1

System Models, Differential Equations, and Initial-Value Problems

1.1 Introduction

The dynamical behavior of systems can be understood by studying their mathematical descriptions. The flight path of an airplane subject to certain engine thrust, rudder and elevator angles, and particular wind conditions, or the behavior of an automobile on cruise control when climbing a certain hill, can be predicted using mathematical descriptions of the pertinent behavior. Mathematical equations, typically differential or difference equations, are used to describe the behavior of processes and to predict their responses to certain inputs. Although computer simulation is an excellent tool for verifying predicted behavior, and thus for enhancing our understanding of processes, it is certainly not an adequate substitute for generating the information captured in a mathematical model, when such a model is available.

This chapter develops mathematical descriptions for linear continuous-time and linear discrete-time finite-dimensional systems. Since such systems are frequently the result of a linearization process of nonlinear systems, or the result of the modeling process of physical systems in which the nonlinear effects have been suppressed or neglected, the origins of these linear systems are frequently nonlinear systems. For this reason, here and in Chapter 4, when we deal with certain qualitative aspects (such as existence, uniqueness, continuation, and continuity with respect to parameters of solutions of system equations, stability of an equilibrium, and so forth), we consider linear as well as nonlinear system models, although the remainder of the book deals exclusively with linear systems.

In this chapter, mathematical models and classification of models are discussed in the remainder of this Introduction, Section 1.1. In Section 1.2, we provide some of the notation used and recall certain facts concerning continuous functions. In Section 1.3 we present the initial-value problem and we give several specific examples in Section 1.4. In Section 1.5 we present results that ensure the existence, continuation, and uniqueness of solutions of initial-value problems and results that ensure that the solutions of inital-value problems

depend continuously on initial conditions and system parameters. In this section we also present the Method of Successive Approximations to determine solutions of intial-value problems. The results in Section 1.5 pertain to differential equations that in general are nonlinear. In Section 1.6 we address linearization of such equations and we provide several specific examples.

We utilize the results of Section 1.5 to establish in Section 1.7 conditions for the existence, uniqueness, continuation, and continuity with respect to initial conditions and parameters of solutions of initial-value problems determined by *linear ordinary* differential equations.

In Section 1.8 we determine the solutions of linear ordinary differential equations and introduce for the first time the notions of state and state transition matrix. We also present the variations of constants formula for solving linear nonhomogeneous ordinary differential equations, and we introduce the notions of homogeneous and particular solutions.

Summarizing, the purpose of Sections 1.3 to 1.8 is to provide material dealing with ordinary differential equations and initial-value problems that is essential in the study of continuous-time finite-dimensional systems. This material will enable us to introduce the state-space equations representation of continuous-time finite-dimensional systems. This introduction will be accomplished in the next chapter.

Physical Processes, Models, and Mathematical Descriptions

A systematic study of (physical) *phenomena* usually begins with a *modeling process*. Examples of models include diagrams of electric circuits consisting of interconnections of resistors, inductors, capacitors, transistors, diodes, voltage or current sources, and so on; mechanical circuits consisting of interconnections of point masses, springs, viscous dampers (dashpots), applied forces, and so on; verbal characterizations of economic and societal systems; among others. Next, appropriate *laws* or *principles* are invoked to generate *equations* that describe the models (e.g., Kirchhoff's current and voltage laws, Newton's laws, conservation laws, and so forth). When using an expression such as "we consider a *system* described by ordinary differential equations," we will have in mind a phenomenon described by an appropriate set of ordinary differential equations (not the description of the physical phenomenon itself).

A physical process (physical system) will typically give rise to several different models, depending on what questions are being asked. For instance, in the study of the voltage-current characteristics of a transistor (the physical process), one may utilize a circuit (the model) that is valid at low frequencies or a circuit (a second model) that is valid at high frequencies; alternatively, if semiconductor impurities are of interest, a third model, quite different from the preceding two, is appropriate.

Over the centuries, a great deal of progress has been made in developing mathematical descriptions of physical phenomena (using models of such phenomena). In doing so, we have invoked laws (or principles) of physics,

chemistry, biology, economics, and so on, to derive mathematical expressions (usually equations) that characterize the evolution (in time) of the variables of interest. The availability of such mathematical descriptions enables us to make use of the vast resources offered by the many areas of applied and pure mathematics to conduct qualitative and quantitative studies of the behavior of processes. *A given model of a physical process may give rise to several different mathematical descriptions.* For example, when applying Kirchhoff's voltage and current laws to the low-frequency transistor model mentioned earlier, one can derive a set of differential and algebraic equations, a set consisting of only differential equations, or a set of integro-differential equations, and so forth. *This process of mathematical modeling, "from a physical phenomenon to a model to a mathematical description," is essential in science and engineering.* To capture phenomena of interest accurately and in tractable mathematical form is a demanding task, as can be imagined, and requires a thorough understanding of the physical process involved. For this reason, the mathematical description of complex electrical systems, such as power systems, is typically accomplished by electrical engineers, the equations of flight dynamics of an aircraft are derived by aeronautical engineers, the equations of chemical processes are arrived at by chemists and chemical engineers, and the equations that characterize the behavior of economic systems are provided by economists. In most nontrivial cases, this type of modeling process is close to an art form since *a good mathematical description must be detailed enough to accurately describe the phenomena of interest and at the same time simple enough to be amenable to analysis.* Depending on the applications on hand, a given mathematical description of a process may be further simplified before it is used in analysis and especially in design procedures. For example, using the finite element method, one can derive a set of first-order differential equations that describe the motion of a space antenna. Typically, such mathematical descriptions contain hundreds of differential equations. Whereas all these equations are quite useful in simulating the motion of the antenna, a lower order model is more suitable for the control design that, for example, may aim to counteract the effects of certain disturbances. Simpler mathematical models are required mainly because of our inability to deal effectively with hundreds of variables and their interactions. In such simplified mathematical descriptions, only those variables (and their interactions) that have significant effects on the phenomena of interest are included.

A point that cannot be overemphasized is that *the mathematical descriptions we will encounter characterize processes only approximately.* Most often, this is the case because the complexity of physical systems defies exact mathematical formulation. In many other cases, however, it is our own choice that a mathematical description of a given process approximate the actual phenomena by only a certain desired degree of accuracy. As discussed earlier, this is done in the interest of mathematical simplicity. For example, in the description of RLC circuits, one could use nonlinear differential equations that take into consideration parasitic effects in the capacitors; however, most often it

suffices to use linear ordinary differential equations with constant coefficients to describe the voltage-current relations of such circuits, since typically such a description provides an adequate approximation and since it is much easier to work with linear rather than nonlinear differential equations.

In this book it will generally be assumed that the mathematical description of a system in question is given. In other words, we assume that the modeling of the process in question has taken place and that equations describing the process are given. Our main objective will be to present a theory of an important class of systems—finite-dimensional linear systems—by studying the equations representing such systems.

Classification of Systems

For our purposes, a comprehensive classification of systems is not particularly illuminating. However, an enumeration of the more common classes of systems encountered in engineering and science may be quite useful, if for no other reason than to show that the classes of systems considered in this book, although very important, are quite specialized.

As pointed out earlier, the particular set of equations describing a given system will in general depend on the effects one wishes to capture. Thus, one can speak of *lumped parameter* or *finite-dimensional systems* and *distributed parameter* or *infinite-dimensional systems*; *continuous-time* and *discrete-time systems*; *linear* and *nonlinear systems*; *time-varying* and *time-invariant systems*; *deterministic* and *stochastic systems*; appropriate combinations of the above, called *hybrid systems*; and perhaps others.

The appropriate mathematical settings for finite-dimensional systems are finite-dimensional vector spaces, and for infinite-dimensional systems they are most often infinite-dimensional linear spaces. Continuous-time finite-dimensional systems are usually described by ordinary differential equations or certain kinds of integral equations, whereas discrete-time finite-dimensional systems are usually characterized by ordinary difference equations or discrete-time counterparts to those integral equations. Equations used to describe infinite-dimensional systems include partial differential equations, Volterra integro-differential equations, functional differential equations, and so forth. Hybrid system descriptions involve two or more different types of equations. Nondeterministic systems are described by stochastic counterparts to those equations (e.g., Ito differential equations).

In a broader context, not addressed in this book, most of the systems described by the equations enumerated generate *dynamical systems*. It has become customary in the engineering literature to use the term "dynamical system" rather loosely, and it has even been applied to cases where the original definition does not exactly fit. (For a discussion of general dynamical systems, refer, e.g., to Michel et al [5].) We will address in this book dynamical systems determined by ordinary differential equations or ordinary difference equations, considered next.

Finite-Dimensional Systems

The dynamical systems we will be concerned with are *continuous-time* and *discrete-time finite-dimensional systems*—primarily *linear systems*. However, since such systems are frequently a consequence of a linearization process, it is important when dealing with fundamental qualitative issues that we have an understanding of the origins of such linear systems. In particular, when dealing with questions of existence and uniqueness of solutions of the equations describing a class of systems, and with stability properties of such systems, we may consider nonlinear models as well.

Continuous-time finite-dimensional dynamical systems that we will consider are described by equations of the form

$$\dot{x}_i = f_i(t, x_1, \ldots, x_n, u_1, \ldots, u_m), \qquad i = 1, \ldots, n, \tag{1.1a}$$

$$y_i = g_i(t, x_1, \ldots, x_n, u_1, \ldots, u_m), \qquad i = 1, \ldots, p, \tag{1.1b}$$

where u_i, $i = 1, \ldots, m$, denote *inputs* or *stimuli*; y_i, $i = 1, \ldots, p$, denote *outputs* or *responses*; x_i, $i = 1, \ldots, n$, denote *state variables*; t denotes *time*; $\dot{x}_i$ denotes the time derivative of x_i; f_i, $i = 1, \ldots, n$, are real-valued functions of $1 + n + m$ real variables; and g_i, $i = 1, \ldots, p$, are real-valued functions of $1 + n + m$ real variables. A complete description of such systems will usually also require a set of *initial conditions* $x_i(t_0) = x_{i0}$, $i = 1, \ldots, n$, where t_0 denotes *initial time*. We will elaborate later on restrictions that need to be imposed on the f_i, g_i, and u_i and on the origins of the term "state variables."

Equations (1.1a) and (1.1b) can be represented in vector form as

$$\dot{x} = f(t, x, u), \tag{1.2a}$$

$$y = g(t, x, u), \tag{1.2b}$$

where x is the *state vector* with components x_i, u is the *input vector* with components u_i, y is the *output vector* with components y_i, and f and g are vector-valued functions with components f_i and g_i, respectively. We call (1.2a) a *state equation* and (1.2b) an *output equation*.

Important special cases of (1.2a) and (1.2b) are the *linear time-varying state equation and output equation* given by

$$\dot{x} = A(t)x + B(t)u, \tag{1.3a}$$

$$y = C(t)x + D(t)u, \tag{1.3b}$$

where A, B, C, and D are real $n \times n, n \times m, p \times n$, and $p \times m$ matrices, respectively, whose elements are time-varying. Restrictions on these matrices will be provided later.

Linear time-invariant state and output equations given by

$$\dot{x} = Ax + Bu, \tag{1.4a}$$

$$y = Cx + Du \tag{1.4b}$$

constitute important special cases of (1.3a) and (1.3b), respectively.

Equations (1.3) and (1.4) may arise in the modeling process, or they may be a consequence of *linearization* of (1.1).

Discrete-time finite-dimensional dynamical systems are described by equations of the form

$$x_i(k+1) = f_i(k, x_1(k), \ldots, x_n(k), u_1(k), \ldots, u_m(k)) \quad i = 1, \ldots, n, \tag{1.5a}$$
$$y_i(k) = g_i(k, x_1(k), \ldots, x_n(k), u_1(k), \ldots, u_m(k)) \quad i = 1, \ldots, p, \tag{1.5b}$$

or in vector form,

$$x(k+1) = f(k, x(k), u(k)), \tag{1.6a}$$
$$y(k) = g(k, x(k), u(k)), \tag{1.6b}$$

where k is an integer that denotes *discrete time* and all other symbols are defined as before. A complete description of such systems involves a set of *initial conditions* $x(k_0) = x_{k_0}$, where k_0 denotes *initial time*. The corresponding linear time-varying and time-invariant state and output equations are given by

$$x(k+1) = A(k)x(k) + B(k)u(k), \tag{1.7a}$$
$$y(k) = C(k)x(k) + D(k)u(k) \tag{1.7b}$$

and

$$x(k+1) = Ax(k) + Bu(k), \tag{1.8a}$$
$$y(k) = Cx(k) + Du(k), \tag{1.8b}$$

respectively, where all symbols in (1.7) and (1.8) are defined as in (1.3) and (1.4), respectively.

This type of system characterization is called *state-space description* or *state-variable description* or *internal description* of finite-dimensional systems. Another way of describing continuous-time and discrete-time finite-dimensional dynamical systems involves operators that establish a relationship between the system inputs and outputs. Such characterization is called *input–output description* or *external description* of a system. In Chapter 2, we will address both the state-variable description and the input–output description of finite-dimensional systems. Before we can do this, however, we will require some background material concerning ordinary differential equations.

1.2 Preliminaries

We will employ a consistent notation and use certain facts from the calculus, analysis, and linear algebra. We will summarize this type of material, as needed, in various sections. This is the first such section.

1.2.1 Notation

Let V and W be *sets*. Then $V \cup W, V \cap W, V - W$, and $V \times W$ denote the *union, intersection, difference*, and *Cartesian product* of V and W, respectively. If V is a *subset* of W, we write $V \subset W$; if x is an *element* of V, we write $x \in V$; and if x is not an element of V, we write $x \notin V$. We let $V', \partial V, \bar{V}$, and int V denote the *complement, boundary, closure*, and *interior* of V, respectively.

