

Fundamentals of Modeling and Analyzing Engineering Systems

Philip D. Cha

James J. Rosenberg

Clive L. Dym

Harvey Mudd College, Claremont, CA



PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>

40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>

10 Stamford Road, Oakleigh, Melbourne 3166, Australia

Ruiz de Alarcón 13, 28014 Madrid, Spain

© Cambridge University Press 2000

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2000

Printed in the United States of America

Typefaces Melior 9.75/13 and Univers and Optima *System* L^AT_EX 2_ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication data

Cha, Philip D.

Fundamentals of modeling and analyzing engineering systems /

Philip D. Cha, James J. Rosenberg, Clive L. Dym.

p. cm.

ISBN 0-521-59443-X (hc.). – ISBN 0-521-59463-4 (pbk.)

1. Mathematical models. 2. System analysis. I. Rosenberg, James
J. II. Dym, Clive L. III. Title.

TA342.C45 2000

620'.001'1 – dc21

99-31144

CIP

ISBN 0 521 59443 X hardback

ISBN 0 521 59463 4 paperback

Contents

<i>Examples</i>	<i>page</i> xiii
<i>Preface</i>	xix
CHAPTER 1	
Fundamental Concepts in Mathematical Modeling	1
1.1 Systems, Modeling, and Analysis	1
1.2 Abstraction	3
1.3 Physical Dimensions and Units	4
1.4 Linearity and Superposition	5
1.5 A Gentle Introduction to Differential Equations	7
1.6 Scaling in Elementary Differential Equations	9
1.7 Balance and Conservation Laws and the System Boundary Approach	11
1.7.1 Balance Equations	12
1.7.2 Conservation Laws	12
1.7.3 Examples	13
1.8 Summary	16
1.9 Problems	17
CHAPTER 2	
Lumped-Element Modeling	18
2.1 Introduction	18
2.2 One-Dimensional Translational Mechanical Systems	19
2.2.1 The Elements Comprising Simple Mechanical Systems	19
2.2.2 Translational Springs	19
2.2.3 Translational Dampers	23
2.2.4 Mass Elements in Translational Motion	24
2.2.5 The Ideal Force and the Ideal Displacement Inputs	25
2.2.6 The Interrelationship Between Forces in Different Elements in a System: Newton's Second Law	26

2.2.7	The Interrelationship Between Deformations of Different Elements in a System: Consistency of Displacements	28
2.2.8	Simplifying Models Through Combination of Elements	29
2.2.9	Examples	32
2.2.10	Summary	37
2.3	RLC Electrical Systems	37
2.3.1	Some Electrical Basics: Charge, Voltage, and Current	37
2.3.2	The Inductive, Resistive, and Capacitive Elements	38
2.3.3	The Ideal Inductor Element	39
2.3.4	The Ideal Resistor Element	39
2.3.5	The Ideal Capacitor Element	40
2.3.6	The Ideal Current and Ideal Voltage Sources	41
2.3.7	The Interrelationship Between Currents in Different Elements in a System	42
2.3.8	The Interrelationships Between the Voltage Differences Across Elements in a System	44
2.3.9	Simplifying Models Through Combination of Elements	45
2.3.10	Examples	50
2.3.11	Summary	57
2.4	Summary	58
2.5	Appendix 2-A: Faraday's Law	58
2.6	Appendix 2-B: Thevenin and Norton Equivalents	59
2.7	Problems	61
CHAPTER 3		
Generalizing Lumped-Element Modeling		67
3.1	Introduction	67
3.2	A Framework for Unifying Lumped-Element Models	67
3.2.1	Some Common Approaches: Virtues and Shortcomings	67
3.2.2	"Basic" Linear Graph Theory	68
3.2.3	Relating Linear Graph Theory to Lumped-Element Models of Physical Systems	73
3.2.4	Manipulation of Graph Theory Rules	74
3.2.5	Examples	79
3.2.6	Summary	84
3.3	Rotational Mechanical Systems	85
3.3.1	The Basics of Rotational Mechanics	85
3.3.2	Rotational Mechanical System Elements	86
3.3.3	Torsional Springs	86
3.3.4	Torsional Damper Elements	87
3.3.5	The Mass Moment of Inertia Element	88
3.3.6	The Ideal Torque and the Ideal Angular Displacement Inputs	89
3.3.7	The Rules Governing Rotational Mechanical Systems	89
3.3.8	Examples	92
3.3.9	Summary	95

