Transfer-Functions (G) Systems.

Such a block represents a system or an element of a system. To simplify our presentation, we shall call the block with a G a system. MATLAB uses sys to represent such a system. The statement

>> sys=tf(num,den)

represents the system. For example, consider the system

$$\frac{Y(s)}{U(s)} = \frac{2s + 25}{s^2 + 4s + 25}$$

This system is represented as two arrays, each containing the coefficients of the polynomials in descending powers of s as follows

MATLAB Program 6-8
>> num=[2 25]; den=[1 4 25]; sys=tf(num,den) % MATLAB will automatically respond with the display Transfer function: 2 s + 25 -------------- s^2 + 4 s + 25

6-5-1 Step Response

The step function plots the unit step response, assuming the I.C's are zero. The basic syntax is `step(sys)`, where sys is the LTI object defined previously.

The basic syntax commands are summarized below

Table 6-1 summarizes these functions.

Command (Basic Syntax)	Use
`>> step(sys)`	generates a plot of a unit step response and displays a response curve on the screen. The computation time interval Δt and the time span of the response tf are determined

	automatically by MATLAB.
`>> step(sys,tf)`	generates a plot of a unit step response and displays a response curve on the screen for the specified final time tf . The computation time interval Δt is determined automatically by MATLAB.
`>> step(sys,t)`	generates a plot of a unit step response and displays a response curve on the screen for the user specified time t where t = 0 : Δt : tf .
`>>[y,t]=step(sys,…)`	Returns the output y, and the time array t used for the simulation. No plot is drawn. The array y is $p \times q \times m$ where p is length(t), q is the number of outputs, and m is the number of inputs.
`>>step(sys1,sys2,…,t)`	Plots the step response of multiple LTI systems on a single plot. The time vector t is optional. You can specify line color, line style and marker for each system.

The steady state response and the time to reach that steady state are automatically determined. The steady state response is indicated by horizontal dotted line.

Example 6-4: The spring-mass-dashpot system mounted on a cart as shown in Figure 6-1.

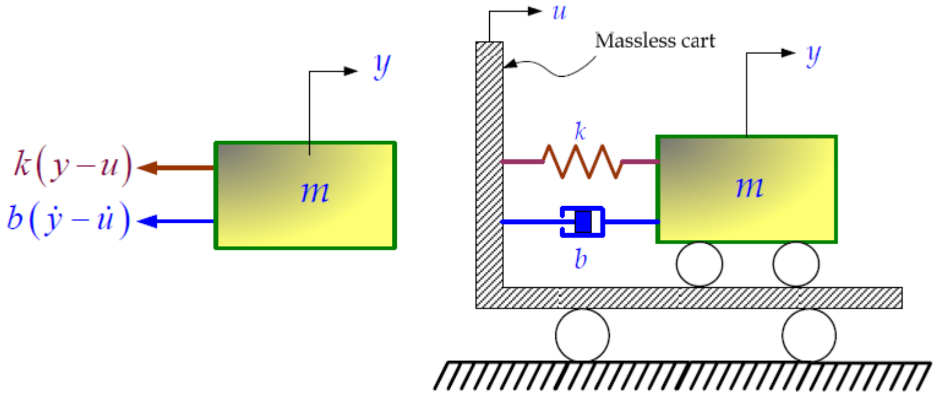


Figure 6-1 Spring-mass-dashpot system mounted on a cart and its FBD.

The transfer function of the system is

$$\text{Transfer Function (G)} = \frac{Y(s)}{U(s)} = \frac{(bs+k)}{(ms^2+bs+k)}$$

For m =10 kg , b = 20N-s/m and k =100N/m. Find the response $y\,(t)$ for a unit step input $u(t) = 1(t)$.

MATLAB Program 6-9
`>> m=10; b=20; k=100; num=[b k]; den=[m b k];`
`>> sys=tf(num,den);`
`>> step(sys)`
`>> grid`

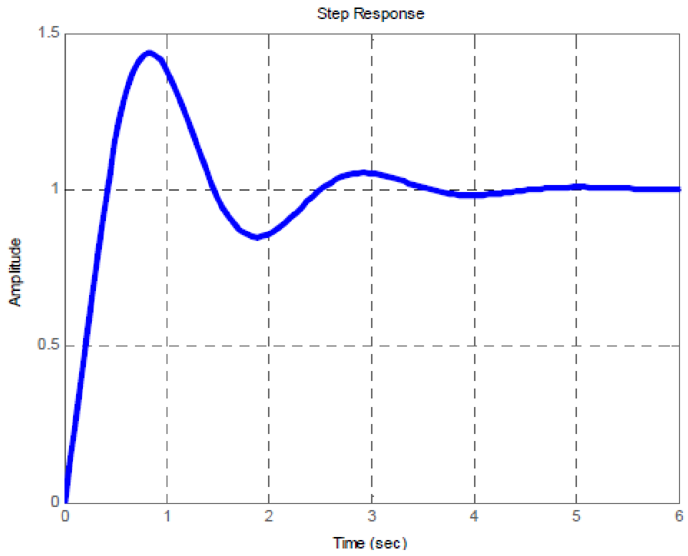


Figure 6-2 Unit step response curve

Example 6-7: The mechanical system shown in Figure 6-2.

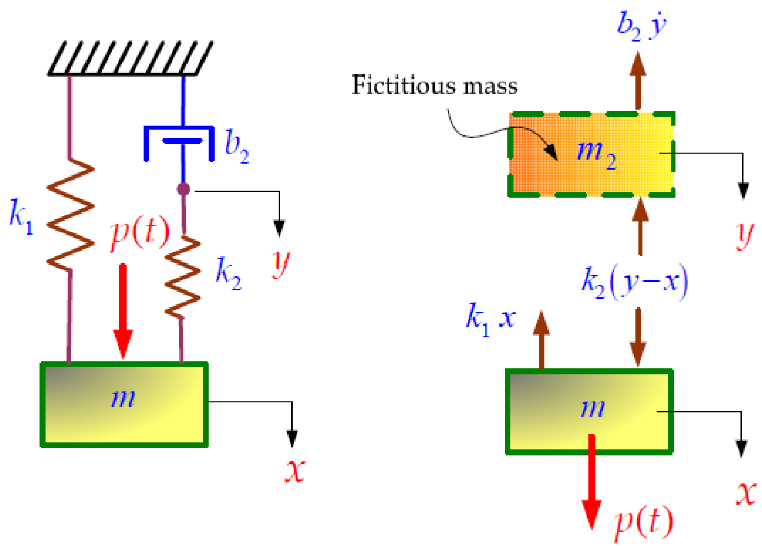


Figure 6-3 Mechanical system and its FBD.

The transfer functions of the system are

$$\frac{X(s)}{P(s)} = \frac{b_2 s + k_2}{m b_2 s^3 + m k_2 s^2 + (k_1 + k_2) b_2 s + k_1 k_2}$$

and

$$\frac{Y(s)}{P(s)} = \frac{k_2}{m b_2 s^3 + m k_2 s^2 + (k_1 + k_2) b_2 s + k_1 k_2}$$

For m = 0.10 kg , b_2 = 0.4N-s/m and k_1 = 6N/m , k_2 = 4N/m and $p(t)$ is a step input of magnitude 10N. Obtain the responses $x(t)$ and $y(t)$.

MATLAB Program 6-10

```
>> m=0.1; b2=0.4; k1=6; k2=4;
>> num1=[b2 k2]; num2=[k2];
>> den=[m*b2   m*k2   k1*b2+k2*b2   k1*k2];
>> sys1=tf(num1,den); sys2=tf(num2,den);
>> step(10*sys1,'r:',10*sys2,'b'); grid
>> gtext('x(t)');gtext('y(t)')
```

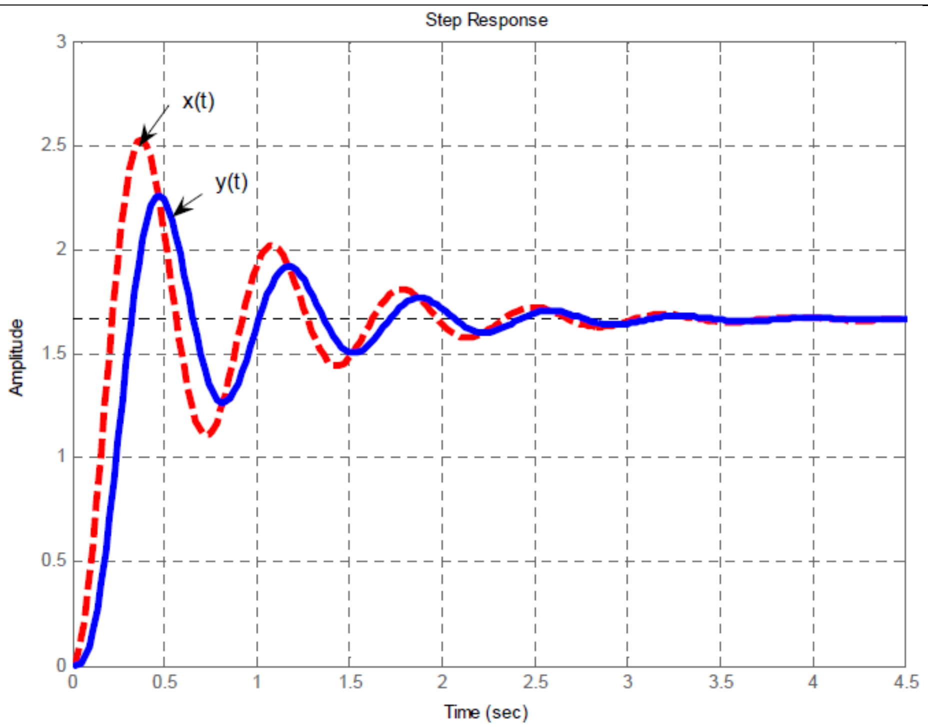


Figure 6-4. Step response curves $x(t)$ and $y\ (t)$

6-5-2 Impulse Response

The impluse function plots the unit-impulse response, assuming the I.C's are zero. The basic syntax is impulse(sys), where sys is the LTI object.

The basic syntax commands are summarized below

Table 6-2 summarizes these functions.