Let ϕ denote the *empty set*, R the *real numbers*, $R^+ = \{x \in R : x \geq 0\}$ (i.e., R^+ denotes the set of nonnegative real numbers), Z the *integers*, and $Z^+ = \{x \in Z : x \geq 0\}$.

We will let $J \subset R$ denote open, closed, or half-open *intervals*. Thus, for $a, b \in R$, $a \leq b$, J may be of the form $J = (a, b) = \{x \in R : a < x < b\}$, $J = [a, b] = \{x \in R : a \leq x \leq b\}$, $J = [a, b) = \{x \in R : a \leq x < b\}$, or $J = (a, b] = \{x \in R : a < x \leq b\}$.

Let R^n denote the real n-space. If $x \in R^n$, then

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

and $x^T = (x_1, \ldots, x_n)$ denotes the *transpose* of the vector x. Also, let $R^{m \times n}$ denote the set of $m \times n$ real matrices. If $A \in R^{m \times n}$, then

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

and $A^T = [a_{ji}] \in R^{n \times m}$ denotes the *transpose* of the matrix A.

Similarly, we let C^n denote the set of n-vectors with complex components and $C^{m \times n}$ denote the set of $m \times n$ matrices with complex elements.

Let $f : V \to W$ denote a *mapping* or *function* from a set V into a set W, and denote by $D(f)$ and $R(f)$ the *domain* and the range of f, respectively. Also, let $f^{-1} : R(f) \to D(f)$, if it exists, denote the *inverse* of f.

1.2.2 Continuous Functions

First, let $J \subset R$ denote an open interval and consider a function $f : J \to R$. Recall that f is said to be *continuous at the point* $t_0 \in J$ if $\lim_{t \to t_0} f(t) = f(t_0)$ exists; i.e., if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(t) - f(t_0)| < \epsilon$ whenever $|t - t_0| < \delta$ and $t \in J$. The function f is said to be *continuous on* J, or simply *continuous*, if it is continuous at each point in J.

In the above definition, δ depends on the choice of t_0 and ϵ; i.e., $\delta = \delta(\epsilon, t_0)$. If at *each* $t_0 \in J$ it is true that there is a $\delta > 0$, independent of t_0 [i.e., $\delta = \delta(\epsilon)$], such that $|f(t) - f(t_0)| < \epsilon$ whenever $|t - t_0| < \delta$ and $t \in J$, then f is said to be *uniformly continuous* (on J).

Let

$$C(J,R) \triangleq \{f : J \to R \mid f \text{ is continuous on } J\}.$$

Now suppose that J contains one or both endpoints. Then continuity is interpreted as being one-sided at these points. For example, if $J = [a,b]$, then $f \in C(J,R)$ will mean that $f \in C((a,b),R)$ and that $\lim_{t\to a^+} f(t) = f(a)$ and $\lim_{t\to b^-} f(t) = f(b)$ exist.

With k any positive integer, and with J an open interval, we will use the notation

$$\begin{aligned} C^k(J,R) \triangleq &\{f : J \to R \mid \text{the derivative } f^{(j)} \text{ exists on } J \text{ and} \\ &f^{(j)} \in C(J,R) \text{ for } j = 0,1,\ldots,k, \text{ where } f^{(0)} \triangleq f\} \end{aligned}$$

and we will call f in this case a C^k*-function*. Also, we will call f a *piecewise* C^k*-function* if $f \in C^{k-1}(J,R)$ and $f^{(k-1)}$ has continuous derivatives for all $t \in J$, with the possible exception of a finite set of points where $f^{(k)}$ may have jump discontinuities. As before, when J contains one or both endpoints, then the existence and continuity of derivatives is one-sided at these points.

For any subset D of the n-space R^n with nonempty interior, we can define $C(D,R)$ and $C^k(D,R)$ in a similar manner as before. Thus, $f \in C(D,R)$ indicates that at every point $x_0 = (x_{10},\ldots,x_{n0})^T \in D, \lim_{x\to x_0} f(x) = f(x_0)$ exists, or equivalently, at every $x_0 \in D$ it is true that for every $\epsilon > 0$ there exists a $\delta = \delta(\epsilon, x_0) > 0$ such that $|f(x) - f(x_0)| < \epsilon$ whenever $|x_1 - x_{10}| + \cdots + |x_n - x_{n0}| < \delta$ and $x \in D$. Also, we define $C^k(D,R)$ as

$$\begin{aligned} C^k(D,R) \triangleq &\left\{f : D \to R \,\middle|\, \frac{\partial^j f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} \in C(D,R), \quad i_1 + \cdots + i_n = j, \right. \\ &\left. j = 1,\ldots,k, \text{ and } f \in C(D,R)\right\} \end{aligned}$$

(i.e., $i_1,\ldots,i_n$ take on all possible positive integer values such that their sum is j). When D contains its boundary (or part of its boundary), then the continuity of f and the existence and continuity of partial derivatives of f, $\frac{\partial^j f}{\partial x_1^{i_1}\ldots\partial x_n^{i_n}}$, $i_1 + \cdots + i_n = j$, $j = 1,\ldots,k$, will have to be interpreted in the appropriate way at the boundary points.

Recall that if $K \subset R^n$, $K \neq \phi$, and K is *compact* (i.e., K is closed and bounded), and if $f \in C(K,R)$, then f is uniformly continuous (on K) and f attains its maximum and minimum on K.

Finally, let D be a subset of R^n with nonempty interior and let $f : D \to R^m$. Then $f = (f_1,\ldots,f_m)^T$ where $f_i : D \to R$, $i = 1,\ldots,m$. We say that $f \in C(D,R^m)$ if $f_i \in C(D,R)$, $i = 1,\ldots,m$, and that for some positive integer k, $f \in C^k(D,R^m)$ if $f_i \in C^k(D,R)$, $i = 1,\ldots,m$.

1.3 Initial-Value Problems

In this section we make precise the meaning of several concepts that arise in the study of continuous-time finite-dimensional dynamical systems.

1.3.1 Systems of First-Order Ordinary Differential Equations

Let $D \subset R^{n+1}$ denote a *domain*, i.e., an open, nonempty, and connected subset of R^{n+1}. We call R^{n+1} the (t,x)-*space*, and we denote elements of R^{n+1} by (t,x) and elements of R^n by $x = (x_1,\ldots,x_n)^T$. Next, we consider the functions $f_i \in C(D,R)$, $i = 1,\ldots,n$, and if x_i is a function of t, we let $x_i^{(n)} = \frac{d^n x_i}{dt^n}$ denote the nth derivative of x_i with respect to t (provided that it exists). In particular, when $n = 1$, we usually write

$$x_i^{(1)} = \dot{x}_i = \frac{dx_i}{dt}.$$

We call the system of equations given by

$$\dot{x}_i = f_i(t, x_1, \ldots, x_n), \quad i = 1,\ldots,n, \tag{1.9}$$

a *system of n first-order ordinary differential equations*. By a *solution* of the system of equations (1.9), we shall mean n continuously differentiable functions $\phi_1,\ldots,\phi_n$ defined on an interval $J = (a,b)$ [i.e., $\phi \in C^1(J,R^n)$] such that $(t,\phi_1(t),\ldots,\phi_n(t)) \in D$ for all $t \in J$ and such that

$$\dot{\phi}_i(t) = f_i(t,\phi_1(t),\ldots,\phi_n(t)), \quad i = 1,\ldots,n,$$

for all $t \in J$.

Next, we let $(t_0, x_{10},\ldots,x_{n0}) \in D$. Then the *initial-value problem* associated with (1.9) is given by

$$\begin{aligned} \dot{x}_i &= f_i(t,x_1,\ldots,x_n), \quad && i = 1,\ldots,n, \\ x_i(t_0) &= x_{i0}, && i = 1,\ldots,n. \end{aligned} \tag{1.10}$$

A set of functions $\{\phi_1,\ldots,\phi_n\}$ is a *solution* of the initial-value problem (1.10) if $\{\phi_1,\ldots,\phi_n\}$ is a solution of (1.9) on some interval J containing t_0 and if $(\phi_1(t_0),\ldots,\phi_n(t_0)) = (x_{10},\ldots,x_{n0})$.

In Figure 1.1 the solution of a hypothetical initial-value problem is depicted graphically when $n = 1$. Note that $\dot{\phi}(\tau) = f(\tau,\tilde{x}) = \tan\alpha$, where α is the slope of the line L that is tangent to the plot of the curve $\phi(t)$ vs. t, at the point $(\tau,\tilde{x})$.

In dealing with systems of equations, we will utilize the vector notation $x = (x_1,\ldots,x_n)^T$, $x_0 = (x_{10},\ldots,x_{n0})^T$, $\phi = (\phi_1,\ldots,\phi_n)^T$, $f(t,x) = (f_1(t,x_1,\ldots,x_n),\ldots,f_n(t,x_1,\ldots,x_n))^T = (f_1(t,x),\ldots,f_n(t,x))^T$, $\dot{x} = (\dot{x}_1,\ldots,\dot{x}_n)^T$, and $\int_{t_0}^{t} f(s,\phi(s))ds = [\int_{t_0}^{t} f_1(s,\phi(s))ds,\ldots,\int_{t_0}^{t} f_n(s,\phi(s))ds]^T$.

With the above notation we can express the system of first-order ordinary differential equations (1.9) by

$$\dot{x} = f(t,x) \tag{1.11}$$

and the initial-value problem (1.10) by

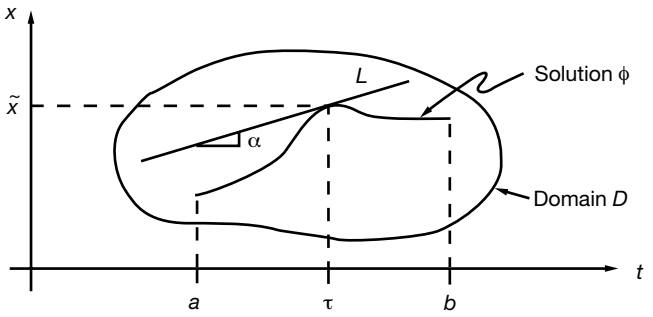

Figure 1.1. Solution of an initial-value problem when $n = 1$

$$\dot{x} = f(t, x), \quad x(t_0) = x_0. \tag{1.12}$$

We leave it to the reader to prove that the initial-value problem (1.12) can be equivalently expressed by the *integral equation*

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds, \tag{1.13}$$

where ϕ denotes a solution of (1.12).

1.3.2 Classification of Systems of First-Order Ordinary Differential Equations

Systems of first-order ordinary differential equations have been classified in many ways. We enumerate here some of the more important cases.

If in (1.11), $f(t, x) \equiv f(x)$ for all $(t, x) \in D$, then

$$\dot{x} = f(x). \tag{1.14}$$

We call (1.14) an *autonomous system* of first-order ordinary differential equations.

If $(t + T, x) \in D$ whenever $(t, x) \in D$ and if $f(t, x) = f(t + T, x)$ for all $(t, x) \in D$, then (1.11) assumes the form

$$\dot{x} = f(t, x) = f(t + T, x). \tag{1.15}$$

We call such an equation a *periodic system* of first-order differential equations with *period* T. The smallest $T > 0$ for which (1.15) is true is called the *least period* of this system of equations.

When in (1.11), $f(t, x) = A(t)x$, where $A(t) = [a_{ij}(t)]$ is a real $n \times n$ matrix with elements a_{ij} that are defined and at least piecewise continuous on a t-interval J, then we have

$$\dot{x} = A(t)x \tag{1.16}$$

and refer to (1.16) as a *linear homogeneous system* of first-order ordinary differential equations.

If for (1.16), $A(t)$ is defined for all real t, and if there is a $T > 0$ such that $A(t) = A(t+T)$ for all t, then we have

$$\dot{x} = A(t)x = A(t+T)x. \tag{1.17}$$

This system is called a *linear periodic system* of first-order ordinary differential equations.

Next, if in (1.11), $f(t,x) = A(t)x + g(t)$, where $A(t)$ is as defined in (1.16), and $g(t) = [g_1(t), \ldots, g_n(t)]^T$ is a real n-vector with elements g_i that are defined and at least piecewise continuous on a t-interval J, then we have

$$\dot{x} = A(t)x + g(t). \tag{1.18}$$

In this case we speak of a *linear nonhomogeneous system* of first-order ordinary differential equations.

Finally, if in (1.11), $f(t,x) = Ax$, where $A = [a_{ij}] \in R^{n\times n}$, then we have

$$\dot{x} = Ax. \tag{1.19}$$

This type of system is called a *linear, autonomous, homogeneous* system of first-order ordinary differential equations.

1.3.3 *n*th-Order Ordinary Differential Equations

Thus far we have been concerned with systems of first-order ordinary differential equations. It is also possible to characterize initial-value problems by means of nth-order ordinary differential equations. To this end we let h be a real function that is defined and continuous on a domain D of the real $(t, y, \ldots, y_n)$-space [i.e., $D \subset R^{n+1}$, D is a domain, and $h \in C(D, R)$]. Then

$$y^{(n)} = h(t, y, y^{(1)}, \ldots, y^{(n-1)}) \tag{1.20}$$

is an nth-*order ordinary differential equation.*

A *solution* of (1.20) is a function $\phi \in C^n(J, R)$ that satisfies $(t, \phi(t), \phi^{(1)}(t), \ldots, \phi^{(n-1)}(t)) \in D$ for all $t \in J$ and

$$\phi^{(n)}(t) = h(t, \phi(t), \phi^{(1)}(t), \ldots, \phi^{(n-1)}(t))$$

for all $t \in J$, where $J = (a, b)$ is a t-interval.