3.4	Hydraulic Systems	96
3.4.1	Basic Physics of Incompressible Fluids	96
3.4.2	Hydraulic System Elements	98
3.4.3	The Pipe Element	98
3.4.4	The Tank Element	99
3.4.5	Ideal Flow Rate and Ideal Pressure Sources	101
3.4.6	The Rules Governing the Hydraulic Model	101
3.4.7	Examples	102
3.4.8	Summary	104
3.5	Thermal Systems	105
3.5.1	Basic Concepts in Heat Transfer	105
3.5.2	Thermal System Elements	106
3.5.3	The Thermal Resistance Element	106
3.5.4	The Thermal Mass Element	106
3.5.5	Ideal Heat Transfer Rate and Ideal Temperature Inputs	107
3.5.6	The Rules Governing the Thermal Model	108
3.5.7	Examples	108
3.5.8	Summary	110
3.6	Drawing Analogies	110
3.7	Summary	117
3.8	Problems	117
 CHAPTER 4		
	First-Order System Models	125
4.1	Governing Equations for First-Order Systems	125
4.2	Canonical Form of First-Order Systems	132
4.3	Classification of Responses and Systems	135
4.4	Solution of Governing Equations	135
4.4.1	Free Response	136
4.4.2	Forced Response	138
4.5	Transient Response Specifications	144
4.6	Experimental Determination of τ	145
4.6.1	Free Response	145
4.6.2	Forced Response	145
4.7	Applications of Superposition in First-Order System Models	146
4.8	Summary	151
4.9	Problems	152
 CHAPTER 5		
	Second-Order Models of Systems	155
5.1	Governing Equations for Second-Order Systems	155
5.2	Canonical Form and Classification of Response	163
5.3	Solution of Governing Equations	166
5.3.1	Free Response	166
5.3.2	Forced Response	170

5.4	Transient Response Specifications	172
5.5	Experimental Determination of ζ	176
5.5.1	Using Free Response to Determine ζ	176
5.5.2	Using Step Response to Determine ζ	178
5.6	Summary	178
5.7	Problems	178
 CHAPTER 6		
	Laplace Transform	185
6.1	Definition of the Laplace Transform	186
6.2	Properties of the Laplace Transform	188
6.2.1	Superposition	188
6.2.2	Shift in t or Time Delay	188
6.2.3	Shift in s	189
6.2.4	Time Derivatives	189
6.2.5	Time Integral	190
6.2.6	Final Value Theorem	191
6.3	The Inverse Laplace Transform	194
6.4	Response of First-Order Systems Using the Laplace Transforms	200
6.4.1	Solution of RC Free Response	201
6.4.2	Solution of RC Step Response	202
6.4.3	Solution of RC Sinusoidal Response	203
6.5	Response of Second-Order Systems Using the Laplace Transform	204
6.5.1	Free Response	204
6.5.2	Forced Unit Step Response	208
6.6	Summary	209
6.7	Appendix 6-A: Complex Arithmetic	209
6.8	Problems	211
 CHAPTER 7		
	Frequency Response of Linear, Time-Invariant Systems	216
7.1	Transfer Functions and the Sinusoidal Steady State	217
7.1.1	Steady-State Sinusoidal Response of First-Order Systems	217
7.1.2	The Transfer Function and Its Relation to the Governing Equation	218
7.1.3	The Relationship Between $T(j\omega)$ and Sinusoidal Steady-State Response	221
7.2	Mathematical Features of the Transfer Function	224
7.2.1	Poles and Zeros	224
7.2.2	First- and Second-Order Factors and Delay Factors	224
7.3	Bode Plots	225
7.3.1	The General Frequency Response Function and Its Factors	226

7.3.2	Gain Factor: k	227
7.3.3	Integral and Derivative Factors: $1/j\omega\tau$ and $j\omega\tau$	228
7.3.4	First-Order Factor: $1 + j\omega\tau$	231
7.3.5	Quadratic Factor: $1 - (\omega/\omega_n)^2 + j2\zeta(\omega/\omega_n)$	234
7.3.6	Delay Factor: $e^{-j\omega t_0}$	237
7.4	Frequency Response of First-Order and Second-Order Systems	238
7.4.1	First-Order Low-Pass Frequency Response Function	239
7.4.2	First-Order High-Pass Frequency Response Function	239
7.4.3	Second-Order Low-Pass Frequency Response Function	240
7.4.4	Second-Order High-Pass Frequency Response Function	244
7.4.5	Second-Order Band-Pass Frequency Response Function	246
7.4.6	Resonance	248
7.4.7	Transfer Functions for Systems with Delay	250
7.4.8	Examples	251
7.5	Impedance	267
7.5.1	Passing the Transform Through the Node and Loop Rule Summations	268
7.5.2	Impedance in Electrical and Mechanical Systems	269
7.5.3	Application of Impedance Techniques	272
7.5.4	Summary	278
7.6	Summary	278
7.7	Problems	279