Command (Basic Syntax)	**Use**
`>> impulse(sys)`	generates a plot of a unit step response and displays a response curve on the screen. The computation time interval Δt and the time span of the response tf are determined automatically by MATLAB.
`>> impulse(sys,tf)`	generates a plot of a unit step response and displays a response curve on the screen for the specified final time tf . The computation time interval Δt is determined automatically by MATLAB.
`>> impulse(sys,t)`	generates a plot of a unit step response and displays a response curve on the screen for the user specified time t where t = 0 : Δt : tf .
`>>[y,t]=impulse(sys,…)`	Returns the output y, and the time array t used for the simulation. No plot is drawn. The array y is $p \times q \times m$ where p is length(t), q is the number of

	outputs, and m is the number of inputs.
`>>impulse(sys1, sys2,…,t)`	Plots the step response of multiple LTI systems on a single plot. The time vector t is optional. You can specify line color, line style and marker for each system.

The steady state response and the time to reach that steady state are automatically determined. The steady state response is indicated by horizontal dotted line.

6-5-3 Impulse Input

The impulse response of a mechanical system can be observed when the system is subjected to a very large force for a very short time, for instance, when the mass of a spring-mass-dashpot system is hit by a hammer or a bullet. Mathematically, such an impulse input can be expressed by an impulse function.

The impulse function is a mathematical function without any actual physical counterpart. However, as shown in Figure 6-5 (a), if the actual input lasts for a short time Δt but has a magnitude h , so that the area $h\ \Delta t$ in a time plot is not negligible, it can be approximated by an impulse function. The impulse input is usually denoted by a vertical arrow, as shown in Figure 6-5 (b), to indicate that it has a very short duration and a very large height.

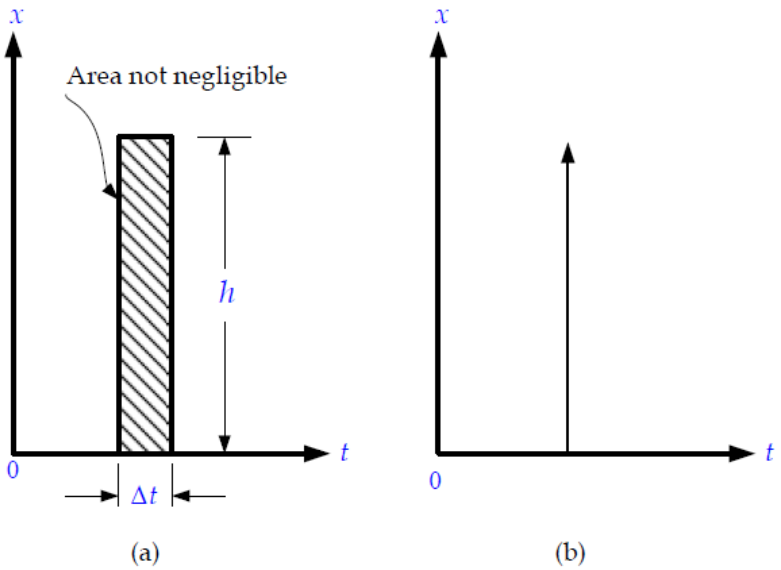


Figure 6-5 Impulse inputs

Example 6-6: [6] Consider the previous Example 6-5 but with an impulse input of magnitude 10 N.

MATLAB Program 6-11

```
>> m=0.1; b2=0.4; k1=6;k2=4;
>> num1=[b2 k2]; num2=[k2]
>> den=[m*b2   m*k2   k1*b2+k2*b2   k1*k2]
>> sys1=tf(num1,den); sys2=tf(num2,den)
>> impulse(10*sys1,'r:',10*sys2,'b'); grid
>> gtext('x(t)');gtext('y(t)')
```

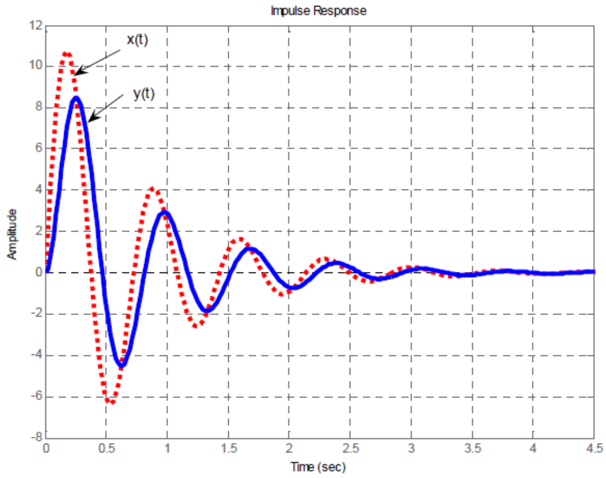


Figure 6-6 Impulse response curves $x(t)$ and $y\ (t)$

6-5-4 Obtaining response to arbitrary input

The `lsim` function plots the response of the system to an arbitrary input. The basic syntax commands is summarized below

Table 6-3 summarizes these functions.

Command (Basic Syntax)	Use
`>> lsim(sys,u,t)`	produces a plot of the time response of the LTI model sys to the input time history t, u. The vector t specifies the time samples for the simulation and consists of regularly spaced time samples. $t = 0 : \Delta t :$ tf The matrix u must have as many rows as time samples (length(t)) and as many columns as system inputs. Each row $u(i,:)$ specifies the input value(s) at the time sample $t(i)$.

Example 6-7: [6] Consider the mass-spring-dashpot system mounted on a cart of Figure 6-1.

The Transfer Function G of the system is

$$\frac{Y(s)}{U(s)} = \frac{(bs + k)}{(ms^2 + bs + k)}$$

where $Y(s)$ is the output $U(s)$ is the input. Assume that m =10kg ,b=20N-s/m and k =100N/m. Find the response $y\ (t)$ for a ramp input with a slope of 2, ($r(t) = 2t$).

MATLAB Program 6-12

```
>> m=10; b=20;
>> k=100;
>> num=[b k];
>> den=[m b k];
>> sys=tf(num,den);
>> t=[0:0.001:3];
>> u=2*t; lsim(sys,u,t);
>> grid;
>> gtext('x(t)');
>> gtext('y(t)')
```

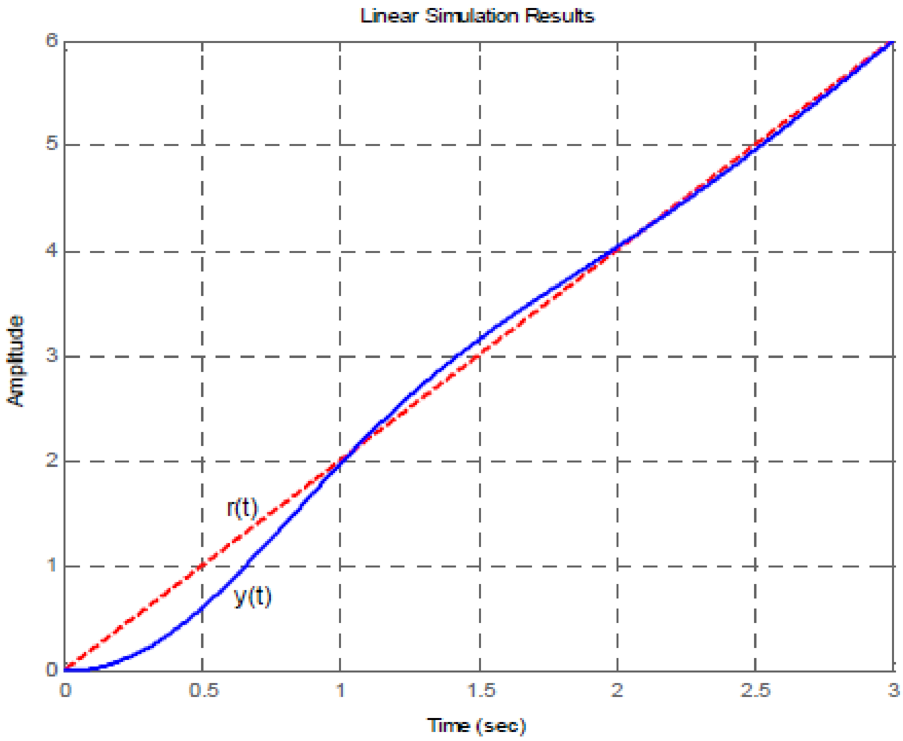


Figure 6-7 Response fora ramp input $r(t) = 2t$

Example 6-8: We find the response $y\ (t)$ of the previous Example 6-7 if the input is shown by the Figure below.

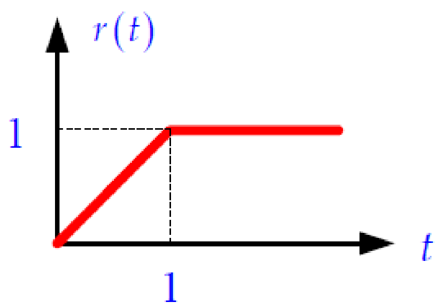


Figure 6-8 Arbitrary input

MATLAB Program 6-13

```
>> m=10; b=20; k=100;
>> num=[b  k]; den=[m  b  k];
>> sys=tf(num,den); t=[0:0.001:5];
>> for k=1:length(t)
>> if t(k) <= 1
     r(k) =t(k);
     else
     r(k)=1;
     end
     end
>> y=lsim(sys,r,t);
>> plot(t,y,t,r,'r:');
>>grid;
>> xlabel('Time (sec)');
>> gtext('r(t)');gtext('y(t)')
```

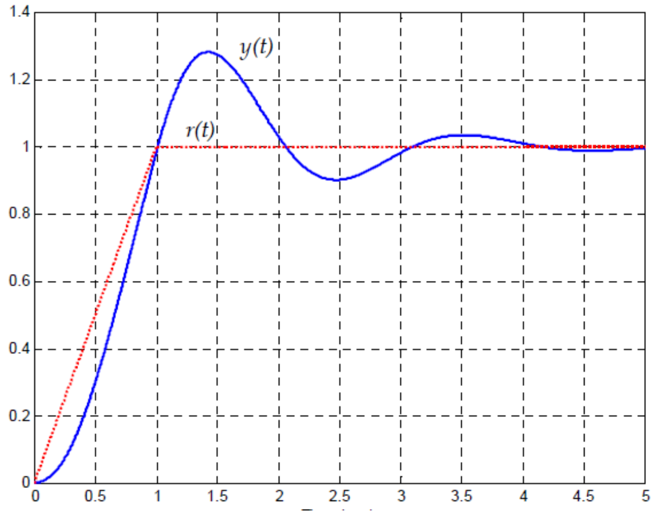


Figure 6-9 Response for an arbitrary input

6-6 Block Diagram Algebra Using MATLAB

MATLAB can be used to perform block diagram algebra if all the gains and transfer function coefficients have numerical values. You can combine blocks in series or in feedback loops using the `series` and `feedback` functions to obtain the transfer function and the state-space model.