Now for a given $(t_0, x_{10}, \ldots, x_{n0}) \in D$, the *initial -value problem* for (1.20) is

$$\begin{aligned} y^{(n)} &= h(t, y, y^{(1)}, \ldots, y^{(n-1)}), \\ y(t_0) &= x_{10}, \ldots, y^{(n-1)}(t_0) = x_{n0}. \end{aligned} \tag{1.21}$$

A function ϕ is a *solution* of (1.21) if ϕ is a solution of (1.20) on some interval containing t_0 and if $\phi(t_0) = x_{10}, \ldots, \phi^{(n-1)}(t_0) = x_{n0}$.

As in the case of systems of first-order ordinary differential equations, we can point to several important special cases. Specifically, we consider equations of the form

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y^{(1)} + a_0(t)y = g(t), \tag{1.22}$$

where $a_i \in C(J, R)$, $i = 0, 1, \ldots, n-1$, and $g \in C(J, R)$. We refer to (1.22) as a *linear nonhomogeneous ordinary differential equation of order n*.

If in (1.22) we let $g(t) \equiv 0$, then

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y^{(1)} + a_0(t)y = 0. \tag{1.23}$$

We call (1.23) a *linear homogeneous ordinary differential equation of order n*.

If in (1.23) we have $a_i(t) \equiv a_i$, $i = 0, 1, \ldots, n-1$, then

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y^{(1)} + a_0y = 0, \tag{1.24}$$

and we call (1.24) a *linear, autonomous, homogeneous ordinary differential equation of order n*.

As in the case of systems of first-order ordinary differential equations, we can define *periodic* and *linear periodic ordinary differential equations of order n* in the obvious way.

It turns out that the theory of nth-order ordinary differential equations can be reduced to the theory of a system of n first-order ordinary differential equations. To demonstrate this, we let $y = x_1, y^{(1)} = x_2, \ldots, y^{(n-1)} = x_n$ in (1.21). We now obtain the system of first-order ordinary differential equations

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_n &= h(t, x_1, \ldots, x_n) \end{aligned} \tag{1.25}$$

that is defined for all $(t, x_1, \ldots, x_n) \in D$. Assume that $\phi = (\phi_1, \ldots, \phi_n)^T$ is a solution of (1.25) on an interval J. Since $\phi_2 = \dot{\phi}_1, \phi_3 = \dot{\phi}_2, \ldots, \phi_n = \phi_1^{(n-1)}$, and since

$$\begin{aligned} h(t, \phi_1(t), \ldots, \phi_n(t)) &= h(t, \phi_1(t), \phi_1^{(1)}(t), \ldots, \phi_1^{(n-1)}(t)) \\ &= \phi_1^{(n)}(t), \end{aligned}$$

it follows that the first component ϕ_1 of the vector ϕ is a solution of (1.20) on the interval J. Conversely, if ϕ_1 is a solution of (1.20) on J, then the vector $(\phi, \phi^{(1)}, \ldots, \phi^{(n-1)})^T$ is clearly a solution of (1.25). Moreover, if $\phi_1(t_0) = x_{10}, \ldots, \phi_1^{(n-1)}(t_0) = x_{n0}$, then the vector ϕ satisfies $\phi(t_0) = x_0 = (x_{10}, \ldots, x_{n0})^T$.

1.4 Examples of Initial-Value Problems

We now give several specific examples of initial-value problems.

Example 1.1. The mechanical system of Figure 1.2 consists of two point masses M_1 and M_2 that are acted upon by viscous damping forces (determined by viscous damping constants B, B_1, and B_2), spring forces (specified by the spring constants K, K_1, and K_2), and external forces f_1 and f_2. The initial displacements of M_1 and M_2 at $t_0 = 0$ are given by $y_1(0)$ and $y_2(0)$, respectively, and their initial velocities are given by $\dot{y}_1(0)$ and $\dot{y}_2(0)$. The arrows in Figure 1.2 indicate positive directions of displacement for M_1 and M_2.

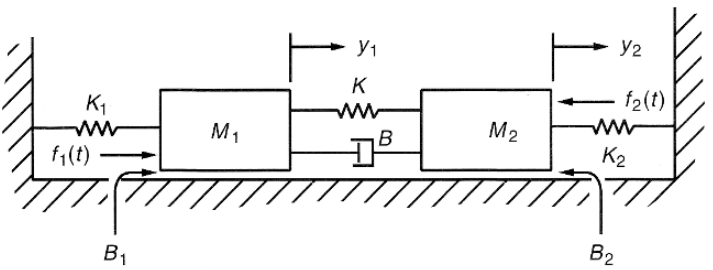

Figure 1.2. An example of a mechanical circuit

Newton's second law yields the following coupled second-order ordinary differential equations that describe the motions of the masses in Figure 1.2 (letting $y^{(2)} = d^2y/dt^2 = \ddot{y}$),

$$\begin{aligned} M_1\ddot{y}_1 + (B + B_1)\dot{y}_1 + (K + K_1)y_1 - B\dot{y}_2 - Ky_2 &= f_1(t) \\ M_2\ddot{y}_2 + (B + B_2)\dot{y}_2 + (K + K_2)y_2 - B_1\dot{y}_1 - Ky_1 &= -f_2(t) \end{aligned} \tag{1.26}$$

with initial data $y_1(0), y_2(0), \dot{y}_1(0)$, and $\dot{y}_2(0)$.

Letting $x_1 = y_1, x_2 = \dot{y}_1, x_3 = y_2$, and $x_4 = \dot{y}_2$, we can express (1.26) equivalently by the system of first-order ordinary differential equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-(K_1+K)}{M_1} & \frac{-(B_1+B)}{M_1} & \frac{K}{M_1} & \frac{B}{M_1} \\ 0 & 0 & 0 & 1 \\ \frac{K}{M_2} & \frac{B}{M_2} & \frac{-(K+K_2)}{M_2} & \frac{-(B+B_2)}{M_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M_1}f_1(t) \\ 0 \\ \frac{-1}{M_2}f_2(t) \end{bmatrix} \tag{1.27}$$

with initial data given by $x(0) = (x_1(0), x_2(0), x_3(0), x_4(0))^T$.

Example 1.2. Using the node voltages v_1, v_2, and v_3 and applying Kirchhoff's current law, we can describe the behavior of the electric circuit given in Figure 1.3 by the system of first-order ordinary differential equations

$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \\ \dot{v}_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_1}\left(\frac{1}{R_1}+\frac{1}{R_2}\right) & \frac{1}{R_2C_1} & 0 \\ -\frac{1}{C_1}\left(\frac{1}{R_1}+\frac{1}{R_2}\right) & -\left(\frac{R_2}{L}-\frac{1}{R_2C_1}\right) & \frac{R_2}{L} \\ \frac{1}{R_2C_2} & -\frac{1}{R_2C_2} & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} \frac{v}{R_1C_1} \\ \frac{v}{R_1C_1} \\ 0 \end{bmatrix}. \quad (1.28)$$

To complete the description of this circuit, we specify the initial data at $t_0 = 0$, given by $v_1(0), v_2(0)$, and $v_3(0)$.

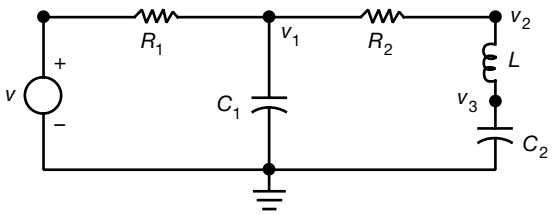

Figure 1.3. An example of an electric circuit

Example 1.3. Figure 1.4 represents a simplified model of an armature voltage-controlled dc servomotor consisting of a stationary field and a rotating armature and load. We assume that all effects of the field are negligible in the description of this system. The various parameters and variables in Figure 1.4 are e_a = externally applied armature voltage, i_a = armature current, R_a = resistance of the armature winding, L_a = armature winding inductance, e_m = back-emf voltage induced by the rotating armature winding, B = viscous damping due to bearing friction, J = moment of inertia of the armature and load, and θ = shaft position. The back-emf voltage (with the polarity as shown) is given by

$$e_m = K_\theta \dot{\theta}, \quad (1.29)$$

where $K_\theta > 0$ is a constant, and the torque T generated by the motor is given by

$$T = K_T i_a. \quad (1.30)$$

Application of Newton's second law and Kirchhoff's voltage law yields

$$J\ddot{\theta} + B\dot{\theta} = T(t) \quad (1.31)$$

and

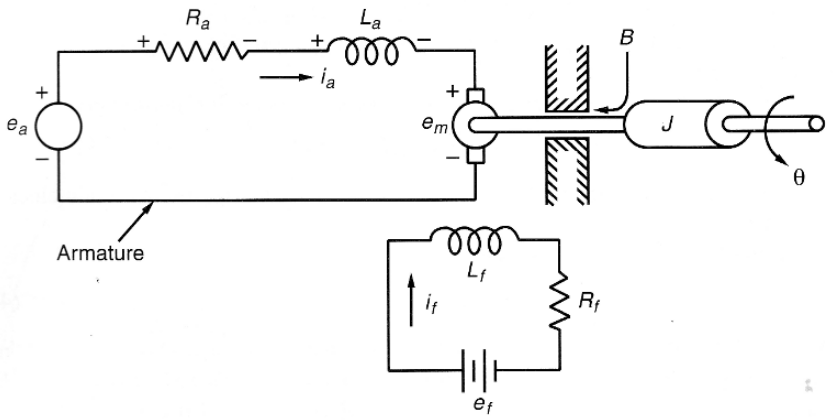

Figure 1.4. An example of an electro-mechanical system circuit

$$L_a \frac{di_a}{dt} + R_a i_a + e_m = e_a. \tag{1.32}$$

Combining (1.29) to (1.32) and letting $x_1 = \theta, x_2 = \dot{\theta}$, and $x_3 = i_a$ yields the system of first-order ordinary differential equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -B/J & K_T/J \\ 0 & -K_\theta/L_a & -R_a/L_a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ e_a/L_a \end{bmatrix}. \tag{1.33}$$

A suitable set of initial data for (1.33) is given by $t_0 = 0$ and $(x_1(0), x_2(0), x_3(0))^T = (\theta(0), \dot{\theta}(0), i_a(0))^T$.

Example 1.4. A much studied ordinary differential equation is given by

$$\ddot{x} + f(x)\dot{x} + g(x) = 0, \tag{1.34}$$

where $f \in C^1(R, R)$ and $g \in C^1(R, R)$.

When $f(x) \geq 0$ for all $x \in R$ and $xg(x) > 0$ for all $x \neq 0$, then (1.34) is called the *Lienard Equation*. This equation can be used to represent, e.g., RLC circuits with nonlinear circuit elements.

Another important special case of (1.34) is the *van der Pol Equation* given by

$$\ddot{x} - \epsilon(1 - x^2)\dot{x} + x = 0, \tag{1.35}$$

where $\epsilon > 0$ is a parameter. This equation has been used to represent certain electronic oscillators.

If in (1.34), $f(x) \equiv 0$, we obtain

$$\ddot{x} + g(x) = 0. \tag{1.36}$$

When $xg(x) > 0$ for all $x \neq 0$, then (1.36) represents various models of so-called "mass on a nonlinear spring." In particular, if $g(x) = k(1 + a^2x^2)x$, where $k > 0$ and $a^2 > 0$ are parameters, then g represents the restoring force of a *hard spring*. If $g(x) = k(1 - a^2x^2)x$, where $k > 0$ and $a^2 > 0$ are parameters, then g represents the restoring force of a *soft spring*. Finally, if $g(x) = kx$, then g represents the restoring force of a *linear spring*. (See Figures 1.5 and 1.6.)

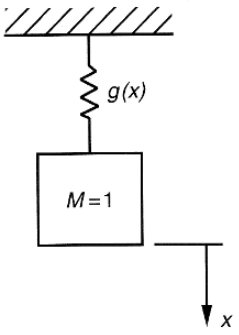

Figure 1.5. Mass on a nonlinear spring

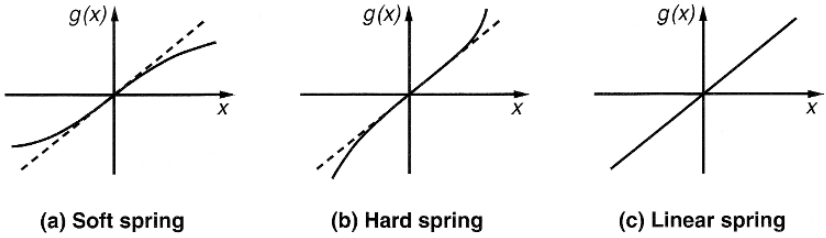

Figure 1.6. Mass on a nonlinear spring

For another special case of (1.34), let $f(x) \equiv 0$ and $g(x) = k \sin x$, where $k > 0$ is a parameter. Then (1.34) assumes the form

$$\ddot{x} + k \sin x = 0. \tag{1.37}$$

This equation describes the motion of a point mass moving in a circular path about the axis of rotation normal to a constant gravitational field, as shown in Figure 1.7. The parameter k depends on the radius l of the circular path, the

gravitational acceleration g, and the mass. The symbol x denotes the angle of deflection measured from the vertical. The present model is called a *simple pendulum*.

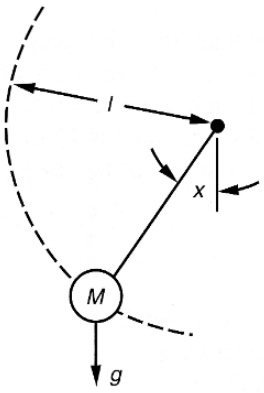

Figure 1.7. Model of a simple pendulum

Letting $x_1 = x$ and $x_2 = \dot{x}$, the second-order ordinary differential equation (1.34) can be represented by the system of first-order ordinary differential equations given by

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -f(x_1)x_2 - g(x_1). \end{aligned} \tag{1.38}$$

The required initial data for (1.38) are given by $x_1(0)$ and $x_2(0)$.