CHAPTER 8

State Space Formulations of Systems Problems		288
8.1	Examples of State Variables and State Equations	288
8.2	Matrix Formulation	293
8.3	Free Response and the Eigenvalue Problem	295
8.4	Stability	298
8.5	Graphical Solution	298
8.6	Forced Response and Response to a Step Input	299
8.7	Examples of State Space Formulations and Solutions	300
8.8	Phase Plane and Stability Considerations	317
8.9	Summary	319
8.10	Appendix 8-A: A Short Introduction to Matrix Manipulation	320
8.11	Problems	323

CHAPTER 9

Relating the Time Domain, Frequency Domain, and State Space		326
9.1	Introduction	326
9.2	The Pole-Zero Plot	326
9.2.1	Relating the Pole-Zero Plot to the Transfer Function, the Governing Equation, and the State Matrix Eigenvalues	327
9.2.2	Relating Pole Locations to System Parameters	328
9.2.3	Examples	329

9.3	Relating Frequency Response to Pole Location	331
9.3.1	The Relationship Between the $ T(s) $ Surface and the Frequency Response Function	331
9.3.2	Higher-Order Systems and Dominant Poles	334
9.3.3	Examples	335
9.4	Transient Response, Poles, and Frequency Response	342
9.4.1	The Relationship Between the Mathematical Form of the Free Response, Pole Location, and System Parameters	342
9.4.2	The Effect of Nondominant Poles on Transient Behavior	345
9.4.3	Examples	348
9.5	State Space Trajectories, Poles, and Transient Response	354
9.6	Summary	356
9.7	Problems	356
CHAPTER 10		
	Feedback Systems	359
10.1	Systems with Feedback	359
10.1.1	What Is Feedback?	359
10.1.2	Some Examples of Systems with Feedback	360
10.2	Representing Systems and Subsystems Using Transfer Functions	365
10.2.1	Representing Inputs, Outputs, and System Behavior in the s -Domain	365
10.2.2	Subsystem Interaction and Loading	369
10.3	Block Diagrams	372
10.3.1	Block Diagram Elements and Structure	372
10.3.2	Simplifying Block Diagrams	374
10.3.3	Examples	378
10.4	Properties of Feedback Systems	383
10.4.1	Effect of Feedback on Overall Gain	383
10.4.2	Effect of Feedback on Parameter Sensitivity	383
10.4.3	Effect of Feedback on External Disturbance or Noise	384
10.4.4	Effect of Feedback on Stability	386
10.5	Relative Stability and the Phase and Gain Margins	388
10.5.1	Relative Stability	388
10.5.2	Phase Margin and Gain Margin	388
10.6	Design of Controllers	399
10.6.1	Feedback Control Systems: General Configuration and Terminology	399
10.6.2	Proportional Control	407
10.6.3	Derivative Control and PD Control	420
10.6.4	Integral Control and PI Control	427
10.6.5	PID Control	435
10.7	Summary	439
10.8	Appendix 10-A: Feedback Circuits Based on Operational Amplifiers	440

10.8.1	Noninverting Amplifier	440
10.8.2	Inverting Amplifier	442
10.9	Appendix 10-B: DC Servomotors and Tachometers	445
10.9.1	DC Servomotors	445
10.9.2	Tachometers	450
10.10	Problems	451
<i>Index</i>		457

Fundamental Concepts in Mathematical Modeling

We will first review some rather specific technical topics that are central to the application of modeling, especially mathematical modeling. Some of these you will have heard of before, in more or less detail, but we think this a good time and place to shore up familiar ideas and to briefly introduce some new ones. They include the definition of a system, modeling, analysis, linearity, the principle of superposition, physical dimensions and units, abstraction, and basic scaling ideas for simple first- and second-order differential equations.

1.1 Systems, Modeling, and Analysis

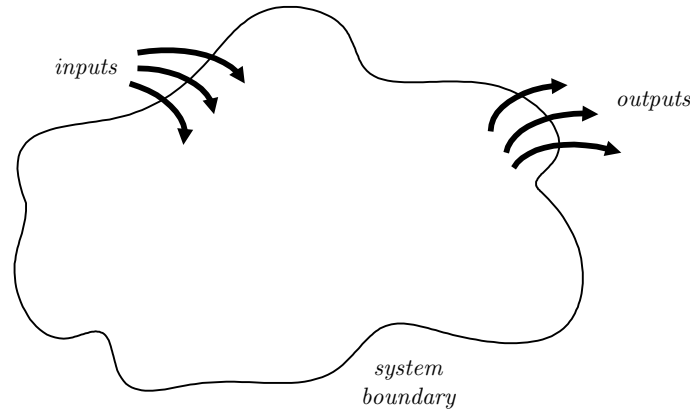
Generally, we may think of a *system* as a combination of components included inside a specified, sometimes arbitrary boundary, that interact in some way. Such a system may be naturally occurring, such as an ecosystem, or may be a human creation, such as a building or a car or a computer. This text is primarily concerned with engineering systems, which are collections of components designed to act together to perform a specific function or set of functions.