If the LTI models `sys1` and `sys2` represent blocks in series, their combined transfer function can be obtained by typing `sys3 = series(sys1,sys2)`. A simple gain need not be converted to a LTI model, and does not require the `series` function. For example, if the first system is a simple gain K, use the multiplication symbol * and enter

```
>>sys3 = K*sys2
```

If the LTI model `sys2` is in a negative feedback loop around the LTI model sys1, then enter

```
>>sys3 = feedback(sys1,sys2)
```

to obtain the LTI model of the closed-loop system.

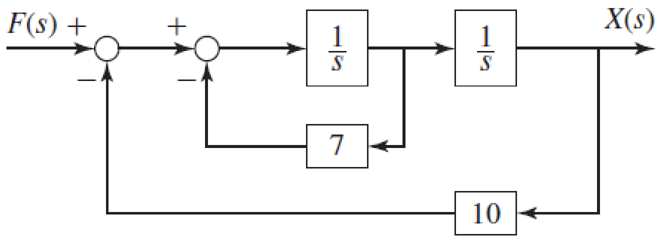


Figure 6-10 A typical block diagram.

If the feedback loop is positive, use the syntax

```
>>sys3 = feedback(sys1,sys2,+1)
```

If you need to obtain the numerator and denominator of the closed-loop transfer function, you can use the `tfdata` function and enter

```
>> [num,den] = tfdata(sys3,'v')
```

You can then find the characteristic roots by entering

```
>> roots(den)
```

Example 6-9: [58] Find the transfer function $X(s)/F(s)$ corresponding to the block diagram shown in Figure 6-10.

Solution. you enter

MATLAB Program 6-14

```
>> sys1=tf(1,[1,0]); sys2=feedback(sys1,7);
>> sys3=series(sys1,sys2); sys4=feedback(sys3,10);
>> [num,den]=tfdata(sys4,'v')
num =   0      0      1
den =   1      7     10
```

The result is num = [0, 0, 1] and den = [1, 7, 10], which corresponds to

$$\frac{X(s)}{F(s)} = \frac{1}{s^2 + 7s + 10}$$

6-7 State-Variable models with MATLAB

The MATLAB `step`, `impulse`, and `lsim` functions, can also be used with state-variable models. However, the `initial` function, which computes the free response, can be used only with a state-variable model. MATLAB also provides functions for converting models between the state-variable and transfer function forms.

Recall that to create an LTI object from the reduced form

$$5\ddot{x} + 7\dot{x} + 4x = f(t)$$

or the transfer function form

$$\frac{X(s)}{F(s)} = \frac{1}{5s^2 + 7s + 4}$$

you use the `tf(num,den)` function by typing:

```
>>sys1 = tf(1, [5, 7, 4]);
```

The result, `sys1`, is the LTI object that describes the system in the transfer function form.

The LTI object `sys2` in transfer function form for the equation

$$8\frac{d^3x}{dt^3} - 3\frac{d^2x}{dt^2} + 5\frac{dx}{dt} + 6x = 4\frac{d^2f}{dt^2} + 3\frac{d^2f}{dt} + 5f$$

is created by typing

```
>>sys2 = tf([4, 3, 5],[8, -3, 5, 6]);
```

6-7-1 LTI Objects And The ss(A,B,C,D) Function

To create an LTI object from a state model, you use the `ss(A,B,C,D)` function, where `ss` stands for state space. The matrix arguments of the function are the matrices in the following standard form of a state model:

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

$$\mathbf{y} = C\mathbf{x} + D\mathbf{u}$$

where **x** is the vector of state variables, **u** is the vector of input functions, and **y** is the vector of output variables. For example, to create an LTI object in state-model form for the system described by

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

$$\dot{x}_2 = \frac{1}{5}f(t) - \frac{4}{5}x_1 - \frac{7}{5}x_2$$

where x_1 is the desired output, the required matrices are

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{4}{5} & -\frac{7}{5} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{5} \end{bmatrix}, \qquad C = [1 \quad 0],$$

In MATLAB you type

```
>>A = [0, 1;-4/5, -7/5];
>>B = [0; 1/5];
>>C = [1, 0];
>>D = 0;
>>sys3 = ss(A,B,C,D);
```

6-7-2 The ss(sys) And ssdata(sys) Functions

An LTI object defined using the `tf` function can be used to obtain an equivalent state model description of the system. To create a state model for the system described by the LTI object `sys1` created previously in transfer function form, you type `ss(sys1)`.You will then see the resulting **A**, **B**, **C**, and **D** matrices on the screen. To extract and save the matrices as `A1`, `B1`, `C1`, and `D1` (to avoid overwriting the matrices from the second example here), use the `ssdata` function as follows.

```
>> [A1, B1, C1, D1] = ssdata(sys1);
```

The results are

$$A1 = \begin{bmatrix} -1.4 & -0.8 \\ 1 & 0 \end{bmatrix}, \qquad B1 = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}$$

$$C1 = [0 \quad 0.4], \qquad D1 = [0]$$

which correspond to the state equations:

$$\dot{x}_1 = -1.4x_1 - 0.8x_2 + 0.5f(t)$$

$$\dot{x}_2 = x_1$$

and the output equation $y = 0.4x_2$.

6-7-3 Relating State Variables To The Original Variables

When using `ssdata` to convert a transfer function form into a state model, note that the output y will be a scalar that is identical to the solution variable of the reduced form; in this case the solution variable of Equation (6-9) is the variable x. To interpret the state model, we need to relate its state variables x_1 and x_2 to x. The values of the matrices **C1** and **D1** tell us that the output variable is $y = 0.4x_2$. Because the output y is the same as x, we then see that $x_2 = x/0.4 = 2.5x$. The other state-variable x_1 is related to x_2 by the second state equation $\dot{x}_2 = x_1$. Thus $x_1 = 2.5\,\dot{x}$.

6-7-4 The `tfdata` Function

To create a transfer function description of the system `sys3`, previously created from the state model, you type $tfsys3 = tf(sys3)$. However, there can be situations where we are given the model `tfsys3` in transfer function form and we need to obtain the numerator and denominator. To extract and save the coefficients of the transfer function, use the `tfdata` function as follows.

```
_[num, den] = tfdata(tfsys3, 'v');
```

The optional parameter 'v' tells MATLAB to return the coefficients as vectors if there is only one transfer function; otherwise, they are returned as cell arrays.

For this example, the vectors returned are `num = [0, 0, 0.2]` and `den =[1, 1.4, 0.8]`. This corresponds to the transfer function

$$\frac{X(s)}{F(s)} = \frac{0.2}{s^2 + 1.4s + 0.8} = \frac{1}{5s^2 + 7s + 4}$$

which is the correct transfer function, as seen from Eq. (6-59).

Example 6-10: [58] Consider the two-mass system shown in Figure 6-11. Suppose the parameter values are $m_1 = 5, m_2 = 3, c_1 = 4, c_2 = 8, k1 = 1, and\ k_2 = 4$. The equations of motion are

$$5\ddot{x}_1 + 12\dot{x}_1 + 5x_1 - 8\dot{x}_2 - 4x_2 = 0 \qquad (6-31)$$

$$3\ddot{x}_2 + 8\dot{x}_2 + 4x_2 - 8\dot{x}_1 - 4x_1 = f(t) \qquad (6-32)$$

Put these equations into state-variable form. And obtain the transfer functions $X_1(s)/F(s)$ and $X_2(s)/F(s)$ of the state-variable model.

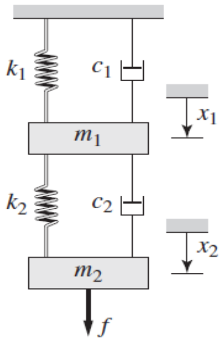


Figure 6-11 A two-mass system.

Solution. Using the system's potential and kinetic energies as a guide, we see that the displacements x_1 and x_2 describe the system's potential energy and that the velocities $\dot{x}_1$ and $\dot{x}$ describe the system's kinetic energy. That is

$$PE = \frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2(x_1 - x_2)^2$$

and

$$KE = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2$$

This indicates that we need four state variables. (Another way to see that we need four variables is to note that the model consists of two coupled second-order equations, and thus is effectively a fourth-order model.) Thus, we can choose the state variables to be

$$x_1 \quad x_2 \quad x_3 = \dot{x}_1 \quad x_4 = \dot{x}_2 \qquad (6-33)$$

Thus, two of the state equations are $\dot{x}_1 = x_3$ and $\dot{x}_2 = x_4$. The remaining two equations can be found by solving equations (6-31) and (6-32) for $\ddot{x}_1$ and $\ddot{x}_2$, noting that $\ddot{x}_1 = x_3$ and $\ddot{x}_2 = x_4 =$, and using the substitutions given by equation (6-33).

$$\dot{x}_3 = 1/5\,(-12x_3 - 5x_1 + 8x_4 + 4x_2)$$
$$\dot{x}_4 = 1/3\,\big(-8x_4 - 4x_2 + 8x_3 + 4x_1 + f(t)\big)$$

Note that the left-hand sides of the state equations must contain only the first-order derivative of each state variable. This is why we divided by 5 and 3, respectively. Note also that the right-hand sides must not contain any derivatives of the state variables. Failure to observe this restriction is a common mistake.

Now list the four state equations in ascending order according to their left-hand sides, after rearranging the right-hand sides so that the state variables appear in ascending order from left to right.

$$\dot{x}_1 = x_3$$
$$\dot{x}_2 = x_4$$
$$\dot{x}_3 = 1/5\,(-5x_1 + 4x_2 - 12x_3 + 8x_4)$$
$$\dot{x}_4 = 1/3\,\big(4x_1 - 4x_2 + 8x_3 - 8x_4 + f(t)\big)$$

These are the state equations in standard form.

The matrices and state vector of the model are

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 4/5 & -12/5 & 8/5 \\ 4/3 & -4/3 & 8/3 & -8/3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/3 \end{bmatrix}$$

and

$$Z = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}$$

Because we want the transfer functions for x_1 and x_2, we must define the **C** and **D** matrices to indicate that z_1 and z_3 are the output variables y_1 and y_2. Thus,

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \ D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The MATLAB program is as follows.