1.5 Solutions of Initial-Value Problems: Existence, Continuation, Uniqueness, and Continuous Dependence on Parameters

The following examples demonstrate that it is necessary to impose restrictions on the right-hand side of equation (1.11) to ensure the existence and uniqueness of solutions of the initial-value problem (1.12).

Example 1.5. For the initial-value problem,

$$\dot{x} = g(x), \quad x(0) = 0, \tag{1.39}$$

where $x \in R$, and

$$g(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0, \end{cases}$$

there exists no differentiable function ϕ that satisfies (1.39). Hence, *no solution* exists for this initial-value problem (in the sense defined in this chapter).

Example 1.6. The initial-value problem

$$\dot{x} = x^{1/3}, \quad x(t_0) = 0, \tag{1.40}$$

where $x \in R$, has at least two solutions given by $\phi_1(t) = [\frac{2}{3}(t - t_0)]^{3/2}$ and $\phi_2(t) = 0$ for $t \geq t_0$.

Example 1.7. The initial-value problem

$$\dot{x} = ax, \quad x(t_0) = x_0, \tag{1.41}$$

where $x \in R$, has a unique solution given by $\phi(t) = e^{a(t-t_0)}x(t_0)$ for $t \geq t_0$.

The following result provides a set of sufficient conditions for the *existence of solutions* of initial-value problem (1.12).

Theorem 1.8. *Let $f \in C(D, R^n)$. Then for any $(t_0, x_0) \in D$, the initial-value problem (1.12) has a solution defined on $[t_0, t_0 + c)$ for some $c > 0$.* ■

For a proof of Theorem 1.8, which is called the *Cauchy–Peano Existence Theorem*, refer to [1, Section 1.6].

The next result provides a set of sufficient conditions for the *uniqueness of solutions* for the initial-value problem (1.12).

Theorem 1.9. *Let $f \in C(D, R^n)$. Assume that for every compact set $K \subset D$, f satisfies the* Lipschitz condition

$$\| f(t, x) - f(t, y) \| \leq L_K \| x - y \| \tag{1.42}$$

for all $(t, x), (t, y) \in K$ where $L_K > 0$ is a constant depending only on K. Then (1.12) has at most one solution on any interval $[t_0, t_0 + c)$, $c > 0$. ■

For a proof of Theorem 1.9, refer to [1, Section 1.8]. In particular, if $f \in C^1(D, R^n)$, then the local Lipschitz condition (1.42) is automatically satisfied.

Now let ϕ be a solution of (1.11) on an interval J. By a *continuation* or *extension* of ϕ, we mean an extension ϕ_0 of ϕ to a larger interval J_0 in such a way that the extension solves (1.11) on J_0. Then ϕ is said to be *continued* or *extended* to the larger interval J_0. When no such continuation is possible, then ϕ is called *noncontinuable*.

Example 1.10. The scalar differential equation

$$\dot{x} = x^2 \tag{1.43}$$

has a solution $\phi(t) = \frac{1}{1-t}$ defined on $J = (-1, 1)$. This solution is continuable to the left to $-\infty$ and is not continuable to the right.

Example 1.11. The differential equation

$$\dot{x} = x^{1/3}, \tag{1.44}$$

where $x \in R$, has a solution $\psi(t) \equiv 0$ on $J = (-\infty, 0)$. This solution is continuable to the right in more than one way. For example, both $\psi_1(t) \equiv 0$ and $\psi_2(t) = (\frac{2t}{3})^{3/2}$ are solutions of (1.44) for $t \geq 0$.

In the next result, ∂D denotes the boundary of a domain D and ∂J denotes the boundary of an interval J.

Theorem 1.12. *If $f \in C(D, R^n)$ and if ϕ is a solution of (1.11) on an open interval J, then ϕ can be continued to a maximal open interval $J^* \supset J$ in such a way that $(t, \phi(t))$ tends to ∂D as $t \to \partial J^*$ when ∂D is not empty and $|t| + |\phi(t)| \to \infty$ if ∂D is empty. The extended solution ϕ^* on J^* is noncontinuable.* ■

For a proof of Theorem 1.12, refer to [1, Section 1.7].

When $D = J \times R^n$ for some open interval J and f satisfies a Lipschitz condition there (with respect to x), we have the following very useful *continuation* result.

Theorem 1.13. *Let $f \in C(J \times R^n, R^n)$ for some open interval $J \subset R$ and let f satisfy a Lipschitz condition on $J \times R^n$ (with respect to x). Then for any $(t_0, x_0) \in J \times R^n$, the initial-value problem (1.12) has a* unique *solution that exists on the* entire *interval J.* ■

For a proof of Theorem 1.13, refer to [1, Section 1.8].

In the next result we address initial-value problems that exhibit dependence on some parameter $\lambda \in G \subset R^m$ given by

$$\begin{aligned} \dot{x} &= f(t, x, \lambda), \\ x(\tau) &= \xi_\lambda, \end{aligned} \tag{1.45}$$

where $f \in C(J \times R^n \times G, R^n), J \subset R$ is an open interval, and ξ_λ depends continuously on λ.

Theorem 1.14. *Let $f \in C(J \times R^n \times G, R^n)$, where $J \subset R$ is an open interval and $G \subset R^m$. Assume that for each pair of compact subsets $J_0 \subset J$ and $G_0 \subset G$, there exists a constant $L = L_{J_0,G_0} > 0$ such that for all $(t, \lambda) \in J_0 \times G_0$, $x, y \in R^n$, the Lipschitz condition*

$$\| f(t, x, \lambda) - f(t, y, \lambda) \| \leq L \| x - y \| \tag{1.46}$$

is true. Then the initial-value problem (1.45) has a unique solution $\phi(t, \tau, \lambda)$, where $\phi \in C(J \times J \times G, R^n)$. Furthermore, if D is a set such that for all $\lambda_0 \in D$ there exists $\epsilon > 0$ such that $[\lambda_0 - \epsilon, \lambda_0 + \epsilon] \cap D \subset D$, then $\phi(t, \tau, \lambda) \to \phi(t, \tau_0, \lambda_0)$ uniformly for $t_0 \in J_0$ as $(\tau, \lambda) \to (\tau_0, \lambda_0)$, where J_0 is any compact subset of J. (Recall that the upper bar denotes closure of a set.) ■

For a proof of Theorem 1.14, refer to [1, Section 1.9].

Note that Theorem 1.14 applies in the case of Example 1.7 and that the solution $\phi(t)$ of (1.41) depends continuously on the parameter a and the initial conditions $x(t_0) = x_0$.

When Theorem 1.9 is satisfied, it is possible to approximate the unique solutions of the initial-value problem (1.12) arbitrarily closely, using the *method of successive approximations* (also known as *Picard iterations*). Let $f \in C(D, R^n)$, let $K \subset D$ be a compact set, and let $(t_0, x_0) \in K$. *Successive approximations* for (1.12), or equivalently for (1.13), are defined as

$$\begin{aligned} \phi_0(t) &= x_0, \\ \phi_{m+1}(t) &= x_0 + \int_{t_0}^{t} f(s, \phi_m(s))ds, \quad m = 0, 1, 2, \ldots \end{aligned} \tag{1.47}$$

for $t_0 \leq t \leq t_0 + c$, for some $c > 0$.

Theorem 1.15. *If $f \in C(D, R^n)$ and if f is Lipschitz continuous on some compact set $K \subset D$ with constant L (with respect to x), then the successive approximations $\phi_m, m = 0, 1, 2, \ldots$ given in (1.47) exist on $[t_0, t_0 + c]$, are continuous there, and converge uniformly, as $m \to \infty$, to the unique solution ϕ of (1.12). (Thus, for every $\epsilon > 0$ there exists $N = N(\epsilon)$ such that for all $t \in [t_0, t_0 + c]$, $\| \phi(t) - \phi_m(t) \| < \epsilon$ whenever $m > N(\epsilon)$.)* ■

For the proof of Theorem 1.15, refer to [1, Section 1.8].

1.6 Systems of Linear First-Order Ordinary Differential Equations

In this section we will address linear ordinary differential equations of the form

$$\dot{x} = A(t)x + g(t) \tag{1.48}$$

and

$$\dot{x} = A(t)x \tag{1.49}$$

and

$$\dot{x} = Ax + g(t) \tag{1.50}$$

and

$$\dot{x} = Ax, \tag{1.51}$$

where $x \in R^n, A(t) = [a_{ij}(t)] \in C(R, R^{n\times n}), g \in C(R, R^n)$, and $A \in R^{n\times n}$.

Linear equations of the type enumerated above may arise in a natural manner in the modeling process of physical systems (see Section 1.4 for specific examples) or in the process of linearizing equations of the form (1.11) or (1.14) or some other kind of form.

1.6.1 Linearization

We consider the system of first-order ordinary differential equations given by

$$\dot{x} = f(t, x), \tag{1.52}$$

where $f : R \times D \to R^n$ and $D \subset R^n$ is some domain.

Linearization About a Solution ϕ

If $f \in C^1(R \times D, R^n)$ and if ϕ is a given solution of (1.52) defined for all $t \in R$, then we can *linearize* (1.52) about ϕ in the following manner. Define $\delta x = x - \phi(t)$ so that

$$\begin{aligned} \frac{d(\delta x)}{dt} \triangleq \delta\dot{x} &= f(t, x) - f(t, \phi(t)) \\ &= f(t, \delta x + \phi(t)) - f(t, \phi(t)) \\ &= \frac{\partial f}{\partial x}(t, \phi(t))\delta x + F(t, \delta x), \end{aligned} \tag{1.53}$$

where $\frac{\partial f}{\partial x}(t, x)$ denotes the *Jacobian matrix* of $f(t, x) = (f_1(t, x), \ldots, f_n(t, x))^T$ with respect to $x = (x_1, \ldots, x_n)^T$; i.e.,

$$\frac{\partial f}{\partial x}(t, x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(t, x) & \cdots & \frac{\partial f_1}{\partial x_n}(t, x) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(t, x) & \cdots & \frac{\partial f_n}{\partial x_n}(t, x) \end{bmatrix} \tag{1.54}$$

and

$$F(t, \delta x) \triangleq [f(t, \delta x + \phi(t)) - f(t, \phi(t)] - \frac{\partial f}{\partial x}(t, \phi(t))\delta x. \tag{1.55}$$

It turns out that $F(t, \delta x)$ is $o(\| \delta x \|)$ as $\| \delta x \| \to 0$ uniformly in t on compact subsets of R; i.e., for any compact subset $I \subset R$, we have

$$\lim_{\|\delta x\|\to 0}\left(\sup_{t\in I}\frac{\| F(t,\delta x)\|}{\|\delta x\|}\right)=0.$$

For a proof of this assertion, we refer the reader to [1, Section 1.11].

Letting

$$\frac{\partial f}{\partial x}(t,\phi(t))=A(t),$$

we obtain from (1.53) the equation

$$\frac{d(\delta x)}{dt}\triangleq\delta\dot{x}=A(t)\delta x+F(t,\delta x). \tag{1.56}$$

Associated with (1.56) we have the linear differential equation

$$\dot{z}=A(t)z, \tag{1.57}$$

called the *linearized equation* of (1.52) about the solution ϕ.

In applications, the linearization (1.57) of (1.52), about a given solution ϕ, is frequently used as a means of approximating a nonlinear process by a linear one (in the vicinity of ϕ). In Chapter 4, where we will study the stability properties of equilibria of (1.52) [which are specific kinds of solutions of (1.52)], we will show under what conditions it makes sense to deduce qualitative properties of a nonlinear process from its linearization.

Of special interest is the case when in (1.52), f is independent of t, i.e.,

$$\dot{x}=f(x) \tag{1.58}$$

and ϕ is a constant solution of (1.58), say, $\phi(t)=x_0$ for all $t\in R$. Under these conditions we have

$$\frac{d(\delta x)}{dt}\triangleq\delta\dot{x}=A\delta x+F(\delta x), \tag{1.59}$$

where

$$\lim_{\|\delta x\|\to 0}\frac{\| F(\delta x)\|}{\|\delta x\|}=0 \tag{1.60}$$

and A denotes the Jacobian $\frac{\partial f}{\partial x}(x_0)$. Again, associated with (1.59) we have the linear differential equation

$$\dot{z}=Az,$$

called the *linearized equation* of (1.58) about the solution $\phi(t)\equiv x_0$.