Complex engineering systems can usually be thought of in terms of simpler subsystems, each of which performs one or more of the functions required of the overall system. An automobile is a prime example of a system whose component subsystems include the body of the car, its frame, its suspension, its tires, its engine, its steering mechanism, and so on, each of which performs a specific function or functions.

The overall system and each of the subsystems has a system boundary that defines what is part of the system or subsystem and what is not. Inputs to the system and outputs from the system (or subsystem) cross this boundary (see Fig. 1.1). The system acts to transform the inputs into the outputs. In the case of an automobile, the motion of the car over an irregular road surface provides an input in the form of a time-varying vertical displacement of the wheels. An output of interest, particularly to the passengers, is the resulting time-varying vertical displacement of the passenger compartment. The suspension subsystem, consisting of tires, wheels, linkages, springs, and shock absorbers, serves the functions of

FIGURE 1.1

Schematic representation of a system and its boundary. System inputs and outputs are also illustrated.



maintaining contact between the tires and the road surface, while simultaneously transforming the sharp, rapid vertical motions of the wheels into gentle vertical motions of the passenger compartment.

The practice of engineering is centered around design. The process of design requires a number of skills, including the ability to understand the end users' needs and wishes, the ability to translate these desires and wishes into the required functions, and the ability to generate potential means to realize these functions. It also requires an ability to determine which of the proposed alternatives will best meet the end users' specifications. This requires an ability to predict how the proposed implementations will behave. The material in this text represents the first step in learning how to make predictions regarding how collections of components or systems behave.

Engineers represent information in many ways: in natural language descriptions, in pictures and drawings, and in mathematics. The primary language for making quantitative predictions regarding component and system behavior is mathematics, because mathematical analysis readily yields quantitative results. It is very important to realize, however, that there is an extremely important step in the process of predicting system behavior that must precede a mathematical analysis. This step is modeling.

In order to apply mathematical tools to predict system behavior, we must first have a mathematical representation of how system components behave. In this text, we may regard modeling as the translation of the physical behavior of components and collections of components into a mathematical representation. This representation must include descriptions of the individual components as well as descriptions of how the components interact.

Returning to the example of an automobile suspension system, we need a mathematical description of the relevant features of the behavior of suspension components such as springs. One relevant feature is the relationship between the forces exerted by the spring and the degree to which the spring is stretched or compressed. We also need a mathematical description of how the forces exerted by the various components interact. This description is provided by Newton's laws.

A good model will provide an accurate description of behavior, while at the same time remaining mathematically simple enough to permit easy calculation. Accuracy and simplicity are often at odds; consequently, modeling often involves some trade-offs. As we will describe in the section on abstraction, we may use different models for the same system, depending on precisely what kind of prediction we need to make.

Once a model has been developed, analysis can proceed. While the result of an analysis is generally a prediction of behavior, this can be broken down into more specific potential goals:

- For a given set of inputs, what are the outputs?
- In order to produce a desired set of outputs for a given set of inputs, what changes must be made to the parameters of the system?
- If some or all of the system parameters cannot be changed, what inputs must be applied in order to produce a desired set of outputs?

The first type of analysis involves a straightforward (although not necessarily easy) process. Given the system model and the resulting governing equation(s), all we need to do is to compute the response of the system for a given set of inputs. The latter two items, on the other hand, require us to solve an inverse problem. Specifically, in order to produce a set of desired system responses, we need to determine what system parameters are required or what type of inputs we need to impose. This type of problem is very common in engineering design, and generally it is a bit more challenging than a straightforward calculation of outputs for a given set of inputs and specific system parameters.

In this text, we have separated the modeling and analysis steps in this process. Our intent here is to emphasize that these are two separate procedures, both of which are required for making predictions of system behavior. The fundamentals of lumped element modeling – that is, the processes of describing system components and their interaction in mathematical language – are covered in Chapters 2 and 3. Chapters 4 through 10 emphasize analysis tools, and Chapter 10 also includes some material on design of feedback control systems.

1.2 Abstraction

The process of deciding the level of detail that is appropriate to describe the problem of interest is called *abstraction*. Abstraction typically requires a very organized and thoughtful approach to describing the phenomena upon which we wish to focus.