MATLAB Program 6-15

```
>> A=[0  0  1  0;0  0  0  1;-1  4/5  -12/5  8/5;4/3  -4/3  8/3  -8/3];
>> B=[0;0;0;1/3]; C=[1 0 0 0;0 1 0 0]; sys4=ss(A,B,C,D);
>> D=[0;0];
>> sys4=ss(A,B,C,D);
>> tfsys4=tf(sys4)
 Transfer function from input to output…
                  0.5333 s + 0.2667
 #1:  ---------------------------------------------
      s^4 + 5.067 s^3 + 4.467 s^2 + 1.6 s + 0.2667
              0.3333 s^2 + 0.8 s + 0.3333
 #2:  ---------------------------------------------
      s^4 + 5.067 s^3 + 4.467 s^2 + 1.6 s + 0.2667
```

The results displayed on the screen are labeled #1 and #2. These correspond to the first and second transfer functions in order. The answers are

$$\frac{X_1(s)}{F(s)} = \frac{0.5333s + 0.2667}{s^4 + 5.067s^3 + 4.467s^2 + 1.6s + 0.2667}$$

$$\frac{X_2(s)}{F(s)} = \frac{0.333s^2 + 0.8s + 0.3333}{s^4 + 5.067s^3 + 4.467s^2 + 1.6s + 0.2667}$$

Table 6-4 summarizes these functions.

Table 6-4 LTI object functions.

Command	Description
`sys = ss(A, B, C, D)`	Creates an LTI object in state-space form, where the matrices A,B, C, and D correspond to those in the model $\dot{x} = Ax + Bu, y = Cx + Du$.
`[A, B, C, D] = ssdata(sys)`	Extracts the matrices A, B, C, and D of the LTI object sys,corresponding to those in the model $\dot{x} = Ax + Bu, y = Cx + Du$.

`Sys = tf(num,den)`	Creates an LTI object in transfer function form, where the vector num is the vector of coefficients of the transfer function numerator, arranged in descending order, and den is the vector of coefficients of the denominator, also arranged in descending order.
`Sys2=tf(sys1)`	Creates the transfer function model `sys2` from the state model `sys1`.
`Sys1=ss(sys2)`	Creates the state model `sys1` from the transfer function model `sys2`.
`[num, den] = tfdata(sys, 'v')`	Extracts the coefficients of the numerator and denominator of the transfer function model `sys`. When the optional parameter '`v`' is used, if there is only one transfer function, the coefficients are returned as vectors rather than as cell arrays.

6-7-5 Linear ODE Solvers

The Control System Toolbox provides several solvers for linear models. These solvers are categorized by the type of input function they can accept: zero input, impulse input, step input, and a general input function.

6-7-6 The `initial` Function

The `initial` function computes and plots the free response of a state model. This is sometimes called the initial condition response or the undriven response in the MATLAB documentation. The basic syntax is `initial(sys,x0)`, where sys is the LTI object in state variable form, and x_0 is the initial condition vector. The time span and number of solution points are chosen automatically.

Example 6-11: [58] Compute the free response $x_1(t)$ and $x_2(t)$ of the state model derived in Example 6-10, for $x_1(0) = 5, \dot{x}_1(0) = -3, x_2(0) = 1, and\ \dot{x}_2(0) = 2$. the model is

$$\dot{x}_1 = x_3$$
$$\dot{x}_2 = x_4$$
$$\dot{x}_3 = \frac{1}{5}(-5x_1 + 4x_2 - 12x_3 + 8x_4)$$
$$\dot{x}_4 = \frac{1}{3}[4x_1 - 4x_2 + 8x_3 - 8x_4 + f(t)]$$

or

$$\dot{x} = Ax + Bf(t)$$

where

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & \frac{4}{5} & -\frac{12}{5} & \frac{8}{5} \\ \frac{4}{3} & -\frac{4}{3} & \frac{8}{3} & -\frac{8}{3} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}$$

and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}$$

Solution. We must first relate the initial conditions given in terms of the original variables to the state variables. From the definition of the state vector **x**, we see that $x_1(0) = 5, x_2(0) = 1, x_3(0) = -3, x_4(0) = 2$. Next we must define the model in state-variable form. The system `sys4` created in Example 6-10 specified two outputs, x_1 and x_2. Because we want to obtain only one output here (x_1), we must create a new state model using the same values for the **A** and **B** matrices, but now using

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The MATLAB program is as follows.

MATLAB Program 6-16

```
>> A=[0 0 1 0;0 0 0 1;-1 4/5 -12/5 8/5;4/3 -4/3 8/3 -8/3];
>> B=[0;0;0;1/3]; C=[1 0 0 0;0 1 0 0]; D=[0;0];
>> sys5=ss(A,B,C,D); initial(sys5,[5,1,-3,2])
```

The plot of $x_1(t)$ and $x_2(t)$ will be displayed on the screen.

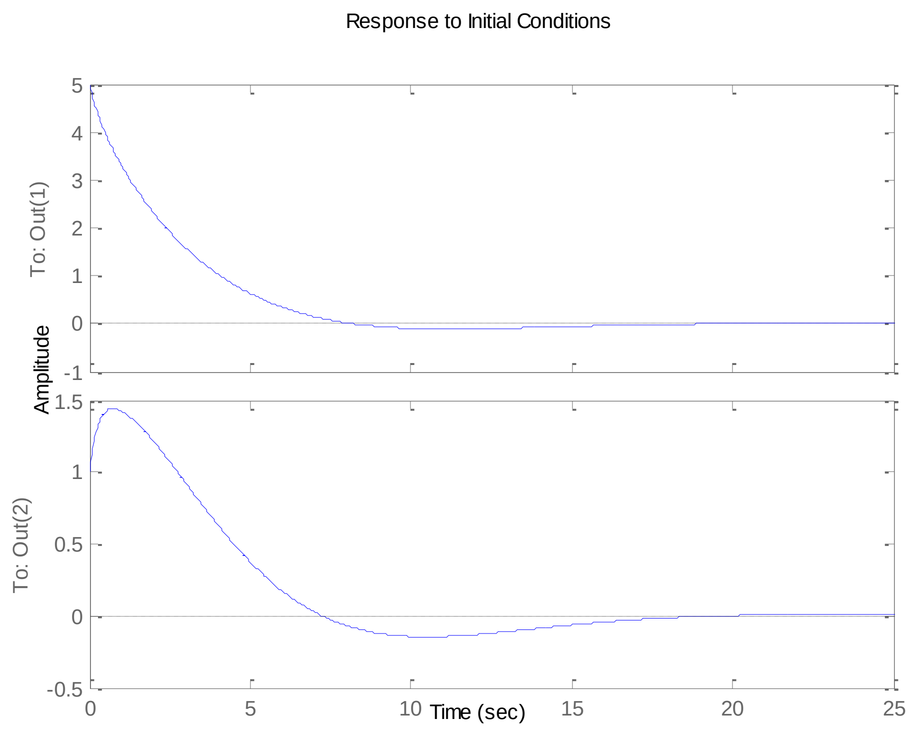


Figure 6-12 Response for Example 6-11 plotted with the `initial` function.

To plot x_1 and x_2 on the same plot you can replace the last line with the following two lines.

MATLAB Program 6-17

```
>> A=[0 0 1 0;0 0 0 1;-1 4/5 -12/5 8/5;4/3 -4/3 8/3 -8/3];
>> B=[0;0;0;1/3]; C=[1 0 0 0;0 1 0 0]; D=[0;0];
>> sys5=ss(A,B,C,D); [y,t]=initial(sys5,[5,1,-3,2]);
>> plot(t,y),gtext('x_1'),gtext('x_2'),xlabel('t')
```

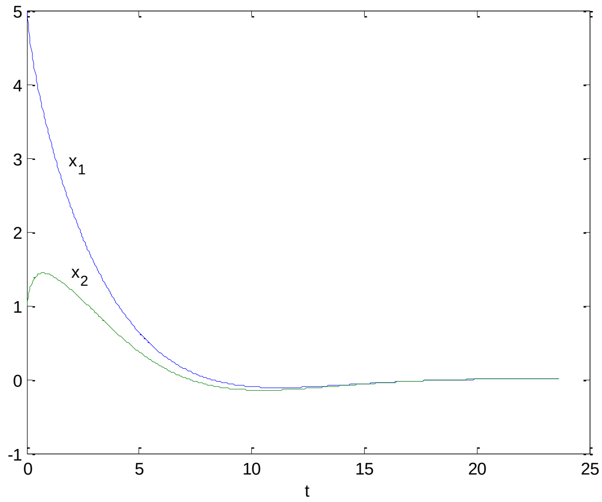


Figure 6-13 Response for Example 6-11 plotted with the plot function.

To specify the final time `tfinal`, use the syntax `initial(sys,x0,tfinal)`. To specify a vector of times of the form `t = (0:dt:tfinal)`, at which to obtain the solution, use the syntax `initial(sys,x0,t)`. When called with left-hand arguments, as `[y, t, x] = initial(sys,x0, ...)`, the function returns the output response y, the time vector t used for the simulation, and the state vector x evaluated at those times. The columns of the matrices y and x are the outputs and the states, respectively. The number of rows in y and x equals `length(t)`. No plot is drawn. The syntax `initial(sys1,sys2, ...,x0,t)` plots the free response of multiple LTI systems on a single plot. The time vector t is optional. You can specify line color, line style, and marker for each system; for example, `initial(sys1,'r',sys2,'y--',sys3,'gx',x0)`.

6-7-7 The `impulse`, `step`, And `lsim` Functions

You may use the `impulse`, `step`, and `lsim` functions with state-variable models the same way they are used with transfer function models. However, when used with state-variable models, there are some additional features available, which we illustrate with the `step` function. When called with left-hand arguments, as `[y, t] = step(sys, ...)`, the function returns the output response y and the time vector t used for the simulation. No plot is drawn. The array y is *(p × q × m)*, where *p* is `length(t)`, *q* is the number of outputs, and *m* is the number of inputs. To obtain the state vector solution for state-space models, use the syntax `[y, t, x] = step(sys, ...)`.

To use the `lsim` function for nonzero initial conditions with a state-space model, use the syntax `lsim(sys,u,t,x0)`. The initial condition vector x_0 is needed only if the initial conditions are nonzero.

These functions are summarized in Table 6-5

Table 6-5 Basic syntax of linear solvers for state variable models.