Linearization About a Solution ϕ and an Input ψ

We can generalize the above to equations of the form

$$\dot{x}=f(t,x,u), \tag{1.61}$$

where $f : R \times D_1 \times D_2 \to R^n$ and $D_1 \subset R^n, D_2 \subset R^m$ are some domains. If $f \in C^1(R \times D_1 \times D_2, R^n)$ and if $\phi(t)$ is a given solution of (1.61) that we assume to exist for all $t \in R$ and that is determined by the initial condition x_0 and the *given specific function* $\psi \in C(R, R^m)$, i.e.,

$$\dot{\phi}(t) = f(t, \phi(t), \psi(t)), \quad t \in R,$$

then we can linearize (1.61) in the following manner. Define $\delta x = x - \phi(t)$ and $\delta u = u - \psi(t)$. Then

$$\begin{aligned}\frac{d(\delta x)}{dt} = \delta\dot{x} &= \dot{x} - \dot{\phi}(t) = f(t, x, u) - f(t, \phi(t), \psi(t)) \\ &= f(t, \delta x + \phi(t), \delta u + \psi(t)) - f(t, \phi(t), \psi(t)) \\ &= \frac{\partial f}{\partial x}(t, \phi(t), \psi(t))\delta x + \frac{\partial f}{\partial u}(t, \phi(t), \psi(t))\delta u \\ &\quad + F_1(t, \delta x, u) + F_2(t, \delta u), \end{aligned} \tag{1.62}$$

where

$$F_1(t, \delta x, u) = f(t, \delta x + \phi(t), u) - f(t, \phi(t), u) - \frac{\partial f}{\partial x}(t, \phi(t), \psi(t))\delta x$$

is $o(||\delta x||)$ as $\| \delta x \| \to 0$, uniformly in t on compact subsets of R for fixed u [i.e., for fixed u and for any compact subset $I \subset R$, $\lim\limits_{\|\delta x\| \to 0} \left(\sup_{t \in I} \frac{\|F_1(t, \delta x, u)\|}{\|\delta x\|} \right) = 0$], where

$$F_2(t, \delta u) = f(t, \phi(t), \delta u + \psi(t)) - f(t, \phi(t), \psi(t)) - \frac{\partial f}{\partial u}(t, \phi(t), \psi(t))\delta u$$

is $o(\| \delta u \|)$ as $\| \delta u \| \to 0$, uniformly in t on compact subsets of R [i.e., for any compact subset $I \subset R, \lim_{\|\delta u\| \to 0} \left(\sup_{t \in I} \frac{\|F_2(t, \delta u)\|}{\|\delta u\|} \right) = 0$], and where $\frac{\partial f}{\partial x}(\cdot)$ and $\frac{\partial f}{\partial u}(\cdot)$ denote the Jacobian matrix of f with respect to x and the Jacobian matrix of f with respect to u, respectively.

Letting

$$\frac{\partial f}{\partial x}(t, \phi(t), \psi(t)) = A(t) \text{ and } \frac{\partial f}{\partial u}(t, \phi(t), \psi(t)) = B(t),$$

we obtain from (1.62),

$$\frac{d(\delta x)}{dt} = \delta\dot{x} = A(t)\delta x + B(t)\delta u + F_1(t, \delta x, u) + F_2(t, \delta u). \tag{1.63}$$

Associated with (1.63), we have

$$\dot{z} = A(t)z + B(t)v. \tag{1.64}$$

We call (1.64) the *linearized equation* of (1.61) about the solution ϕ and the input function ψ.

As in the case of the linearization of (1.52) by (1.49), the linearization (1.64) of system (1.61) about a given solution ϕ and a given input ψ is often used in attempting to capture the qualitative properties of a nonlinear process by a linear process (in the vicinity of ϕ and ψ). In doing so, great care must be exercised to avoid erroneous conclusions.

The motivation of linearization is of course very obvious: much more is known about linear ordinary differential equations than about nonlinear ones. For example, the explicit forms of the solutions of (1.51) and (1.50) are known; the structures of the solutions of (1.49), (1.48), and (1.64) are known; the qualitative properties of the solutions of linear equations are known; and so forth.

1.6.2 Examples

We now consider some specific cases.

Example 1.16. We consider the *simple pendulum* discussed in Example 1.4 and described by the equation

$$\ddot{x} + k \sin x = 0, \tag{1.65}$$

where $k > 0$ is a constant. Letting $x_1 = x$ and $x_2 = \dot{x}$, (1.65) can be expressed as

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -k \sin x_1. \end{aligned} \tag{1.66}$$

It is easily verified that $\phi_1(t) \equiv 0$ and $\phi_2(t) \equiv 0$ is a solution of (1.66). Letting $f_1(x_1, x_2) = x_2$ and $f_2(x_1, x_2) = -k \sin x_1$, the Jacobian of $f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))^T$ evaluated at $(x_1, x_2)^T = (0, 0)^T$ is given by

$$J(0) \triangleq A = \begin{bmatrix} 0 & 1 \\ -k \cos x_1 & 0 \end{bmatrix}_{\begin{bmatrix} x_1=0 \\ x_2=0 \end{bmatrix}} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix}.$$

The linearized equation of (1.66) about the solution $\phi_1(t) \equiv 0, \phi_2(t) \equiv 0$ is given by

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Example 1.17. The system of equations

$$\begin{aligned} \dot{x}_1 &= a x_1 - b x_1 x_2 - c x_1^2, \\ \dot{x}_2 &= d x_2 - e x_1 x_2 - f x_2^2 \end{aligned} \tag{1.67}$$

describes the growth of two competing species (e.g., two species of small fish) that prey on each other (e.g., the adult members of one species prey on the young members of the other species, and vice versa). In (1.67) a, b, c, d, e, and f are positive parameters and it is assumed that $x_1 \geq 0$ and $x_2 \geq 0$. For (1.67), $\phi_1(t) = \phi_1(t, 0, 0) \equiv 0$ and $\phi_2(t) = \phi_2(t, 0, 0) \equiv 0, t \geq 0$, is a solution of (1.67). A simple computation yields

$$A = \frac{\partial f}{\partial x}(0) = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix},$$

and thus the system of equations

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

constitutes the linearized equation of (1.67) about the solution $\phi_1(t) = 0, \phi_2(t) = 0, t \geq 0$.

Example 1.18. Consider a unit mass subjected to an inverse square law force field, as depicted in Figure 1.8. In this figure, r denotes radius and θ denotes angle, and it is assumed that the unit mass (representing, e.g., a satellite) can thrust in the radial and in the tangential directions with thrusts u_1 and u_2, respectively. The equations that govern this system are given by

$$\begin{aligned} \ddot{r} &= r\dot{\theta}^2 - \frac{k}{r^2} + u_1, \\ \ddot{\theta} &= \frac{-2\dot{\theta}\dot{r}}{r} + \frac{1}{r}u_2. \end{aligned} \tag{1.68}$$

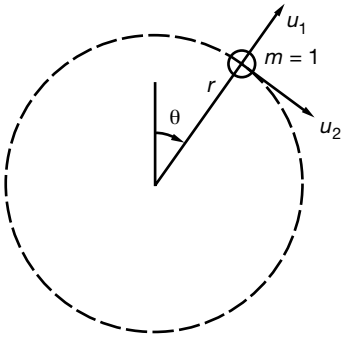

Figure 1.8. A unit mass subjected to an inverse square law force field

When $r(0) = r_0, \dot{r}(0) = 0, \theta(0) = \theta_0, \dot{\theta}(0) = \omega_0$, and $u_1(t) \equiv 0$, $u_2(t) \equiv 0$ for $t \geq 0$, it is easily verified that the system of equations (1.68) has as a solution the circular orbit given by

$$\begin{aligned} r(t) &\equiv r_0 = \text{ constant}, \\ \dot{\theta}(t) &= \omega_0 = \text{ constant} \end{aligned} \tag{1.69}$$

for all $t \geq 0$, which implies that

$$\theta(t) = \omega_0 t + \theta_0, \tag{1.70}$$

where $\omega_0 = (k/r_0^3)^{1/2}$.

If we let $x_1 = r$, $x_2 = \dot{r}$, $x_3 = \theta$, and $x_4 = \dot{\theta}$, the equations of motion (1.68) assume the form

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_1 x_4^2 - \frac{k}{x_1^2} + u_1, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -\frac{2x_2 x_4}{x_1} + \frac{u_2}{x_1}. \end{aligned} \tag{1.71}$$

The linearized equation of (1.71) about the solution (1.70) [with $u_1(t) \equiv 0, u_2(t) \equiv 0$] is given by

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega_0^2 & 0 & 0 & 2r_0\omega_0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{-2\omega_0}{r_0} & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & \frac{1}{r_0} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Example 1.19. In this example we consider systems described by equations of the form

$$\dot{x} + Af(x) + Bg(x) = u, \tag{1.72}$$

where $x \in R^n, A = [a_{ij}] \in R^{n\times n}, B = [b_{ij}] \in R^{n\times n}$ with $a_{ii} > 0$, $b_{ii} > 0$, $1 \leq i \leq n$, $f, g \in C^1(R^n, R^n)$, $u \in C(R^+, R^n)$, and $f(x) = 0$, $g(x) = 0$ if and only if $x = 0$.

Equation (1.72) can be used to model a great variety of physical systems. In particular, (1.72) has been used to model a large class of integrated circuits consisting of (nonlinear) transistors and diodes, (linear) capacitors and resistors, and current and voltage sources. (Figure 1.9 gives a symbolic representation of such circuits.) For such circuits, we assume that $f(x) = [f_1(x_1), \ldots, f_n(x_n)]^T$.

If $u(t) = 0$ for all $t \geq 0$, then $\phi_i(t) = 0$, $t \geq 0$, $1 \leq i \leq n$, is a solution of (1.72).

The system of equations (1.72) can be expressed equivalently as

$$\dot{x}_i = -\sum_{j=1}^{n} \left[a_{ij} \frac{f_j(x_j)}{x_j} + b_{ij} \frac{g_j(x_j)}{x_j} \right] x_j + u_i, \tag{1.73}$$

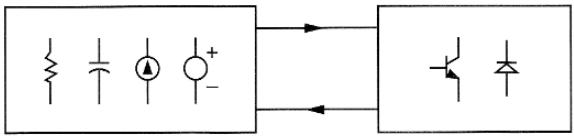

Figure 1.9. Integrated circuit

$i = 1, \ldots, n$. The linearized equation of (1.73) about the solution $\phi_i(t) = 0$, and the input $u_i(t) = 0$, $t \geq 0$, $i = 1, \ldots, n$, is given by

$$\dot{z}_i = -\sum_{j=1}^{n} \left[a_{ij} f_j'(0) + b_{ij} g_j'(0)\right] z_j + v_i, \tag{1.74}$$

where $f_j'(0) = \frac{df_j}{dx_j}(0)$ and $g_j'(0) = \frac{dg_j}{dx_j}(0), i = 1, \ldots, n$.

1.7 Linear Systems: Existence, Uniqueness, Continuation, and Continuity with Respect to Parameters of Solutions

In this section we address nonhomogeneous systems of first-order ordinary differential equations given by

$$\dot{x} = A(t)x + g(t), \tag{1.75}$$

where $x \in R^n, A(t) = [a_{ij}(t)]$ is a real $n \times n$ matrix and g is a real n-vector-valued function.

Theorem 1.20. *Suppose that $A \in C(J, R^{n\times n})$ and $g \in C(J, R^n)$, where J is some open interval. Then for any $t_0 \in J$ and any $x_0 \in R^n$, equation (1.75) has a unique solution satisfying $x(t_0) = x_0$. This solution exists on the* entire *interval J and is continuous in (t, t_0, x_0).*

Proof. The function $f(t, x) = A(t)x + g(t)$ is continuous in (t, x), and moreover, for any compact subinterval $J_0 \subset J$, there is an $L_0 \geq 0$ such that

$$\begin{aligned} \| f(t,x) - f(t,y) \|_1 &= \| A(t)(x-y) \|_1 \leq \| A(t) \|_1 \| x - y \|_1 \\ &\leq \left(\sum_{i=1}^{n} \max_{1\leq j\leq n} |a_{ij}(t)| \right) \| x - y \|_1 \leq L_0 \| x - y \|_1 \end{aligned}$$

for all $(t, x), (t, y) \in J_0 \times R^n$, where L_0 is defined in the obvious way. Therefore, f satisfies a Lipschitz condition on $J_0 \times R^n$.

If $(t_0, x_0) \in J_0 \times R^n$, then the continuity of f implies the existence of solutions (Theorem 1.8), whereas the Lipschitz condition implies the uniqueness of solutions (Theorem 1.9). These solutions exist for the entire interval J_0 (Theorem 1.13). Since this argument holds for *any* compact subinterval $J_0 \subset J$, the solutions exist and are unique for all $t \in J$. Furthermore, the solutions are continuous with respect to t_0 and x_0 (Theorem 1.14 modified for the case where A and g do not depend on any parameters λ). ■

For the case when in (1.75) the matrix A and the vector g depend continuously on parameters λ and μ, respectively, it is possible to modify Theorem 1.20, and its proof, in the obvious way to show that the unique solutions of the system of equations

$$\dot{x} = A(t, \lambda)x + g(t, \mu) \tag{1.76}$$

are continuous in λ and μ as well. [Assume that $A \in C(J \times R^l, R^{n\times n})$ and $g \in C(J \times R^m, R^n)$ and follow a procedure that is similar to the proof of Theorem 1.20.]

1.8 Solutions of Linear State Equations

In this section we determine the specific form of the solutions of systems of linear first-order ordinary differential equations. We will revisit this topic in much greater detail in Chapter 3.