Consider, for example, Hooke's law, $F = kx$, for describing the behavior of a spring. This relation can be used to model more than just the relation between force and extension of a simple coiled spring. We could also use Hooke's law to describe the static response of a diving board, but the corresponding spring constant k would then need to reflect the stiffness of the diving board taken as a whole, which in turn reflects more detailed properties of the board, including the material of which it is made and its physical dimensions. The appropriate spring

constant k for the diving board can be approximated experimentally by measuring the board's tip deflection for divers of different weight. By plotting the weight of the various divers as a function of the tip deflection, we can easily calculate the slope of the force–deflection curve, which corresponds approximately to the stiffness or spring constant of the diving board.

Hooke's law can also be used to model the static behavior of a tall building, perhaps to model wind loading, perhaps as part of analyzing how the building would respond to an earthquake. We see in these examples that we can use a very simple abstract model by subsuming various details of the behavior of the building within the parameters of that model. On the other hand, this also limits the applicability of the information that we derive from applying the model. For both the diving board and the tall building we would need some very detailed expressions of how their respective stiffnesses depended on the properties of each. Without such relations, we could not do a detailed design of either the board or of the building.

Another facet of this abstraction process is that in each case we are taking some “real,” three-dimensional object and saying that, for certain well-defined purposes, it behaves like a simple spring element. We are thus introducing the concept of a lumped element wherein the actual physical properties of some real object or device are aggregated or lumped into a more abstract and less detailed expression.

Consider another example: An airplane can be modeled in very different ways, depending on our modeling goals. For example, to lay out a trajectory or a flight plan, the airplane can simply be considered as a point mass moving with respect to a spherical coordinate system. Here the mass of the point can simply be taken as the total mass of the plane, and the effect of any retarding drag force can be modeled as acting on the mass point itself with a magnitude related to the speed at which the mass is moving. If we wanted to model and analyze the more immediate, more local effects of the air movement over the plane's wings, then we obviously need a model that accounts for the wing's surface area and that is complex enough to incorporate the aerodynamics that occur in different flight regimes. If we wanted to model and design the flaps that are used to control the plane's ascent and descent, we then need a model complex enough to incorporate the system required to control the flaps and the dynamics of the wing's strength and vibration response, as well as include some of the modeling issues already accounted for. In each case, it is desirable to use the simplest model that yields a sufficient degree of accuracy in relating the inputs of interest to the outputs of interest.

1.3 Physical Dimensions and Units

A central issue in modeling is related to dimensions and units. All equations representing physical phenomena must be dimensionally compatible or consistent. We cannot, for example, equate a quantity with dimensions of volume to one having dimensions of area. This proves to be very useful, especially when we want to check the validity of a newly developed mathematical model or before we begin calculations based on formulas and equations.

Physical quantities that are used to describe or model a problem come in two types. They consist of either *fundamental* or *primary* quantities, or they are *derived* quantities. In mechanical problems, for example, mass, length, and time are generally taken as the primary mechanical variables, while force is derived from Newton's second law of motion. It is equally correct to take force, length, and time as the fundamental quantities and to derive mass from Newton's second law. However, taking a quantity as fundamental generally means that it can be assigned a standard of measurement independent of that chosen for the other fundamental quantities. For any given problem, of course, there must be a sufficient number of primary quantities so that each derived quantity can be expressed in terms of these primary quantities. Finally, we note that primary quantities are chosen arbitrarily, while the derived quantities are selected to satisfy physical laws or relevant definitions.

A *dimension* is the measure by which a physical quantity is expressed. A *unit* is a way of assigning a number to the dimension. Thus, length is a dimension chosen as the primary quantity describing such variables as distance, width and displacement, and the corresponding units can be described by meters, centimeters, or inches. The magnitude or size of the attached number to a given dimension obviously depends on the unit chosen, and this dependence often suggests a choice of units to facilitate calculation for a given problem. In our work we will typically use SI (or *Système International*) units.

1.4 Linearity and Superposition

A very important distinction in modeling is whether or not a system model is *linear*. We say that system or device models are linear when their basic equations – whether algebraic, differential, or integral – are such that the magnitude of the behavior or response produced is directly proportional to the excitation or input that drives the system. Consider a system whose governing equation is described by

$$y = f(x) \quad (1.1)$$

where y represents the *response* or *output* of the system and x denotes the *excitation* or *input* that drives the system, such that

$$x = \sum_{i=1}^N a_i x_i \quad (1.2)$$

where the a_i represent a set of constants, and the x_i represent a set of arbitrary inputs, each of which produces its own system output:

$$y_i = f(x_i) \quad (1.3)$$

The system is said to be *linear* if for the input x of Eq. (1.2), the output of Eq. (1.1) is

$$y = f(x) = f\left(\sum_{i=1}^N a_i x_i\right) = \sum_{i=1}^N a_i f(x_i) = \sum_{i=1}^N a_i y_i \quad (1.4)$$

Thus, for a linear system, we can obtain the response of that system to the sum of N inputs by adding or *superposing* the responses of the system to each individual input considered separately. This result is called the *principle of superposition*. It is also quite evident that if we multiply the input to a linear system by a constant α , the output of the system will be proportional by the same constant. Any system not satisfying the relationship of Eq. (1.4) is said to be *nonlinear*.