Command	Description
`initial(sys,x0,tfinal)`	Generates a plot of the free response of the state variable model `sys`, for the initial conditions specified in the array `x0`. The final time `tfinal` is optional.
`initial(sys,x0,t)`	Generates the free response plot using the user-supplied array of regularly-spaced time values t.
`[y,t,x]=initial(sys,x0, ...)`	Generates and saves the free response in the array y of the output variables, and in the array x of the state variables. No plot is produced.
`step(sys)`	Generates a plot of the unit step response of the LTI model `sys`.
`step(sys,t)`	Generates a plot of the unit step response using the user-supplied array of regularly-spaced time values t.
`[y,t]= step(sys)`	Generates and saves the unit step response in the arrays `y` and `t`. No plot is produced.
`[y,t,x]=step(sys,...)`	Generates and saves the free response in the array y of the output variables, and in the array x of the state variables, which is optional. No plot is produced.
`impulse(sys)`	Generates and plots the unit impulse response of the LTI model `sys`. The extended syntax is identical to that of the `step` function.
`lsim(sys,u,t,x0)`	Generates a plot of the total response of the state variable model `sys`. The array u contains the values of the forcing function, which must have the same number of values as the regularly-spaced time values in the array t. The initial conditions are specified in the array x0, which is optional if the initial conditions are zero.
`[y,x]= lsim(sys,u,t,x0)`	Generates and saves the total response in the array y of the output variables, and in the array x of the state variables, which is optional. No plot is produced.

Example 6-12: [58] Obtain the total response $x_1(t)$ and $x_2(t)$ of the two-mass model given in Example 6-11, using the same initial conditions but now subjected to a step input of magnitude 3.

Solution. We first define the **A**, **B**, **C**, and **D** matrices and then create the LTI model `sys`. Then we compute the step response, saving it in the arrays `ystep` and t. Note that the `step` function automatically selects a time span and a time spacing for the array t. We then use this array to compute the free response `yfree`. Finally we add the two arrays `ystep` and `yfree` to obtain the total response. Note that we could not add these two arrays if they did not have the same number of points. Note also, that if they had different time increments, we could add them, but the sum would be meaningless. The following script file shows the procedure. The resulting plot is shown Figure 6-10.

MATLAB Program 6-18

```
% InitialPlusStep.m
>> A=[0 0 1 0;0 0 0 1;-1 4/5 -12/5 8/5;4/3 -4/3 8/3 -8/3]; B=[0;0;0;1/3];
>> C=[1 0 0 0;0 1 0 0]; D=[0;0]; sys=ss(A,B,C,D);
>> [ystep,t]=step(3*sys); yfree=initial(sys,[5,1,-3,2],t);
y=yfree+ystep;
plot(t,y),xlabel('t'),gtext('x_1'),gtext('x_2')
```

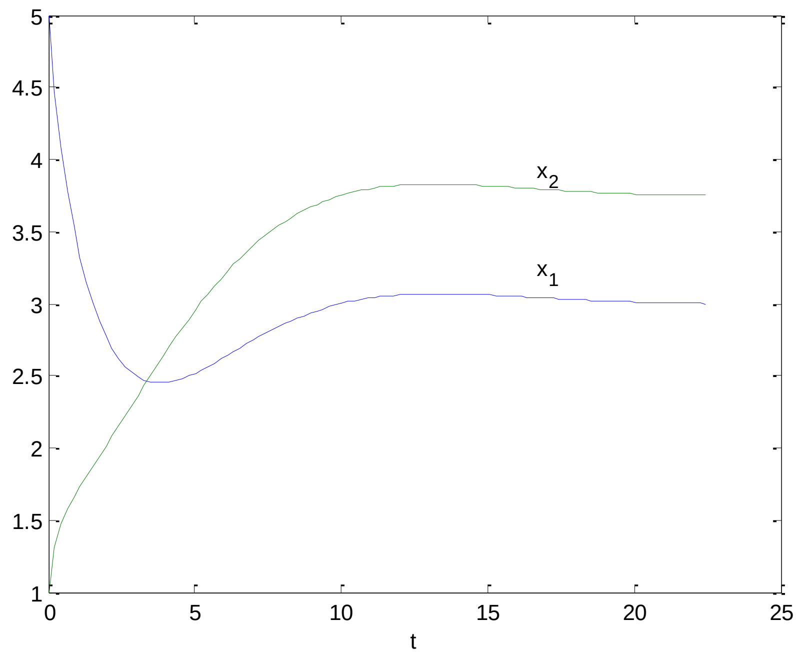


Figure 6-14 Step plus free response for Example 6-12.

CHAPTER SEVEN

Existence results

7-1 Differential Equation to State Space

Consider the transfer function with a constant numerator. We'll use a third order equation, thought it generalizes to n^th^ order in the obvious way [61].

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0}{s^3 + a_1 s^2 + a_2 s + a_3}$$

$$(s^3 + a_1 s^2 + a_2 s + a_3)Y(s) = b_0 U(s)$$

For such systems (no derivatives of the input) we can choose as our n state variables the variable y and its first n-1 derivatives (in this case the first two derivatives)

$$x_1(t) = y(t) \qquad X_1(s) = Y(s)$$

$$x_2(t) = \dot{y}(t) \qquad X_2(s) = sY(s)$$

$$x_3(t) = \ddot{y}(t) \qquad X_3(s) = s^2Y(s)$$

Taking the derivatives we can develop our state space model (which is exactly the same as when we started from the differential equation)

$$sX_1(s) = X_2(s) = sY(s)$$

$$sX_2(s) = X_3(s) = s^2Y(s)$$

$$sX_3(s) = s^3Y(s) = -a_1 s^2 Y(s) - a_2 sY(s) - a_3 Y(s) + b_0 u$$

$$= -a_1 s^2 X_3(s) - a_2 s X_2(s) - a_3 X_1(s) + b_0 u$$

$$sX(s) = AX(s) + BU(s) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} X(s) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} U(s)$$

$$Y(s) = CX(s) + DU(s) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} X(s) + 0.U(s)$$

Note: For an n^{th} order system the matrices generalize in the obvious way (**A** has ones above the main diagonal and the coefficients of the denominator polynomial for the last row, **B** is all zeros with b_0 (the numerator coefficient) in the bottom row, **C** is zero except for the leftmost element which is one, and D is zero. If we try this method on a slightly more complicated system, we find that it initially fails (though we can succeed with a little cleverness) [61].

Consider the differential equation with second derivative on the right hand side.

$$\dddot{y} + a_1\ddot{y} + a_2\dot{y} + a_3 y = b_0\ddot{u} + b_1\dot{u} + b_2 u \qquad (7-1)$$

We can try the same method as before:

$$x_1 = y$$

$$x_2 = \dot{y}$$

$$x_3 = \ddot{y}$$

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

$$\dot{x}_2 = x_3 = \ddot{y}$$

$$\dot{x}_3 = \dddot{y} = -a_3 y - a_2\dot{y} - a_1\ddot{y} + b_0\ddot{u} + b_1\dot{u} + b_2 u$$

$$= -a_3 x_1 - a_2 x_2 - a_1 x_3 + b_0\ddot{u} + b_1\dot{u} + b_2 u$$

The method has failed because there is a derivative of the input on the right hand, and that is not allowed in a state space model [60].

Fortunately we can solve our problem by revising our choice of state variables.

$$x_1 = y$$

$$x_2 = \dot{y} - b_0 u = \dot{x}_1 - b_0 u$$

$$x_3 = \ddot{y} - b_0\dot{u} - (b_1 - a_1 b_0)u = \dot{x}_2 - (b_1 - a_1 b_0)u$$

From the definition of state variables x_2 and x_3 , we have:

$$\dot{x}_1 = x_2 + b_0 u \qquad (7-2)$$

$$\dot{x}_2 = x_3 + (b_1 - a_1 b_0)u \qquad (7-3)$$

To derive the equation for $\dot{x}_3$ we first note from Equation (7-1) that

$$\dddot{y} = -a_3 y - a_2\dot{y} - a_1\ddot{y} + b_0\ddot{u} + b_1\dot{u} + b_2 u$$

Since

$$x_3 = \ddot{y} - b_0\dot{u} - (b_1 - a_1 b_0)u$$

We have

$$\dot{x}_3 = \dddot{y} - b_0\ddot{u} - (b_1 - a_1 b_0)\dot{u} \qquad (7-4)$$

$$= -a_3 y - a_2\dot{y} - a_1\ddot{y} + b_0\dddot{u} + b_1\ddot{u} + b_2\dot{u} - b_0\dddot{u} - (b_1 - a_1 b_0)\ddot{u}$$

The equations (7-2), (7-3) and (7-4) are not correct, because $\dot{y}, \ddot{y}$ are not one of the state variables. However we can make use of the fact:

$$x_2 = \dot{y} - b_0 u$$

$$x_3 = \ddot{y} - b_0\dot{u} - (b_1 - a_1 b_0)u, \quad \text{so}$$

$$\dot{y} = x_2 + b_0 u$$

$$\ddot{y} = x_3 + b_0\dot{u} + (b_1 - a_1 b_0)u$$

The state variable equations then becomes

$$\dot{x}_1 = x_2 + b_0 u$$

$$\dot{x}_2 = x_3 + (b_1 - a_1 b_0)u$$

In the third state variable equation we have successfully removed the derivative of the input from the right side of the third equation, and we can get rid of the ÿ term using the same substitution we used for the second state variable.

$$\dot{x}_3 = -a_3 y - a_2\dot{y} - a_1\ddot{y} + b_0\ddot{u} + b_1\dot{u} + b_2 u - b_0\ddot{u} - (b_1 - a_1 b_0)\dot{u}$$

$$= -a_3 x_1 - a_2(x_2 + b_0 u) - a_1(x_3 + b_0\dot{u} + (b_1 - a_1 b_0)u)$$

$$+ b_0\ddot{u} + b_1\dot{u} + b_2 u - b_0\ddot{u} - (b_1 - a_1 b_0)\dot{u}$$

$$= -a_3 x_1 - a_2 x_2 - a_1 x_3 - a_2 b_0 u - a_1 b_0\dot{u} - a_1 b_1 u + a_1^2 b_0 u$$

$$+ b_0\ddot{u} + b_1\dot{u} + b_2 u - b_0\ddot{u} - b_1\dot{u} + a_1 b_0\dot{u}$$

$$= -a_3 x_1 - a_2 x_2 - a_1 x_3 + (b_0 - b_0)\ddot{u} + (b_1 - b_1 - a_1 b_0 + a_1 b_0)\dot{u}$$

$$+ (b_2 - a_1 b_1 - a_2 b_0 + a_1^2 b_0)u$$

Hence, we get

$$\dot{x}_3 = -a_3 x_1 - a_2 x_2 - a_1 x_3 + (b_2 - a_1 b_1 - a_2 b_0 + a_1^2 b_0)u$$

or

$$\dot{x}_1 = x_2 + \alpha_0 u$$

$$\dot{x}_2 = x_3 + \alpha_1 u$$

$$\dot{x}_3 = -a_3 x_1 - a_2 x_2 - a_1 x_3 + \alpha_2 u \qquad (7-5)$$

where, $\alpha_0 = b_0,\quad \alpha_1 = b_1 - a_1 b_0,\ \ \alpha_2 = b_2 - a_1 b_1 - a_2 b_0 + a_1^2 b_0$

Combining Equations (7-2), (7-3), and (7-5) into a vector-matrix equation, we obtain Equation (7-6). Also, from the definition of state variable x_1 , we get the output equation given by Equation (7-7).

$$\dot{x}(t) = Ax(t) + Bu(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} x(t) + \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} u(t) \quad (7-6)$$

$$y(t) = Cx(t) + Du(t) = [1 \quad 0 \quad 0]x(t) + 0.u \qquad (7-7)$$

in general Consider the differential equation system that involves derivatives of the forcing function, such as

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}\dot{y} + a_n y$$
$$= b_0 u^m + b_1 u^{(m-1)} + \cdots + b_{m-1}\dot{u} + b_m u \quad n > m \qquad (7-8)$$

The main problem in defining the state variables for this case lies in the derivative terms of the input u. The state variables must be such that they will eliminate the derivatives of u in the state equation.