Homogeneous Equations

We begin by considering linear homogeneous systems

$$\dot{x} = A(t)x, \tag{1.77}$$

where $A \in C(R, R^{n\times n})$. By Theorem 1.20, for every $x_0 \in R^n$, (1.77) has a unique solution that exists for all $t \in R$. We will now use Theorem 1.15 to derive an expression for the solution $\phi(t, t_0, x_0)$ for (1.77) for $t \in R$ with $\phi(t_0, t_0, x_0) = x_0$. In this case the successive approximations given in (1.47) assume the form

$$\begin{aligned}
\phi_0(t, t_0, x_0) &= x_0, \\
\phi_1(t, t_0, x_0) &= x_0 + \int_{t_0}^{t} A(s)x_0 ds, \\
\phi_2(t, t_0, x_0) &= x_0 + \int_{t_0}^{t} A(s)\phi_1(s, t_0, x_0) ds, \\
&\cdots \\
\phi_m(t, t_0, x_0) &= x_0 + \int_{t_0}^{t} A(s)\phi_{m-1}(s, t_0, x_0) ds,
\end{aligned}$$

or

$$\begin{aligned}\phi_m(t,t_0,x_0) &= x_0 + \int_{t_0}^{t} A(s_1)x_0 ds_1 + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)x_0 ds_2 ds_1 + \cdots \\ &\quad + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)\cdots\int_{t_0}^{s_{m-1}} A(s_m)x_0 ds_m \cdots ds_1 \\ &= \left[I + \int_{t_0}^{t} A(s_1)ds_1 + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)ds_2 ds_1 + \cdots \right. \\ &\quad \left. + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)\cdots\int_{t_0}^{s_{m-1}} A(s_m)ds_m \cdots ds_1\right] x_0,\end{aligned} \tag{1.78}$$

where I denotes the $n \times n$ identity matrix. By Theorem 1.15, the sequence $\{\phi_m\}, m = 0, 1, 2, \ldots$ determined by (1.78) converges uniformly, as $m \to \infty$, to the unique solution $\phi(t, t_0, x_0)$ of (1.77) on compact subsets of R. We thus have

$$\phi(t,t_0,x_0) = \Phi(t,t_0)x_0, \tag{1.79}$$

where

$$\begin{aligned}\Phi(t,t_0) &= I + \int_{t_0}^{t} A(s_1)ds_1 + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)ds_2 ds_1 \\ &\quad + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)\int_{t_0}^{s_2} A(s_3)ds_3 ds_2 ds_1 + \cdots \\ &\quad + \int_{t_0}^{t} A(s_1)\int_{t_0}^{s_1} A(s_2)\ldots\int_{t_0}^{s_{m-1}} A(s_m)ds_m ds_{m-1}\cdots ds_1 + \cdots .\end{aligned} \tag{1.80}$$

Expression (1.80) is called the *Peano–Baker series*.

From expression (1.80) we immediately note that

$$\Phi(t,t) = I. \tag{1.81}$$

Furthermore, by differentiating expression (1.80) with respect to time and substituting into (1.77), we obtain that

$$\dot{\Phi}(t,t_0) = A(t)\Phi(t,t_0). \tag{1.82}$$

From (1.79) it is clear that once the initial data are specified and once the $n \times n$ matrix $\Phi(t, t_0)$ is known, the entire behavior of system (1.77) evolving in time t is known. This has motivated the *state* terminology: $x(t_0) = x_0$ is the *state of the system (1.77) at time* t_0, $\phi(t, t_0, x_0)$ is the *state of the system (1.77) at time* t, the solution ϕ is called the *state vector* of (1.77), the components of ϕ are called the *state variables* of (1.77), and the matrix $\Phi(t, t_0)$ that maps $x(t_0)$ into $\phi(t, t_0, x_0)$ is called the *state transition matrix*

for (1.77). Also, the vector space containing the state vectors is called the *state space* for (1.77).

We can specialize the preceding discussion to linear systems of equations

$$\dot{x} = Ax. \tag{1.83}$$

In this case the mth term in (1.80) assumes the form

$$\int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2) \int_{t_0}^{s_2} A(s_3) \dots \int_{t_0}^{s_{m-1}} A(s_m) ds_m \cdots ds_1$$
$$= A^m \int_{t_0}^{t} \int_{t_0}^{s_1} \int_{t_0}^{s_2} \dots \int_{t_0}^{s_{m-1}} 1 ds_m \cdots ds_1 = \frac{A^m (t - t_0)^m}{m!},$$

and expression (1.78) for ϕ_m assumes now the form

$$\phi_m(t, t_0, x_0) = \left[I + \sum_{k=1}^{m} \frac{A^k (t - t_0)^k}{k!} \right] x_0.$$

We conclude once more from Theorem 1.15 that $\{\phi_m\}$ converges uniformly as $m \to \infty$ to the unique solution $\phi(t, t_0, x_0)$ of (1.83) on compact subsets of R. We have

$$\phi(t, t_0, x_0) = \left[I + \sum_{k=1}^{\infty} \frac{A^k (t - t_0)^k}{k!} \right] x_0$$
$$= \Phi(t, t_0) x_0 \triangleq \Phi(t - t_0) x_0, \tag{1.84}$$

where $\Phi(t - t_0)$ denotes the state transition matrix for (1.83). [Note that by writing $\Phi(t, t_0) = \Phi(t - t_0)$, we have used a slight abuse of notation.] By making the analogy with the scalar $e^a = 1 + \sum_{k=1}^{\infty} \frac{a^k}{k!}$, usage of the notation

$$e^A = I + \sum_{k=1}^{\infty} \frac{A^k}{k!} \tag{1.85}$$

should be clear. We call e^A a *matrix exponential*. In Chapter 3 we will explore several ways of determining e^A for a given A.

Nonhomogeneous Equations

Next, we consider linear nonhomogeneous systems of ordinary differential equations

$$\dot{x} = A(t)x + g(t), \tag{1.86}$$

where $A \in C(R, R^{n \times n})$ and $g \in C(R, R^n)$. Again, by Theorem 1.20, for every $x_0 \in R^n$, (1.86) has a unique solution that exists for all $t \in R$. Instead of *deriving* the complete solution of (1.86) for a given set of initial data

$x(t_0) = x_0$, we will *guess* the solution and verify that it indeed satisfies (1.86). To this end, let us assume that the solution is of the form

$$\phi(t, t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)g(s)ds, \tag{1.87}$$

where $\Phi(t, t_0)$ denotes the state transition matrix for (1.77).

To show that (1.87) is indeed the solution of (1.86), we first let $t = t_0$. In view of (1.81) and (1.87), we have $\phi(t_0, t_0, x_0) = x_0$. Next, by differentiating both sides of (1.87) and by using (1.81), (1.82), and (1.87), we have

$$\begin{aligned}
\dot{\phi}(t, t_0, x_0) &= \dot{\Phi}(t, t_0)x_0 + \Phi(t, t)g(t) + \int_{t_0}^{t} \dot{\Phi}(t, s)g(s)ds \\
&= A(t)\Phi(t, t_0)x_0 + g(t) + \int_{t_0}^{t} A(t)\Phi(t, s)g(s)ds \\
&= A(t)[\Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)g(s)ds] + g(t) \\
&= A(t)\phi(t, t_0, x_0) + g(t);
\end{aligned}$$

i.e., $\phi(t, t_0, x_0)$ given in (1.87) satisfies (1.86). Therefore, $\phi(t, t_0, x_0)$ is the unique solution of (1.86). Equation (1.87) is called the *variation of constants formula*, which is discussed further in Chapter 3. In the exercise section of Chapter 3 (refer to Exercise 3.13), we ask the reader (with hints) to *derive* the variation of constants formula (1.87), using a *change of variables*.

We note that when $x_0 = 0$, (1.87) reduces to

$$\phi(t, t_0, 0) \triangleq \phi_p(t) = \int_{t_0}^{t} \Phi(t, s)g(s)ds \tag{1.88}$$

and when $x_0 \neq 0$ but $g(t) = 0$ for all $t \in R$, (1.87) reduces to

$$\phi(t, t_0, x_0) \triangleq \phi_h(t) = \Phi(t, t_0)x_0. \tag{1.89}$$

Therefore, the *total solution* of (1.86) may be viewed as consisting of a component that is due to the initial conditions (t_0, x_0) and another component that is due to the *forcing term* $g(t)$. This type of separation is in general possible only in linear systems of differential equations. We call ϕ_p a *particular solution* of the nonhomogeneous system (1.86) and ϕ_h the *homogeneous solution*.

From (1.87) it is clear that for given initial conditions $x(t_0) = x_0$ and given forcing term $g(t)$, the behavior of system (1.86), summarized by ϕ, is known for all t. Thus, $\phi(t, t_0, x_0)$ specifies the *state vector*of (1.86) at time t. The components ϕ_i of ϕ, $i = 1, \ldots, n$, represent the *state variables* for (1.86), and the vector space that contains the state vectors is the *state space* for (1.86).

Before closing this section, it should be pointed out that in applications the matrix $A(t)$ and the vector $g(t)$ in (1.86) may be only *piecewise continuous* rather than continuous, as assumed above [i.e., $A(t)$ and $g(t)$ may have

(at most) a finite number of discontinuities over any finite time interval]. In such cases, the derivative of x with respect to t [i.e., the right-hand side in (1.86)] will be discontinuous at a finite number of instants over any finite time interval; however, the state itself, x, will still be continuous at these instants [i.e., the solutions of (1.86) will still be continuous over R]. In such cases, all the results presented concerning existence, uniqueness, continuation of solutions, and so forth, as well as the explicit expressions of solutions of (1.86), are either still valid or can be modified in the obvious way. For example, should $g(t)$ experience a discontinuity at, say, $t_1 > t_0$, then expression (1.87) will be modified to read as follows:

$$\phi(t, t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)g(s)ds, \quad t_0 \le t < t_1, \tag{1.90}$$

$$\phi(t, t_1, x_1) = \Phi(t, t_1)x_1 + \int_{t_1}^{t} \Phi(t, s)g(s)ds, \quad t \ge t_1, \tag{1.91}$$

where $x_1 = \lim_{t \to t_1^-} \phi(t, t_0, x_0)$.

1.9 Summary and Highlights

- *Initial-value problem*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{1.12}$$

or

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds, \tag{1.13}$$

where $\phi(t)$ is a solution of (1.12).

- *Successive approximations*

$$\begin{aligned} \phi_0(t) &= x_0, \\ \phi_{m+1}(t) &= x_0 + \int_{t_0}^{t} f(s, \phi_m(s))ds, \quad m = 0, 1, 2, \ldots. \end{aligned} \tag{1.47}$$

Under certain conditions (see Theorem 1.15) ϕ_m, $m = 1, 2$, converges uniformly (on compact sets) as $m \to \infty$ to the unique solution of (1.12).

- *Linearization*

Given is $\dot{x} = f(t, x)$ and a solution ϕ. The Jacobian matrix is

$$\frac{\partial f}{\partial x}(t, x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(t, x) & \cdots & \frac{\partial f_1}{\partial x_n}(t, x) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(t, x) & \cdots & \frac{\partial f_n}{\partial x_n}(t, x) \end{bmatrix}. \tag{1.54}$$

For $A(t) = \frac{\partial f}{\partial x}(t, \phi(t))$,

$$\dot{z} = A(t)z \tag{1.57}$$

is the linearized equation about the solution ϕ.

- *Existence and uniqueness of solutions* of

$$\dot{x} = A(t)x + g(t). \tag{1.75}$$

See Theorem 1.20.

- The *solution* of

$$\dot{x} = A(t)x + g(t), \tag{1.86}$$

with $x(t_0) = x_0$, is given by the variation of constants formula

$$\phi(t, t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)g(s)ds, \tag{1.87}$$

where the state transition matrix $\Phi(t, t_0)$ is given by

$$\Phi(t, t_0) = I + \int_{t_0}^{t} A(s_1)ds_1 + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2)ds_2ds_1 + \cdots \tag{1.80}$$

the Peano–Baker series.

- In the *time-invariant case* $\dot{x} = Ax$,

$$\begin{aligned} \phi(t, t_0, x_0) &= \left[I + \sum_{k=1}^{\infty} \frac{A^k(t - t_0)^k}{k!} \right] x_0 \\ &= \Phi(t, t_0)x_0 \triangleq \Phi(t - t_0)x_0 \\ &= e^{A(t-t_0)}x_0, \end{aligned} \tag{1.84}$$

where

$$e^A = I + \sum_{k=1}^{\infty} \frac{A^k}{k!}. \tag{1.85}$$

1.10 Notes

For a classic reference on ordinary differential equations, see Coddington and Levinson [3]. Other excellent sources include Brauer and Nohel [2], Hartman [4], and Simmons [7]. Our treatment of ordinary differential equations in this chapter was greatly influenced by Miller and Michel [6].

References

1. P.J. Antsaklis and A.N. Michel, *Linear Systems*, Birkhäuser, Boston, MA, 2006.
2. F. Brauer and J.A. Nohel, *Qualitative Theory of Ordinary Differential Equations*, Benjamin, New York, NY, 1969.
3. E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, NY, 1955.

4. P. Hartman, *Ordinary Differential Equations*, Wiley, New York, NY, 1964.
5. A.N. Michel, K. Wang, and B. Hu, *Qualitative Theory of Dynamical Systems, 2nd Edition*, Marcel Dekker, New York, NY, 2001.
6. R.K. Miller and A.N. Michel, *Ordinary Differential Equations*, Academic Press, New York, NY, 1982.
7. G.F. Simmons, *Differential Equations*, McGraw-Hill, New York, NY, 1972.

Exercises

1.1. (*Hamiltonian dynamical systems*) *Conservative dynamical systems*, also called *Hamiltonian dynamical systems*, are those systems that contain no energy-dissipating elements. Such systems with n degrees of freedom can be characterized by means of a *Hamiltonian function* $H(p,q)$, where $q^T = (q_1, \ldots, q_n)$ denotes n generalized position coordinates and $p^T = (p_1, \ldots, p_n)$ denotes n generalized momentum coordinates. We assume that $H(p,q)$ is of the form

$$H(p,q) = T(q,\dot{q}) + W(q), \tag{1.92}$$

where T denotes the kinetic energy and W denotes the potential energy of the system. These energy terms are obtained from the path-independent line integrals

$$T(q,\dot{q}) = \int_0^{\dot{q}} p(q,\xi)^T d\xi = \int_0^{\dot{q}} \sum_{i=1}^{n} p_i(q,\xi) d\xi_i, \tag{1.93}$$

$$W(q) = \int_0^{q} f(\eta)^T d\eta = \int_0^{q} \sum_{i=1}^{n} f_i(\eta) d\eta_i, \tag{1.94}$$

where f_i, $i = 1, \ldots, n$, denote generalized potential forces.