As a guide to recognizing whether a system is linear or not, we can look at the exponent or power to which the response or output is raised in the system equation. For a system to be linear, the exponent or power of the output must be equal to unity. Let us consider some input–output relations and see if they are linear or nonlinear. For definiteness, let us consider the constitutive equation for a classical elastic spring, of stiffness k , whose force–deflection relation is governed by the familiar Hooke’s law:

$$F = kx \quad (1.5)$$

where F represents the force exerted on the spring and x denotes the extension or relative deflection between the two ends of the spring from its static equilibrium. Clearly if we double the extension x of the spring, then the force F in the spring also doubles. Furthermore, if we impose an extension on the spring by a distance x_1 and then by an additional distance x_2 , the total force required to extend the spring by the distance $x_1 + x_2$ can be calculated by adding or superposing the results for each; that is,

$$F = k(x_1 + x_2) = kx_1 + kx_2 = F_1 + F_2 \quad (1.6)$$

Thus, Eq. (1.5) is linear. Any relation involving a transcendental function or a power greater than unity is nonlinear. Consider now the force–deflection relation of a spring of the form

$$F = k_1 x + k_2 x^3 \quad (1.7)$$

where k_1 and k_2 are constants. For this spring, it is not difficult to show that

$$F(x_1 + x_2) = k_1(x_1 + x_2) + k_2(x_1 + x_2)^3 \neq F(x_1) + F(x_2) \quad (1.8)$$

A spring whose constitutive relation is defined by Eq. (1.7) is said to be nonlinear.

The principle of superposition, which applies to linear systems, is one of the most powerful tools in system analysis. It allows us to say that the response of a system to a sum of inputs is equal to the sum of the responses of the system to the inputs taken individually. This has very deep implications for analysis. If a complicated input to a linear system can be represented as a sum of simpler inputs, then the response of the system to the simpler inputs can be calculated separately and then added to get the response of the system to the complicated input. If a linear system has a number of separate inputs, we can find the response of the system to the inputs taken one at a time, and then we can add them to get the overall response. This can be extremely useful in analyzing complicated systems having multiple inputs.

1.5 A Gentle Introduction to Differential Equations

Differential equations are of great importance in engineering and science, because many physical laws and relations appear in the form of differential equations. In addition, the most useful mathematical models used to predict the behavior of systems are commonly described in terms of differential equations. Consider a physical system whose dynamical behavior is governed by the following linear, constant coefficient differential equation of order n as follows:

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = f(t) \quad (1.9)$$

where $y(t)$ and $f(t)$ represent the output and input of the system, respectively. By definition, the order of a system corresponds to the highest derivative in the system's governing equation; a system is said to be linear if the output variable, $y(t)$, and its derivatives are raised to the first power; and a governing equation is said to have constant coefficients if $a_i = \text{constant}$, for $i = 0, \dots, n$.

For a physical system, we are generally interested in its output $y(t)$ to a given input $f(t)$. The relationship between the output and input, or the response and excitation, can be described by introducing the following *linear differential operator*:

$$D = a_n \frac{d^n}{dt^n} + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} + \cdots + a_1 \frac{d}{dt} + a_0 \quad (1.10)$$

Equation (1.10) allows us to rewrite Eq. (1.9) compactly as

$$D[y(t)] = f(t) \quad (1.11)$$

which implies that the operation D on $y(t)$ leads to Eq. (1.9). A differential operator is said to be *linear* if $D[y(t)]$ contains only the function $y(t)$ and its time derivatives raised to the first power. If $f(t) = 0$, the differential equation is said to be *homogeneous*; if $f(t) \neq 0$, it is said to be *nonhomogeneous*.

We can use the linear differential operator previously defined to introduce the principle of superposition for differential equations. Let $y_i(t)$ represent the response to an excitation $f_i(t)$. Mathematically, we can write

$$D[y_i(t)] = f_i(t) \quad (1.12)$$

Let $y(t)$ be the output or response to the sum of N inputs $\sum_{i=1}^N f_i(t)$. Because the system is linear, by summing all the equations given by Eq. (1.12), it is not difficult to show that

$$y(t) = \sum_{i=1}^N y_i(t) \quad (1.13)$$

As mentioned in Section 1.4, the principle of superposition is of great importance in the analysis of linear systems. We shall make use of it often, particularly when we want to predict the response of a system to a complicated input which can

be broken down or decomposed into a set of simpler inputs. This simplifies our calculations substantially.