To obtain a state equation and output equation for this case is to define the following n variables as a set of n state variables:

$$x_1 = y$$

$$x_2 = \dot{y}$$

$$x_3 = \ddot{y}$$

$$\vdots$$

$$x_{n-(m-1)} = y^{(n-m)} - \alpha_0 u \qquad (7-9)$$

$$x_{n-(m-2)} = y^{(n-(m-1))} - \alpha_0 \dot{u} - \alpha_0 u$$

$$\vdots$$

$$x_n = y^{(n-1)} - \alpha_0 u^{(m-1)} - \alpha_1 u^{(m-2)} - \cdots - \alpha_{m-2}\dot{u} - \alpha_{m-1}u$$
$$= \dot{x}_{n-1} - \alpha_{m-1}u$$

where $\alpha_0, \alpha_1, \alpha_2 \dots, \alpha_{m-1}$are determined from

$$\alpha_0 = b_0$$

$$\alpha_1 = b_1 - a_1\alpha_0$$

$$\alpha_2 = b_2 - a_1\alpha_1 - a_2\alpha_0$$

$$\alpha_3 = b_3 - a_1\alpha_2 - a_2\alpha_1 - a_3\alpha_0 \qquad (7-10)$$

$$\vdots$$

$$\alpha_{m-1} = b_{m-1} - a_1\alpha_{m-2} - a_2\alpha_{m-3} - \cdots - a_{m-2}\alpha_1 - a_{m-1}\alpha_0$$

With the present choice of state variables, we obtain

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

$$\dot{x}_2 = x_3$$

$$\dot{x}_3 = x_4$$

$$\vdots$$

$$\dot{x}_{n-m} = x_{n-(m-1)} + \alpha_0 u$$

$$\dot{x}_{n-(m-1)} = x_{n-(m-2)} + \alpha_1 u \qquad (7-11)$$

$$\vdots$$

$$\dot{x}_{n-1} = x_n + \alpha_{m-1}u$$

$$\dot{x}_n = -a_n x_1 - a_{n-1}x_2 - \cdots - -a_1 x_n + \alpha_m u$$

where α_m is given by

$$\alpha_m = b_m - a_1\alpha_{m-1} - a_2\alpha_{m-2} - \cdots - a_{m-1}\alpha_1 - a_m\alpha_0 \qquad (7-12)$$

In terms of vector-matrix equations, Equation (7-11) and the output equation can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_{n-m} \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-m} \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ \alpha_0 \\ \vdots \\ \alpha_m \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-m} \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u$$

or

$$\dot{\mathrm{x}} = \mathrm{Ax} + \mathrm{B}u \qquad (7-13)$$

$$y = \mathrm{Cx} + \mathrm{D}u \qquad (7-14)$$

where

$$\mathrm{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{n-m} \\ \vdots \\ x_n \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ \vdots \\ \alpha_0 \\ \vdots \\ \alpha_m \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, \qquad D = 0$$

In this state-space representation, matrices A and C are exactly the same as those for the system in the case The derivatives on the right-hand side is equal to The derivatives on the left-hand side. The derivatives on the right-hand side of Equation (7-8) affect only the elements of the **B** matrix [61].

Note that the state-space representation for the transfer function

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1} s + b_m}{s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n} \qquad (7-15)$$

is given also by Equations (7-13) and (7-14).

Mechanical systems

The fundamental law governing mechanical systems is Newton’s second law

Example 7-1: [6] Consider the spring-mass-dashpot system mounted on a massless cart as shown in Figure 7–1. Let us obtain mathematical models of this system by assuming that the cart is standing still for $t < 0$ and the spring-mass-dashpot system on the cart is also standing still for $t < 0$. In this system, $u(t)$ is the displacement of the cart and is the input to the system. At $t = 0$, the cart is moved at a constant speed, or $u = constant$.The displacement $y(t)$ of the mass is the output. (The displacement is relative to the ground.) In this system, m denotes the mass, b denotes the viscous-friction coefficient, and k denotes the spring constant. We assume that the friction force of the dashpot is proportional to $\dot{y} - \dot{u}$ and that the spring is a linear spring; that is, the spring force is proportional to $y - u$.

For translational systems, Newton's second law states that

$$ma = \sum F \qquad (7-16)$$

where m is a mass, a is the acceleration of the mass, and $\sum F$ is the sum of the forces acting on the mass in the direction of the acceleration a. Applying Newton's second law to the present system and noting that the cart is massless, we obtain

$$m\frac{d^2y}{dt^2} = -b\left(\frac{dy}{dt} - \frac{du}{dt}\right) - k(y-u) \qquad (7-17)$$

or

$$m\frac{d^2y}{dt^2} + b\frac{dy}{dt} + ky = b\frac{du}{dt} + ku \qquad (7-18)$$

This equation represents a mathematical model of the system considered. Taking the Laplace transform of this last equation, assuming zero initial condition, gives

$$(ms^2 + bs + k)Y(s) = (bs + k)U(s) \qquad (7-19)$$

Taking the ratio of $Y(s)$ to $U(s)$, we find the transfer function of the system to be

$$\text{Transfer function} = G(s) = \frac{Y(s)}{U(s)} = \frac{bs + k}{ms^2 + bs + k} \qquad (7-20)$$

Such a transfer-function representation of a mathematical model is used very frequently in control engineering.

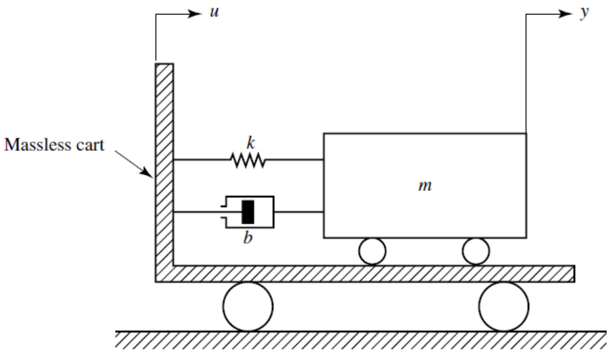


Figure 7–1 Spring-massdashpot system mounted on a cart.

Next we shall obtain a state-space model of this system. We shall first compare the differential equation for this system

$$\ddot{y} + \frac{b}{m}\dot{y} + \frac{k}{m}y = \frac{b}{m}\dot{u} + \frac{k}{m}u \qquad (7-21)$$

with the standard form

$$\ddot{y} + a_1\dot{y} + a_2 y = b_0\ddot{u} + b_1\dot{u} + b_2 u \qquad (7-22)$$

and identify a_1, a_2, b_0, b_1 , and b_2 as follows:

$$a_1 = \frac{b}{m}, \qquad a_2 = \frac{k}{m}, \qquad b_0 = 0, \qquad b_1 = \frac{b}{m}, \qquad b_2 = \frac{k}{m}$$

we have

$$\beta_0 = b_0 = 0, \qquad \beta_1 = b_1 - a_1\beta_0 = \frac{b}{m},$$
$$\beta_2 = b_2 - a_1\beta_1 - a_2\beta_0 = \frac{k}{m} - \left(\frac{b}{m}\right)^2$$

define

$$x_1 = y - \beta_0 u = y, \qquad x_2 = \dot{x}_1 - \beta_1 u = \dot{x}_1 - \frac{b}{m}u$$

With the present choice of state variables, we obtain

$$\dot{x}_1 = x_2 + \beta_1 u = x_2 + \frac{b}{m}u$$

$$\dot{x}_2 = -a_2 x_1 - a_1 x_2 + \beta_2 u = -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \left[\frac{k}{m} - \left(\frac{b}{m}\right)^2\right]u$$

and the output equation becomes

$$y = x_1$$

or

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{b}{m} \\ \frac{k}{m} - \left(\frac{b}{m}\right)^2 \end{bmatrix} u \qquad (7-23)$$

and

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (7-24)$$

Equations (7–23) and (7–24) give a state-space representation of the system. (Note that this is not the only state-space representation. There are infinitely many state-space representations for the system.)

7-2 Methods for Electrical Systems Modeling

One of the most utilized procedures in modeling electrical networks is the *mesh analysis method*, and another method is the *node analysis*; both methods are briefly presented.

7-2-1 The Mesh Analysis Method

The *mesh analysis method* is normally used for electrical networks where the input is provided by voltage sources. The voltages supplied by the sources are known, as well as the parameters defining the electrical components making up the network. In mesh analysis, therefore, Kirchhoff's second law is applied to express voltage balances for each mesh, aided by Kirchhoff's first law to relate the currents converging at common nodes. In such a network, the unknowns are usually the currents or the charges. For circuits containing only resistors, the mathematical model consists of an algebraic equations system. When the circuits also contain capacitors or inductors, the mathematical model consists of differential (or differential-integral) equations.

Example 7-2: [60] Determine the currents set up in the circuit of Figure 7-2 using the mesh analysis method. Known are $R_1 = 20\mathbf{\Omega}, R_2 = 30\mathbf{\Omega}, R_3 = 10\mathbf{\Omega}, R_4 = 40\mathbf{\Omega}, R_5 = 5\mathbf{\Omega},$ and $v = 40$.