For the integral (1.93) to be path-independent, it is necessary and sufficient that

$$\frac{\partial p_i(q,\dot{q})}{\partial \dot{q}_j} = \frac{\partial p_j(q,\dot{q})}{\partial \dot{q}_i}, \quad i,j = 1, \ldots, n. \tag{1.95}$$

A similar statement can be made about (1.94).

Conservative dynamical systems are described by the system of $2n$ ordinary differential equations

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}(p,q), \qquad i = 1, \ldots, n, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}(p,q), \quad i = 1, \ldots, n. \end{aligned} \tag{1.96}$$

Note that if we compute the derivative of $H(p,q)$ with respect to t for (1.96) [along the solutions $q_i(t), p_i(t)$, $i = 1, \ldots, n$], then we obtain, by the chain rule,

$$\begin{aligned}\frac{dH}{dt}(p(t),q(t)) &= \sum_{i=1}^{n}\frac{\partial H}{\partial p_i}(p,q)\dot{p}_i + \sum_{i=1}^{n}\frac{\partial H}{\partial q_i}(p,q)\dot{q}_i \\ &= \sum_{i=1}^{n} -\frac{\partial H}{\partial p_i}(p,q)\frac{\partial H}{\partial q_i}(p,q) + \sum_{i=1}^{n}\frac{\partial H}{\partial q_i}(p,q)\frac{\partial H}{\partial p_i}(p,q) \\ &= -\sum_{i=1}^{n}\frac{\partial H}{\partial p_i}(p,q)\frac{\partial H}{\partial q_i}(p,q) + \sum_{i=1}^{n}\frac{\partial H}{\partial p_i}(p,q)\frac{\partial H}{\partial q_i}(p,q) \equiv 0.\end{aligned}$$

In other words, in a conservative system (1.96), the Hamiltonian, i.e., the total energy, will be constant along the solutions (1.96). This constant is determined by the initial data $(p(0), q(0))$.

(a) In Figure 1.10, M_1 and M_2 denote point masses; K_1, K_2, K denote spring constants; and x_1, x_2 denote the displacements of the masses M_1 and M_2. Use the Hamiltonian formulation of dynamical systems described above to derive a system of first-order ordinary differential equations that characterize this system. Verify your answer by using Newton's second law of motion to derive the same system of equations. Assume that $x_1(0)$, $\dot{x}_1(0)$, $x_2(0)$, and $\dot{x}_2(0)$ are given.

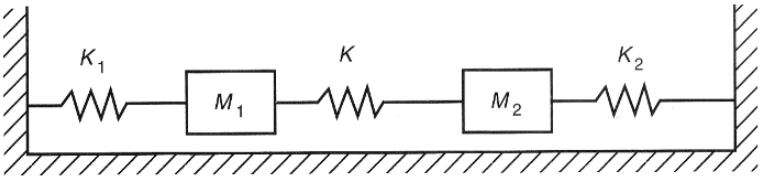

Figure 1.10. Example of a conservative dynamical system

(b) In Figure 1.11, a point mass M is moving in a circular path about the axis of rotation normal to a constant gravitational field (this is called the *simple pendulum problem*). Here l is the radius of the circular path, g is the gravitational acceleration, and θ denotes the angle of deflection measured from the vertical. Use the Hamiltonian formulation of dynamical systems described above to derive a system of first-order ordinary differential equations that characterize this system. Verify your answer by using Newton's second law of motion to derive the same system of equations. Assume that $\theta(0)$ and $\dot{\theta}(0)$ are given.

(c) Determine a system of first-order ordinary differential equations that characterizes the two-link pendulum depicted in Figure 1.12. Assume that $\theta_1(0)$, $\theta_2(0)$, $\dot{\theta}_1(0)$, and $\dot{\theta}_2(0)$ are given.

1.2. (*Lagrange's equation*) If a dynamical system contains elements that dissipate energy, such as viscous friction elements in mechanical systems and

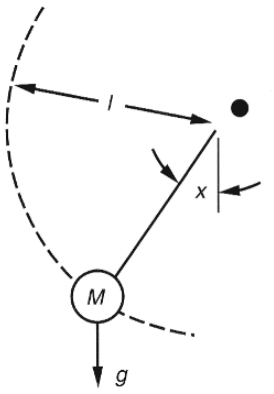

Figure 1.11. Simple pendulum

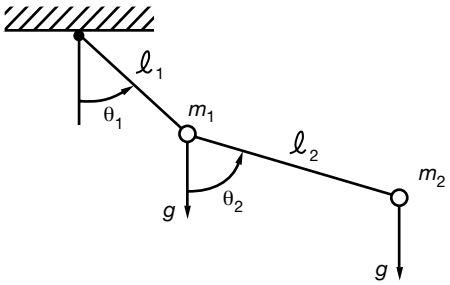

Figure 1.12. Two link pendulum

resistors in electric circuits, then we can use *Lagrange's equation* to describe such systems. (In the following, we use some of the same notation used in Exercise 1.1.) For a system with n degrees of freedom, this equation is given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}(q,\dot{q})\right) - \frac{\partial L}{\partial q}(q,\dot{q}) + \frac{\partial D}{\partial \dot{q}_i}(\dot{q}) = f_i, \quad i = 1,\ldots,n, \tag{1.97}$$

where $q^T = (q_1,\ldots,q_n)$ denotes the generalized position vector. The function $L(q,\dot{q})$ is called the *Lagrangian* and is defined as

$$L(q,\dot{q}) = T(q,\dot{q}) - W(q),$$

i.e., the difference between the kinetic energy T and the potential energy W.

The function $D(\dot{q})$ denotes *Rayleigh's dissipation function*, which we shall assume to be of the form

$$D(\dot{q}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij}\dot{q}_i\dot{q}_j,$$

where $[\beta_{ij}]$ is a positive semidefinite matrix (i.e., $[\beta_{ij}]$ is symmetric and all of its eigenvalues are nonnegative). The dissipation function D represents one-half the rate at which energy is dissipated as heat. It is produced by friction in mechanical systems and by resistance in electric circuits.

Finally, f_i in (1.97) denotes an applied force and includes all external forces associated with the q_i coordinate. The force f_i is defined as being positive when it acts to increase the value of the coordinate q_i.

(a) In Figure 1.13, M_1 and M_2 denote point masses; K_1, K_2, K denote spring constants; y_1, y_2 denote the displacements of masses M_1 and M_2, respectively; and B_1, B_2, B denote viscous damping coefficients. Use the Lagrange formulation of dynamical systems described above to derive two second-order differential equations that characterize this system. Transform these equations into a system of first-order ordinary differential equations. Verify your answer by using Newton's second law of motion to derive the same system equations. Assume that $y_1(0)$, $\dot{y}(0)$, $y_2(0)$, and $\dot{y}(0)$ are given.

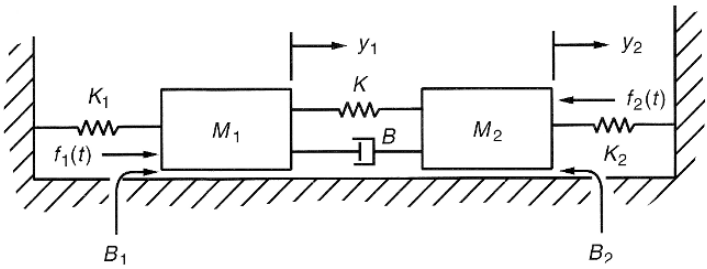

Figure 1.13. An example of a mechanical system with energy dissipation

(b) Consider the capacitor microphone depicted in Figure 1.14. Here we have a capacitor constructed from a fixed plate and a moving plate with mass M. The moving plate is suspended from the fixed frame by a spring with a spring constant K and has some damping expressed by the damping constant B. Sound waves exert an external force $f(t)$ on the moving plate. The output voltage v_s, which appears across the resistor R, will reproduce electrically the sound-wave patterns that strike the moving plate.
When $f(t) \equiv 0$ there is a charge q_0 on the capacitor. This produces a force of attraction between the plates that stretches the spring. When sound waves exert a force on the moving plate, there will be a resulting motion displacement x that is measured from the equilibrium position. The distance between the plates will then be $x_0 - x$, and the charge on the plates will be $q_0 + q$.
When displacements are small, the expression for the capacitance is given approximately by

$$C = \frac{\epsilon A}{x_0 - x}$$

with $C_0 = \epsilon A / x_0$, where $\epsilon > 0$ is the dielectric constant for air and A is the area of the plate.

Use the Lagrange formulation of dynamical systems to derive two second-order ordinary differential equations that characterize this system. Transform these equations into a system of first-order ordinary differential equations. Verify your answer by using Newton's laws of motion and Kirchhoff's voltage/current laws. Assume that $x(0), \dot{x}(0), q(0)$, and $\dot{q}(0)$ are given.

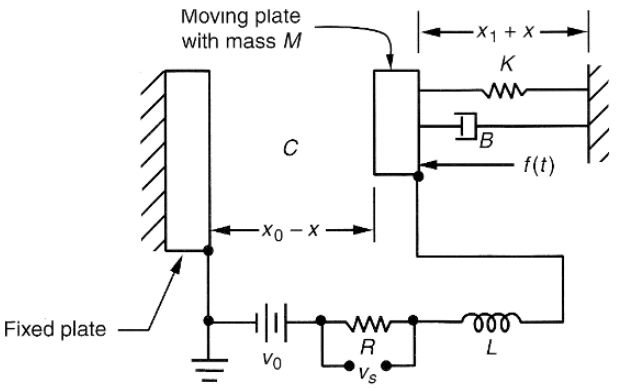

Figure 1.14. Capacitor microphone

(c) Use the Lagrange formulation to derive a system of first-order differential equations for the system given in Example 1.3.

1.3. Find examples of initial-value problems for which (a) no solutions exist; (b) more than one solution exists; (c) one or more solutions exist, but cannot be continued for all $t \in R$; and (d) unique solutions exist for all $t \in R$.

1.4. (*Numerical solution of ordinary differential equations—Euler's method*) An approximation to the solution of the *scalar* initial-value problem

$$\dot{y} = f(t, y), \quad y(t_0) = y_0 \tag{1.98}$$

is given by *Euler's method*,

$$y_{k+1} = y_k + hf(t_k, y_k), \quad k = 0, 1, 2, \ldots, \tag{1.99}$$

where $h = t_{k+1} - t_k$ is the (constant) integration step. The interpretation of this method is that the area below the solution curve is approximated by a sequence of sums of rectangular areas. This method is also called the *forward rectangular rule* (of integration).

(a) Use Euler's method to determine the solution of the initial-value problem

$$\dot{y} = 3y, \quad y(t_0) = 5, \quad t_0 = 0, \quad t_0 \le t \le 10.$$

(b) Use Euler's method to determine the solution of the initial-value problem

$$\ddot{y} = t(\dot{y})^2 - y^2, \quad y(t_0) = 1, \quad \dot{y}(t_0) = 0, \quad t_0 = 0, t_0 \le t \le 10.$$

Hint: In both cases, use $h = 0.2$. For part (b), let $y = x_1$, $\dot{x}_1 = x_2$, $\dot{x}_2 = tx_2^2 - x_1^2$, and apply (1.99), appropriately adjusted to the vector case. In both cases, plot y_k vs. $t_k, k = 0, 1, 2, \ldots$.
Remark:. Euler's method yields arbitrarily close approximations to the solutions of (1.98), by making h sufficiently small, *assuming infinite (computer) word length.* In practice, however, where truncation errors (quantization) and round-off errors (finite precision operations) are a reality, extremely small values of h may lead to numerical instabilities. Therefore, Euler's method is of limited value as a means of solving initial-value problems numerically.

1.5. (*Numerical solution of ordinary differential equations—Runge–Kutta methods*) The Runge–Kutta family of integration methods are among the most widely used techniques to solve initial-value problems (1.98). A simple version is given by

$$y_{i+1} = y_i + k,$$

where

$$k = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

with

$$\begin{aligned} k_1 &= hf(t_i, y_i), \\ k_2 &= hf(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1), \\ k_3 &= hf(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2), \\ k_4 &= hf(t_i + h, y_i + k_3), \end{aligned}$$

and $t_{i+1} = t_i + h, y(t_0) = y_0$.

The idea of this method is to probe ahead (in time) by one-half or by a whole step h to determine the values of the derivative at several points, and then to form a weighted average.

Runge–Kutta methods can also be applied to higher order ordinary differential equations. For example, after a change of variables, suppose that a second-order differential equation has been changed to a system of two first-order differential equations, say,

$$\begin{aligned} \dot{x}_1 &= f_1(t, x_1, x_2), \quad x_1(t_0) = x_{10}, \\ \dot{x}_2 &= f_2(t, x_1, x_2), \quad x_2(t_0) = x_{20}. \end{aligned} \tag{1.100}$$

In solving (1.100), a simple version of the Runge–Kutta method is given by

$$y_{i+1} = y_i + \underline{k},$$

where

$$y_i = (x_{1i}, x_{2i})^T \text{ and } \underline{k} = (k, l)^T$$

with

$$k = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad l = \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4)$$

and

$$\begin{aligned}
k_1 &= hf_1(t_i, x_{1i}, x_{2i}), & l_1 &= hf_2(t_i, x_{1i}, x_{2i}),\\
k_2 &= hf_1(t_i + \tfrac{1}{2}h, x_{1i} + \tfrac{1}{2}k_1, x_{2i} + \tfrac{1}{2}l_1), & l_2 &= hf_2(t_i + \tfrac{1}{2}h, x_{1i} + \tfrac{1}{2}k_1, x_{2i} + \tfrac{1}{2}l_1),\\
k_3 &= hf_1(t_i + \tfrac{1}{2}h, x_{1i} + \tfrac{1}{2}k_2, x_{2i} + \tfrac{1}{2}l_2), & l_3 &= hf_2(t_i + \tfrac{1}{2}h, x_{1i} + \tfrac{1}{2}k_2, x_{2i} + \tfrac{1}{2}l_2),\\
k_4 &= hf_1(t_i + h, x_{1i} + k_3, x_{2i} + l_3), & l_4 &= hf_2(t_i + h, x_{1i} + k_3, x_{2i} + l_3).
\end{aligned}$$

Use the Runge–Kutta method described above to obtain numerical solutions to the initial-value problems given in parts (a) and (b) of Exercise 1.4. Plot your data.
Remark: Since Runge–Kutta methods do not use past information, they constitute attractive starting methods for more efficient numerical integration schemes (e.g., predictor–corrector methods) . We note that since there are no built-in accuracy measures in the Runge–Kutta methods, significant computational efforts are frequently expended to achieve a desired accuracy.