The principle of superposition can also be used to decompose a complicated solution into the sum of simpler parts that have specific physical meaning. In this way the principle of superposition can also aid in the interpretation of solutions. We shall now use the superposition principle to show that the complete response to Eq. (1.9) can be expressed as

$$y(t) = y_h(t) + y_p(t) \quad (1.14)$$

where $y_h(t)$ represents the *homogeneous solution* and $y_p(t)$ represents the *particular solution*.

The homogeneous solution is obtained by solving the following homogeneous equation:

$$D[y_h(t)] = 0 \quad (1.15)$$

The homogeneous equation, in and of itself, does not have a single, unique solution. The homogeneous solution is a general solution that contains as many arbitrary constants as the order of the differential equation. Thus, a first-order differential equation has one arbitrary constant, a second-order differential equation has two arbitrary constants, and so on. We obtain a unique solution only when we also take into account the *initial conditions* – that is, the value of $y(t)$ and its first $n - 1$ derivatives (where n is the order of the differential equation) at some specific value of t , which is typically taken at $t = 0$. Taking these initial conditions into account allows us to evaluate the arbitrary constants in the homogeneous solution. This then provides a solution that is not only characteristic of the behavior of the system in a general sense, but also takes into account the specifics of the situation in which the system operates.

The particular solution is found by solving the nonhomogeneous equation

$$D[y_p(t)] = f(t) \quad (1.16)$$

The particular solution depends on the input, $f(t)$, to the system but is independent of any prescribed initial conditions. The particular solution is chosen to satisfy the differential equation (1.16) only. We defer further discussion of how to obtain $y_p(t)$ until we reach specific problems with specific inputs or excitations. Because the system is linear, it is easy to verify that Eq. (1.14) is the solution to Eq. (1.11). Summing Eqs. (1.15) and (1.16) and recalling Eq. (1.14), we recover Eq. (1.11).

Given that $y_p(t)$ is the solution of the nonhomogeneous equation [see Eq. (1.16), which has the same form as the governing equation (1.11)], we might wonder why we need to include $y_h(t)$ in our expression for the complete solution, $y(t)$. Recall that $y_p(t)$ solves the nonhomogeneous equation, but does not necessarily yield a function that meets the initial conditions. Mathematically, the addition of the homogeneous solution, $y_h(t)$, allows us to uniquely satisfy the imposed initial conditions, while still maintaining $y(t)$ as a solution to the nonhomogeneous equation.

In many problems in which time, t , is the independent variable, we are interested in the *steady-state* response, or the response as $t \rightarrow \infty$. Generally, as $t \rightarrow \infty$, the effect of the initial conditions disappears, so that the steady-state solution is typically dependent only on the particular solution:

$$\lim_{t \rightarrow \infty} y(t) = y_p(\infty) \quad (1.17)$$

When $f(t)$ remains a constant for all values of time $t \geq 0$, the particular solution remains a constant for $t \geq 0$ as well. We shall denote this steady-state or limiting value to a constant input as y_{ss} .

There are also cases where the input function is a pure sinusoid that continues indefinitely. In these cases the response of a linear system will be a pure sinusoid of the same frequency that oscillates forever. This solution also represents a steady-state solution because it is always bounded by the peak amplitudes of the sinusoid. To distinguish from the previous steady-state value to a constant input, y_{ss} , we will denote this sinusoidal steady-state response as $y_{ss}(t)$, which appears as an explicit function of time.

Thus, we see that the particular solution represents the long-term response of the system to a continuing input, while the homogeneous solution represents a passing or transient response of the system to conditions at the outset of the observation period. There are situations in engineering design and analysis when each of these types of responses – the initial response to a sudden change in input and the long-term response to an ongoing input – are of interest. The principle of superposition allows us to decompose the overall solution into homogeneous and particular solutions, providing a means for solving the initial value problem and for interpreting the results.

1.6 Scaling in Elementary Differential Equations

In this text we deal primarily with models of physical systems that are expressed as first-order or second-order differential equations. In these models the independent variable will usually be time, which we denote as t . For example, we will show in Chapter 4 that a charged capacitor draining through a resistor loses voltage $v(t)$ at a rate proportional to the actual value of the voltage at any given instant. The mathematical model would be written as

$$\frac{dv(t)}{dt} = -\frac{1}{\tau}v(t) \quad (1.18)$$

where $\tau = RC$ denotes the time constant of the system, and R and C represent the resistance and capacitance of the resistor and capacitor, respectively (these will be detailed in Chapter 4). We can rewrite this equation in an equivalent form:

$$\frac{dv(t)}{v(t)} = -\frac{1}{\tau}dt \quad (1.19)$$

Because both $dv(t)$ and $v(t)$ have units of volts, in order for Eq. (1.19) to be dimensionally compatible, the quantity τ must have physical dimensions of time – that is, units of sec. We shall confirm this in another way after we solve Eq. (1.18).