Solution. The circuit of Figure 7-2 indicates the currents and their arbitrary directions, as well as the arbitrary positive directions for each of the two meshes (the curved arrowed lines inside each mesh). The following equations are obtained through application of Kirchhoff's second law:

$$\begin{cases} R_1 i_1 + R_4 i_3 + R_5 i_1 = V \\ R_2 i_2 + R_3 i_2 - R_4 i_3 = 0 \end{cases} \qquad (7-25)$$

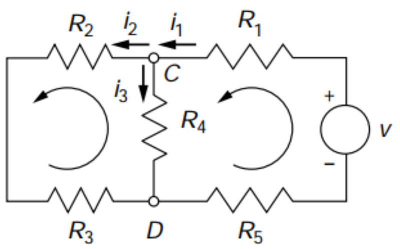


Figure 7-2 Electrical Network Comprising Resistors and a Voltage Source.

Kirchhoff's first law is applied at node *C* (or node *D*), which results in

$$i_1 = i_2 + i_3 \qquad (7-26)$$

Equations (7-25) and (7-26) are rearranged in a system of three algebraic equations where the unknowns are the currents i_1, i_2, and i_3:

$$\begin{cases} (R_1 + R_5) i_1 + R_4 i_3 = V \\ (R_2 + R_3) i_2 - R_4 i_3 = 0 \\ i_1 - i_2 - i_3 = 0 \end{cases} \qquad (7-27)$$

The equations system (7-27) can be written in vector-matrix form as

$$\begin{bmatrix} R_1 + R_5 & 0 & R_4 \\ 0 & R_2 + R_3 & -R_4 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} V \\ 0 \\ 0 \end{bmatrix} \qquad (7-28)$$

and the unknown vector containing the currents is determined by matrix algebra as

$$\begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} R_1 + R_5 & 0 & R_4 \\ 0 & R_2 + R_3 & -R_4 \\ 1 & -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} V \\ 0 \\ 0 \end{bmatrix} \qquad (7-29)$$

The solution to Eqs. (7-29) is obtained using the symbolic calculation capabilities of MATLAB® by means of the following code:

MATLAB Program 6–2

```
>> syms  r1 r2 r3 r4 r5 v
>> a = [r1+r5 0 r4;0 r2+r3 –r4;1  -1 -1]; f = [v; 0; 0];
>> i=inv(a)*f
i =
```

```
 (v*(r2 + r3 + r4))/(r1*r2 + r1*r3 + r1*r4 + r2*r4 + r2*r5 + r3*r4 + r3*r5 + r4*r5)
 (r4*v)/(r1*r2 + r1*r3 + r1*r4 + r2*r4 + r2*r5 + r3*r4 + r3*r5 + r4*r5)
 (v*(r2 + r3))/(r1*r2 + r1*r3 + r1*r4 + r2*r4 + r2*r5 + r3*r4 + r3*r5 + r4*r5)
```

After some algebraic conditioning, the returned currents are

$$\begin{cases} i_1 = \dfrac{(R_2+R_3+R_4)V}{R_1(R_2+R_3+R_4)+(R_2+R_3)(R_4+R_5)+R_4R_5} \\ i_2 = \dfrac{R_4V}{R_1(R_2+R_3+R_4)+(R_2+R_3)(R_4+R_5)+R_4R_5} \\ i_3 = \dfrac{(R_2+R_3)V}{R_1(R_2+R_3+R_4)+(R_2+R_3)(R_4+R_5)+R_4R_5} \end{cases} \qquad (7-30)$$

Numerically, the following values are obtained: $i_1 = 0.89\ A, i_2 = i_3 = 0.44\ A$.

7-2-2 The Node Analysis Method

In the *node analysis method*, voltages are associated with each node of the network where current change occurs, then nodal equations are formulated by using Kirchhoff's node law. With this method, we select one node to be the reference node, and all the voltages are expressed in terms of the reference node's voltage. Usually, it is computationally preferable to select as the reference node the one with the largest number of element branches connected to it. It is also customary to consider the voltage of the reference node to be zero.

Example 7-3: [60] Calculate the currents produced by the current and voltage sources in the circuit sketched in Figure 7-3 employing the node analysis method. Known are $R_1 = 50\ \Omega, R_2 = 70\ \Omega, R_3 = 60\ \Omega, R_4 = 40\ \Omega,\ i = 0.1\ A,$ and $v = 80\ V$.

Solution. For the circuit of Figure 7-3, if node *B* is considered to be the reference node with zero voltage, it follows that the entire line *BD* is grounded; therefore, the voltage of node *D* is also zero. Using Ohm's law, according to which a current is equal to the voltage difference across it divided by resistance, and Kirchhoff's node law, two node equations can be written when the voltages v_A and v_B are associated with nodes *A* and *B*, respectively. The equations are

$$\begin{cases} i = i_1 + i_2;\ \ or\ \ i = \dfrac{v_A - 0}{R_1} + \dfrac{v_A - v_C}{R_2} \\ i_2 = i_3 + i_4;\ \ or\ \ \dfrac{v_A - v_C}{R_2} = \dfrac{v_C - 0}{R_3} + \dfrac{v_C - v}{R_4} \end{cases} \qquad (7-31)$$

The unknowns v_A and v_C are calculated symbolically from Eq. (7-31) using MATLAB® as:

$$\begin{cases} v_A = \dfrac{R_1(R_2R_3 + R_3R_4 + R_4R_2)}{(R_1 + R_2)(R_3 + R_4) + R_3R_4} i + \dfrac{R_1R_3}{(R_1 + R_2)(R_3 + R_4) + R_3R_4} v \\ v_C = \dfrac{R_1R_3R_4}{(R_1 + R_2)(R_3 + R_4) + R_3R_4} i + \dfrac{R_3(R_1 + R_2)}{(R_1 + R_2)(R_3 + R_4) + R_3R_4} v \end{cases} \quad (7-32)$$

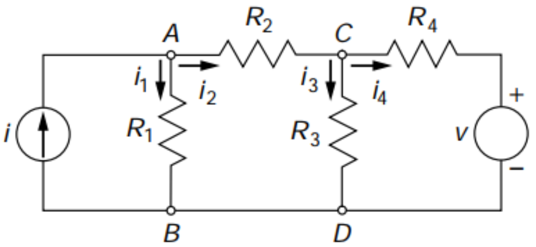


Figure 7-3 Electrical Network Comprising Resistors, a Voltage Source, and a Current Source.

The numerical values of the voltages of Eqs. (7-32) are $v_A = 19.93\,V$ and $v_C = 40.83\,V$. The four currents are determined from Eqs. (7-31) as $i_1 = 0.40A$, $i_2 = -0.30A$, $i_3 = 0.68A$, $i_4 = -0.98A$. The minus signs of i_2 and i_4 indicate that these currents have directions opposite to the ones arbitrarily chosen in Figure 7-3.

7-3 Free Response

We studies the natural response and the free damped response of electrical systems, both describing the behavior of electrical networks in the absence of voltage or current sources.

7-3-1 Natural Response

Electrical systems have a natural response when no voltage or current source is involved and no energy dissipation occurs (which translates in the absence of resistors). Electrical circuits that contain only capacitors and inductors are conservative systems; they display a natural response consisting of one or more natural frequencies, depending on the number of *degrees of freedom* (DOFs). The differential equation(s) defining the natural response of an electrical system can be derived using Kirchhoff's laws or applying the energy method, similarly to mechanical systems. MATLAB® can also be utilized to determine the eigen-frequencies and eigenvectors associated to the natural response. These methods are discussed next for single- and multiple-DOF electrical systems.

7-3-2 Single-DOF Conservative Electrical Systems

For single-DOF conservative electrical systems, the natural frequency is calculated by searching for sinusoidal (harmonical) solutions of the mathematical model differential equation.

Example 7-4: [60] Consider the LC (resonant) circuit sketched in Figure 7-4(a), which is formed of an inductor and a capacitor. Derive its mathematical model using Kirchhoff's second law and also by the energy method. Determine the natural frequency of this electrical system for $L = 1H$ and $C = 4\mu F$.

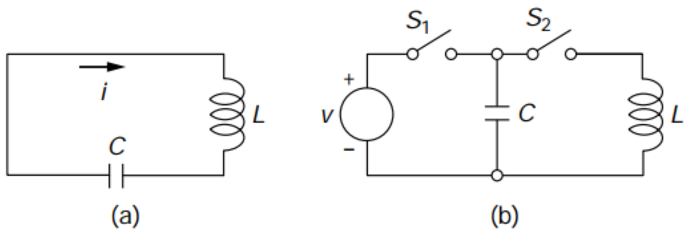


Figure 7-4 *LC* (Inductor-Capacitor) Circuit: (a) Actual Circuit; (b) Schematic for Charging the Circuit.

Solution. The single-DOF electrical system shown in Figure 7-4(a) is a conservative one, as no voltage or current sources input energy into the system and no resistors draw energy from the system. We can assume that the capacitor is charged separately from a voltage source (when switch S_1 is closed and switch S_2 is open) then disconnected from the source and connected to the inductor (by opening switch S_1 and closing S_2), as shown in Figure 7-4(b), so that the capacitor discharges on the inductor and a current *i* is produced through the circuit of Figure 7-4(a).

Application of Kirchhoff's second law to the electrical circuit of Figure 7-4(a) yields

$$L\frac{di(t)}{dt} + \frac{1}{C}\int i(t)dt = 0 \qquad (7-33)$$

which can be written in terms of the charge q as

$$L\ddot{q} + \frac{1}{C}q = 0 \qquad (7-34)$$

Or

$$\ddot{q} + \frac{1}{LC}q = 0 \qquad (7-35)$$

The electrical energy totaled by the capacitor and inductor in the circuit of Figure 7-4(a) is

$$E = \frac{1}{2}Li^2 + \frac{1}{2}\frac{q^2}{C} = \frac{1}{2}L\dot{q}^2 + \frac{1}{2}\frac{q^2}{C} \qquad (7-36)$$

Because the total electrical energy is conserved, the time derivative of the energy is zero; therefore, the following equation results from Eq. (7-36):

$$\dot{q}\left(L\ddot{q} + \frac{1}{C}q\right) = 0 \qquad (7-37)$$

The condition posed by Eq. (7-37) should be valid at all times, but the charge rate (the current) is not zero at all times; therefore, the only way that Eq. (7-37) is satisfied is when

$$L\ddot{q} + \frac{1}{C}q = 0 \qquad (7-38)$$

which is identical to Eq. (7-34): They represent the mathematical model of the electrical system shown in Figure 7-4(a). By comparing Eq. (7-35) with the generic equation that modeled the free undamped response of a single-DOF mechanical system and was of the form

$$\ddot{x} + \omega_n^2 x = 0 \qquad (7-39)$$

it follows that the natural frequency ω_n of the electrical system of Figure 7-4(a) is

$$\omega_n = \frac{1}{\sqrt{LC}} \qquad (7-40)$$

and its numerical value is $\omega_n = 500\text{rad/s}$. The free response of the electrical system consists of a harmonic (sinusoidal or cosinusoidal) vibration at the natural frequency, quite similar to the case of a spring-mass mechanical system.