1.6. (*Numerical solution of ordinary differential equations—Predictor–Corrector methods*) A common predictor–corrector technique for solving initial-value problems determined by ordinary differential equations, such as (1.98), is the *Milne* method, which we now summarize. In this method, $\dot{y}_{i-1}$ denotes the value of the first derivative at time t_{i-1}, where t_i is the time for the ith iteration step, $\dot{y}_{i-2}$ is similarly defined, and y_{i+1} represents the value of y to be determined. The details of the Milne method are:

1. $y_{i+1,p} = y_{i-3} + \frac{4h}{3}(2\dot{y}_{i-2} - \dot{y}_{i-1} + 2\dot{y}_i)$ (*predictor*)
2. $\dot{y}_{i+1,p} = f(t_{i+1}, y_{i+1,p})$
3. $y_{i+1,c} = y_{i-1} + \frac{h}{3}(\dot{y}_{i-1} + 4\dot{y}_i + \dot{y}_{i+1,p})$ (*corrector*)
4. $\dot{y}_{i+1,c} = f(t_{i+1}, y_{i+1,c})$
5. $y_{i+1,c} = y_{i-1} + \frac{h}{3}(\dot{y}_{i-1} + 4\dot{y}_i + \dot{y}_{i+1,c})$ (*iterating corrector*)

The first step is to obtain a predicted value of y_{i+1} and then to substitute $y_{i+1,p}$ into the given differential equation to obtain a predicted value of $\dot{y}_{i+1,p}$, as indicated in the second equation above. This predicted value, $\dot{y}_{i+1,p}$ is then used in the second equation, the corrector equation, to obtain a corrected value of y_{i+1}. The corrected value, $y_{i+1,c}$ is next substituted into the differential equation to obtain an improved value of $\dot{y}_{i+1}$, and so on. If necessary, an

iteration process involving the fourth and fifth equations continues until successive values of y_{i+1} differ by less than the value of some desirable tolerance. With y_{i+1} determined to the desired accuracy, the method steps forward one h increment.

A more complicated predictor–corrector method that is more reliable than the Milne method is the *Adams–Bashforth–Moulton* method, the essential equations of which are

$$y_{i+1,p} = y_i + \frac{h}{24}(55\dot{y}_i - 59\dot{y}_{i-1} + 37\dot{y}_{i-2} - 9\dot{y}_{i-3}),$$
$$y_{i+1,c} = y_i + \frac{h}{24}(9\dot{y}_{i+1} + 19\dot{y}_i - 5\dot{y}_{i-1} + \dot{y}_{i-2}),$$

where in the corrector equation, $\dot{y}_{i+1}$ denotes the predicted value.

The application of predictor–corrector methods to systems of first-order ordinary differential equations is straightforward. For example, the application of the Milne method to the second-order system in (1.100) yields from the predictor step

$$x_{k,i+1,p} = x_{k,i-3} + \frac{4h}{3}(2\dot{x}_{k,i-2} - \dot{x}_{k,i-1} + 2\dot{x}_{k,i}), \quad k = 1, 2.$$

Then

$$\dot{x}_{k,i+1,p} = f_k(t_{i+1}, x_{1,i+1,p}, x_{2,i+1,p}), \quad k = 1, 2,$$

and the corrector step assumes the form

$$x_{k,i+1,c} = x_{k,i-1} + \frac{h}{3}(\dot{x}_{k,i-1} + 4\dot{x}_{k,i} + \dot{x}_{k,i+1}), \quad k = 1, 2,$$

and

$$\dot{x}_{k,i+1,c} = f_k(t_{i+1}, x_{1,i+1,c}, x_{2,i+1,c}), \quad k = 1, 2.$$

Use the Milne method and the Adams–Bashforth–Moulton method described above to obtain numerical solutions to the initial-value problems given in parts (a) and (b) of Exercise 1.4. To initiate the algorithm, refer to the Remark in Exercise 1.5.

Remark. Derivations and convergence properties of numerical integration schemes, such as those discussed here and in Exercises 1.4 and 1.5, can be found in many of the standard texts on numerical analysis.

1.7. Use Theorem 1.15 to solve the initial-value problem $\dot{x} = ax + t, x(0) = x_0$ for $t \geq 0$. Here $a \in R$.

1.8. Consider the initial-value problem

$$\dot{x} = Ax, \quad x(0) = x_0, \tag{1.101}$$

where $x \in R^2$ and $A \in R^{2\times 2}$. Let λ_1, λ_2 denote the eigenvalues of A; i.e., λ_1 and λ_2 are the roots of the equation $\det(A - \lambda I) = 0$, where det denotes determinant, λ is a scalar, and I denotes the 2×2 identity matrix. Make specific choices of A to obtain the following cases:

1. $\lambda_1 > 0, \lambda_2 > 0$, and $\lambda_1 \neq \lambda_2$
2. $\lambda_1 < 0, \lambda_2 < 0$, and $\lambda_1 \neq \lambda_2$
3. $\lambda_1 = \lambda_2 > 0$
4. $\lambda_1 = \lambda_2 < 0$
5. $\lambda_1 > 0, \lambda_2 < 0$
6. $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta, i = \sqrt{-1}, \alpha > 0$
7. $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta, \alpha < 0$
8. $\lambda_1 = i\beta, \lambda_2 = -i\beta$

Using t as a parameter, plot $\phi_2(t, 0, x_0)$ vs. $\phi_1(t, 0, x_0)$ for $0 \leq t \leq t_f$ for every case enumerated above. Here $[\phi_1(t, t_0, x_0), \phi_2(t, t_0, x_0)]^T = \phi(t, t_0, x_0)$ denotes the solution of (1.101). On your plots, indicate increasing time t by means of arrows. Plots of this type are called *trajectories* for (1.101), and sufficiently many plots (using different initial conditions and sufficiently large t_f) make up a *phase portrait* for (1.101). Generate a phase portrait for each case given above.

1.9. Write two first-order ordinary differential equations for the *van der Pol Equation* (1.35) by choosing $x_1 = x$ and $x_2 = \dot{x}_1$. Determine by simulation *phase portraits* (see Exercise 1.8) for this example for the cases $\epsilon = 0.05$ and $\epsilon = 10$ (refer also to Exercises 1.5 and 1.6 for numerical methods for solving differential equations). The periodic solution to which the trajectories of (1.35) tend to is an example of a *limit cycle*.

1.10. Consider a system whose state-space description is given by

$$\dot{x} = -k_1 k_2 \sqrt{x} + k_2 u(t),$$
$$y = k_1 \sqrt{x}.$$

Linearize this system about the nominal solution

$$u_0 \equiv 0, \quad 2\sqrt{x_0(t)} = 2\sqrt{k} - k_1 k_2 t,$$

where $x_0(0) = k$.

1.11. For (1.36) consider the *hard, linear, and soft spring models* given by

$$g(x) = k(1 + a^2 x^2)x,$$
$$g(x) = kx,$$
$$g(x) = k(1 - a^2 x^2)x,$$

respectively, where $k > 0$ and $a^2 > 0$. Write two first-order ordinary differential equations for (1.36) by choosing $x_1 = x$ and $x_2 = \dot{x}$. Pick specific values for k and a^2. Determine by simulation *phase portraits* (see Exercise 1.8) for this example for the above three cases.

1.12. (a) Show that $x^T = (0, 0)$ is a solution of the system of equations

$$\dot{x}_1 = x_1^2 + x_2^2 + x_2 \cos x_1,$$
$$\dot{x}_2 = (1 + x_1)x_1 + (1 + x_2)x_2 + x_1 \sin x_2.$$

Linearize this system about the point $x^T = (0, 0)$. By means of computer simulations, compare solutions corresponding to different initial conditions in the vicinity of the origin of the above system of equations and its linearization.

(b) Linearize the (bilinear control) system

$$\ddot{x} + (3 + \dot{x}^2)\dot{x} + (1 + x + x^2)u = 0$$

about the solution $x = 0, \dot{x} = 0$, and the input $u(t) \equiv 0$. As in part (a), compare (by means of computer simulations) solutions of the above equation with corresponding solutions of its linearization.

(c) In the circuit given in Figure 1.15, $v_i(t)$ is a voltage source and the nonlinear resistor obeys the relation $i_R = 1.5v_R^3$ [$v_i(t)$ is the circuit input and $v_R(t)$ is the circuit output]. Derive the differential equation for this circuit. Linearize this differential equation for the case when the circuit operates about the point $v_i = 14$.

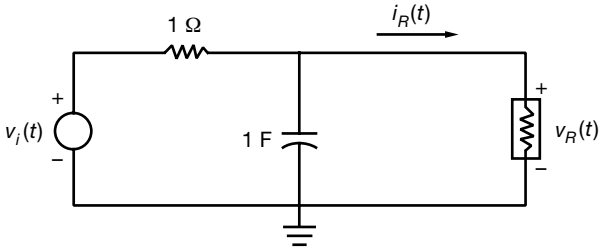

Figure 1.15. Nonlinear circuit

1.13. (*Inverted pendulum*) The inverted pendulum on a moving carriage subjected to an external force $\mu(t)$ is depicted in Figure 1.16.

The moment of inertia with respect to the center of gravity is J, and the coefficient of friction of the carriage (see Figure 1.16) is F. From Figure 1.17 we obtain the following equations for the dynamics of this system

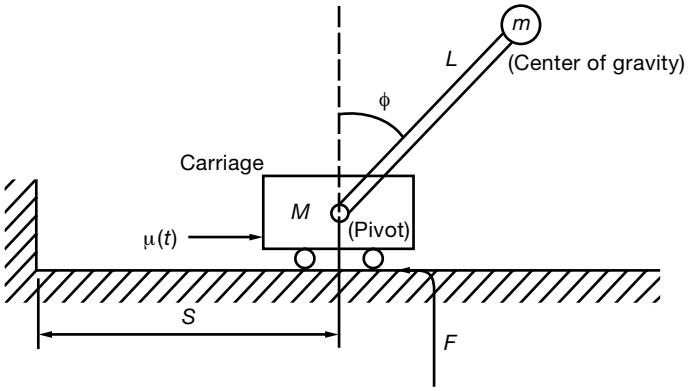

Figure 1.16. Inverted pendulum

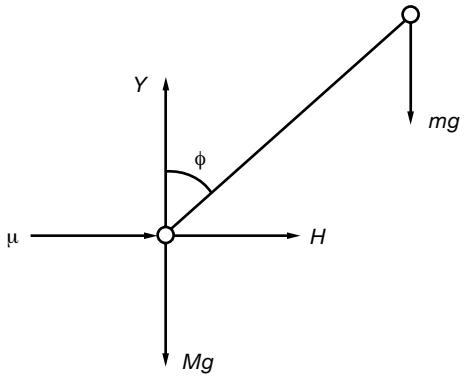

Figure 1.17. Force diagram of the inverted pendulum

$$m\frac{d^2}{dt^2}(S + L\sin\phi) \triangleq H, \tag{1.102a}$$

$$m\frac{d^2}{dt^2}(L\cos\phi) \triangleq Y - mg, \tag{1.102b}$$

$$J\frac{d^2\phi}{dt^2} = LY\sin\phi - LH\cos\phi, \tag{1.102c}$$

$$M\frac{d^2S}{dt^2} = \mu(t) - H - F\frac{dS}{dt}. \tag{1.102d}$$

Assuming that $m << M$, (1.102d) reduces to

$$M\frac{d^2S}{dt^2} = \mu(t) - F\frac{dS}{dt}. \tag{1.102e}$$

Eliminating H and Y from (1.102a) to (1.102c), we obtain

$$(J + mL^2)\ddot{\phi} = mgL\sin\phi - mL\ddot{S}\cos\phi. \tag{1.102f}$$

Thus, the system of Figure 1.16 is described by the equations

$$\begin{aligned} \ddot{\phi} - (g/L')\sin\phi + (1/L')\ddot{S}\cos\phi = 0, \\ M\ddot{S} + F\dot{S} = \mu(t), \end{aligned} \tag{1.102g}$$

where

$$L' = \frac{J + mL^2}{mL}$$

denotes the effective pendulum length.

Linearize system (1.102g) about $\phi = 0$.

1.14. (*Simple pendulum*) A system of first-order ordinary differential equations that characterize the simple pendulum considered in Exercise 1.1b is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{g}{l}\sin x_1 \end{bmatrix},$$

where $x_1 \triangleq \theta$ and $x_2 \triangleq \dot{\theta}$ with $x_1(0) = \theta(0)$ and $x_2(0) = \dot{\theta}(0)$ specified. A linearized model of this system about the solution $x = [0, 0]^T$ is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{l} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Let $g = 10\ (m/sec^2)$ and $l = 1\ (m)$.

(a) For the case when $x(0) = [\theta_0, 0]^T$ with $\theta_0 = \pi/18$, $\pi/12$, $\pi/6$, and $\pi/3$, plot the states for $t \geq 0$ for the nonlinear model.
(b) Repeat (a) for the linear model.
(c) Compare the results in (a) and (b).