Integrating Eq. (1.19) with respect to time, t , leads to

$$\ln |v(t)| = -\frac{t}{\tau} + A_1 \quad (1.20)$$

where A_1 is the arbitrary constant of integration. If we assume that our capacitor was initially charged at a voltage v_0 – that is, $v(t = 0) = v_0$ – then it follows that $\ln v_0 = A_1$, which implies that

$$v(t) = v_0 e^{-t/\tau} \quad (1.21)$$

Equation (1.21) is said to be the solution to Eq. (1.18).

We see in Eq. (1.21) further confirmation of the dimensional nature of the quantity τ . Recall that functions such as sinusoids, logarithms, and exponentials are called transcendental functions. Using Taylor's theorem, these transcendental functions can be represented by power series, which means that their arguments must be dimensionless – without which property the power series cannot be added. The presence of the quantity τ in Eq. (1.21) serves to render the argument of the exponential function dimensionless. Additionally, the quantity τ represents a characteristic aspect of the problem being modeled, so that a ratio such as t/τ becomes a useful measure of whether a time duration is truly long or short with respect to the particular system being modeled. For the discharging capacitor, the parameter τ provides a measure of time, called the *time constant*, that characterizes the system being modeled. For example, we could define a decay time as the time it takes for the voltage to decrease to a specified fraction of its initial value. Suppose we choose that specified value to be one-tenth. The characteristic or decay time of the charged capacitor would be defined by

$$v(t_{\text{decay}}) \equiv \frac{v_0}{10} \quad (1.22)$$

which, in terms of our solution (1.12), means that

$$t_{\text{decay}} \equiv -\tau \ln \frac{1}{10} = 2.303\tau \quad (1.23)$$

Another interesting illustration of scaling and dimensionality that we see quite often is the following second-order differential equation that describes the behavior of a simple spring–mass system (see Chapter 5 for its derivation):

$$m \frac{d^2 y(t)}{dt^2} + ky(t) = 0 \quad (1.24)$$

Here $y(t)$ represents the displacement of a mass m attached to the free end (the other end of the spring is fixed) of a linear spring of stiffness k . If we divide through by the mass, we find

$$\frac{d^2 y(t)}{dt^2} + \frac{k}{m} y(t) = 0 \quad (1.25)$$

or

$$\frac{d^2 y(t)}{dt^2} + \omega_n^2 y(t) = 0 \quad (1.26)$$

where ω_n is the natural frequency of the spring-mass system represented by Eq. (1.24); that is,

$$\omega_n \equiv \sqrt{\frac{k}{m}} \quad (1.27)$$

It is clear from either Eq. (1.25) or Eq. (1.26) and from a dimensional analysis of the definition (1.27) that ω_n has the physical dimensions of 1/time. Furthermore, the product $\omega_n t$ must be dimensionless because, as we will show in Chapter 5, the solution to Eq. (1.25) is expressed in terms of transcendental functions that have $\omega_n t$ as their argument.

The time constant, τ , and the natural frequency, ω_n , are parameters that describe the behaviors of first- and second-order systems, respectively (to be discussed later in Chapters 4 and 5). They also offer another useful application in that they allow us to introduce the notion of nondimensionalized coordinates. For instance, in plotting the solution given by Eq. (1.21), we can use the normalized voltage $v(t)/v_0$ for the ordinate and t/τ for the abscissa. With proper application of the resultant plot, we can avoid having to make separate calculations for each charged capacitor draining through a resistor. Similarly, in plotting the response of a simple spring-mass system given by Eq. (1.24), we can use $\omega_n t$ for the abscissa, thereby extending the usefulness of the curves. This will be demonstrated later in Chapters 4 and 5.

1.7 Balance and Conservation Laws and the System Boundary Approach

In this section we will describe the *system boundary* approach to modeling physical systems. In this approach, a *control volume* is defined that contains a system or a part of a system. System boundary modeling is based on the simple premise that we can describe the rate of accumulation of some particular property within the control volume by accounting for all possible means by which the amount of this property can increase or decrease within the control volume.

Much of the modeling we do will be based on conservation laws, such as conservation of mass, conservation of electrical charge, and conservation of momentum. Some quantities we will discuss in this course are, to the best of our ability to determine, absolutely conserved. Examples include linear momentum, electrical charge, and (neglecting nuclear reactions) mass; these appear never to be created nor destroyed. Other quantities are *not* conserved, such as the number of animals in a population or the amount of a chemical species that may be produced or consumed by a chemical reaction. System boundary modeling can be applied to situations in which the property of interest is conserved as well as when it is not.