Comments and conclusions

Engineering system dynamics is a discipline that focuses on deriving mathematical models based on simplified physical representations of actual systems, such as mechanical, electrical, fluid, or thermal, and on solving the mathematical models (most often consisting of differential equations).The resulting solution (which reflects the system response or behavior) is utilized in design or analysis before producing and testing the actual system. Because dynamic systems are characterized by similar mathematical models, a unitary approach can be used to characterize individual systems pertaining to different fields as well as to consider the interaction of systems from multiple fields as in coupled-field problems.

This study is a modern treatment of system dynamics and its relation to traditional mechanical engineering problems as well as modern microscale devices and machines. It provides an excellent course of study for students who want to grasp the fundamentals of dynamic systems and it covers a signifcant amount of material also taught in engineering modeling, systems dynamics, and vibrations, all combined in a dense form.

References

[1] ASTRÖM, Karl Johan; MURRAY, Richard M. Feedback systems: an introduction for scientists and engineers. Princeton university press, 2010.

[2] BAKSHI, Uday A.; BAKSHI, M. V. *Modern control theory*. Technical Publications, 2009

[3] BROGAN, William L. *Modern control theory*. Pearson education india, 1982.

[4] CARLOS, A. B. C. S.; CORRIPIO, Armando B. Principles and practice of automatic process control. *Jhon Wiley &Sons*, 1997.

[5] DORF, Richard C.; BISHOP, Robert H. Mathematical models of systems. *Modern control systems*, 2011, 49-160.

[6] Ogata, Katsuhiko, and Yanjuan Yang. *Modern control engineering*. Vol. 4. India: Prentice hall, 2002..

[7] PHILLIPS, Charles L.; PARR, John M.; RISKIN, Eve A. *Signals, systems, and transforms*. Prentice Hall, 1995.

[8] Anderson, B. D. O., and J. B. Moore, ***Linear Optimal Control***. Upper Saddle River, NJ: Prentice Hall, 1971.

[9] Athans, M., and P. L. Falb, ***Optimal Control: An Introduction to the Theory and Its Applications***. New York: McGraw-Hill Book Company, 1965.

[10] Barnet, S., “Matrices, Polynomials, and Linear Time Invariant Systems,” ***IEEE Trans. Automatic Control***, AC-18 (1973), pp. 1–10.

[11] Bayliss, L. E., ***Living Control Systems.*** London: English Universities Press Limited, 1966.

[12] Bellman, R., ***Introduction to Matrix Analysis***. New York: McGraw-Hill Book Company, 1960.

[13] Bode, H. W., ***Network Analysis and Feedback Design***. New York: Van Nostrand Reinhold, 1945.

[14] Brogan, W. L., ***Modern Control Theory***. Upper Saddle River, NJ: Prentice Hall, 1985.

[15] Butman, S., and R. Sivan (Sussman),“On Cancellations, Controllability and Observability,” ***IEEE Trans. Automatic Control***, AC-9 (1964), pp. 317–8.

[16] Campbell, D. P, ***Process Dynamics***. New York: John Wiley & Sons, Inc., 1958.

[17] Cannon, R., ***Dynamics of Physical Systems***. New York: McGraw-Hill Book Company, 1967.

[18] Chang, P. M., and S. Jayasuriya, “An Evaluation of Several Controller Synthesis Methodologies Using a Rotating Flexible Beam as a Test Bed,” ***ASME J. Dynamic Systems, Measurement, and Control***, 117 (1995), pp. 360–73.

[19] Cheng, D. K., ***Analysis of Linear Systems***. Reading, MA: Addison-Wesley Publishing Company, Inc., 1959.

[20] Churchill, R.V., ***Operational Mathematics***, 3rd ed. New York: McGraw-Hill Book Company, 1972.

[21] Coddington, E.A., and N. Levinson, ***Theory of Ordinary Differential Equations***. New York: McGraw-Hill Book Company, 1955.

[22] Craig, J. J., ***Introduction to Robotics, Mechanics and Control***. Reading, MA:AddisonWesley Publishing Company, Inc., 1986.

[23] Cunningham, W J., ***Introduction to Nonlinear Analysis.*** New York: McGraw-Hill Book Company, 1958.

[24] Dorf, R. C., and R. H. Bishop, ***Modern Control Systems***, 9th ed. Upper Saddle River, NJ: Prentice Hall, 2001.

[25] Enns, M., J. R. Greenwood III, J. E. Matheson, and F. T. Thompson, "Practical Aspects ofState-Space Methods Part I: System Formulation and Reduction," ***IEEE Trans. Military Electronics***, MIL-8 (1964), pp. 81–93.

[26] Evans, W. R., "Graphical Analysis of Control Systems," ***AIEE Trans. Part II***, 67 (1948), pp. 547-51.

[27] Evans, W. R., "Control System Synthesis by Root Locus Method," ***AIEE Trans Part II***, 69 (1950), pp. 66–9.

[28] Evans,W. R.,"The Use of Zeros and Poles for Frequency Response or Transient Response," ***ASME Trans***. 76 (1954), pp. 1135–44.

[29] Evans, W. R., ***Control System Dynamics***. New York: McGraw-Hill Book Company, 1954.

[30] Franklin, G. F, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley Publishing Company, Inc., 1994.

[31] Friedland, B., *Control System Design*. New York: McGraw-Hill Book Company, 1986.

[32] Fu, K. S., R. C. Gonzalez, and C. S. G. Lee, Robotics: Control, Sensing, Vision, and Intelligence. New York: McGraw-Hill Book Company, 1987.

[33] Gantmacher, F. R., *Theory of Matrices*, Vols. I and II. NewYork: Chelsea Publishing Company, Inc., 1959.

[34] Gardner, M. F, and J. L. Barnes, *Transients in Linear Systems.* New York: John Wiley & Sons, Inc., 1942.

[35] Gibson, J. E., *Nonlinear Automatic Control*. New York: McGraw-Hill Book Company, 1963.

[36] Gilbert, E. G.,"Controllability and Observability in Multivariable Control Systems," *J.SIAM Control*, ser. A, **1** (1963) , pp. 128–51.

[37] Graham, D., and R. C. Lathrop, "The Synthesis of Optimum Response: Criteria and Standard Forms," *AIEE Trans. Part II*, **72** (1953), pp. 273–88.

[38] Hahn, W., *Theory and Application of Liapunov's Direct Method*. Upper Saddle River, NJ: Prentice Hall, 1963.

[39] Halmos, P. R., *Finite Dimensional Vector Spaces*. New York:Van Nostrand Reinhold, 1958.

[40] Kailath, T., *Linear Systems*. Upper Saddle River, NJ: Prentice Hall, 1980.

[41] Kalman, R. E., "Contributions to the Theory of Optimal Control," *Bol. Soc Mat. Mex.*, **5** (1960), pp. 102–19.

[42] Kalman, R. E.,"On the General Theory of Control Systems," *Proc. First Intern. Cong. IFAC, Moscow*, 1960, *Automatic and Remote Control*. London: Butterworths & Company Limited, 1961, pp. 481–92.

[43] Kalman, R. E.,"Canonical Structure of Linear Dynamical Systems," Proc. Natl. Acad. Sci., USA, **48** (1962), pp. 596–600.

[44] Kautsky, J., and N. Nichols, "Robust Pole Assignment in Linear State Feedback," *Intern. J. Control*, **41** (1985), pp 1129–55.

[45] Kreindler, E., and P. E. Sarachick, "On the Concepts of Controllability and Observability of Linear Systems," *IEEE Trans. Automatic Control*, **AC-9** (1964), pp. 129–36.

[46] Kuo, B. C., *Automatic Control Systems*, 6th ed. Upper Saddle River, NJ: Prentice Hall, 1991.

[47] LaSalle, J. P, and S. Lefschetz, *Stability by Liapunov's Direct Method with Applications*. New York: Academic Press, Inc., 1961.

[48] Levin, W. S., *The Control Handbook*. Boca Raton, FL: CRC Press, 1996.

[49] Levin, W. S. *Control System Fundamentals*. Boca Raton, FL: CRC Press, 2000.

[50] Luenberger, D. G., "An Introduction to Observers," *IEEE Trans. Automatic Control*, **AC-16** (1971), pp. 596–602.

[51] Minorsky, N., *Nonlinear Oscillations*. New York: Van Nostrand Reinhold, 1962.

[52] Monopoli, R.V.,"Controller Design for Nonlinear and Time-Varying Plants," *NASA* **CR152**, Jan., 1965.

[53] Nyquist, H., "Regeneration Theory," *Bell System Tech. J.*, **11** (1932), pp. 126–47.

[54] Ogata, K., *State Space Analysis of Control Systems*. Upper Saddle River, NJ: Prentice Hall, 1967.

[55] Ogata, K., *Solving Control Engineering Problems with MATLAB*. Upper Saddle River, NJ: Prentice Hall, 1994.

[56] Ogata, K., *Designing Linear Control Systems with MATLAB*. Upper Saddle River, NJ: Prentice Hall, 1994.

[57] Ogata, K., *System Dynamics*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2004.

[58] Palm, William John. *System dynamics*. McGraw-Hill Higher Education, 2005.

[59] Golnaraghi, Farid, and Benjamin C. Kuo. *Automatic control systems*. John Wiley & Sons CN6, 1972.

[60] Lobontiu, Nicolae. *System dynamics for engineering students: concepts and applications*. Academic Press, 2010.

[61] Abdelrahim Bashir and Musa Adam, Transfer Functions and Mathematical Modeling of Dynamic Systems with MATLAB, Phd. Thesis, Sudan University of Science and Technology (2018).