2 Mathematical Preliminary

2.1 PHYSICAL SYSTEMS AND MODELS

This text is concerned with analytical study of control systems. Roughly speaking, it consists of four parts:

1. Modeling
2. Development of mathematical equations
3. Analysis
4. Design

This chapter discusses the first two parts. The distinction between physical systems and models is fundamental in engineering. In fact, the circuits and control systems studied in most texts are models of physical systems. For example, a resistor with a constant resistance is a model; the power limitation of the resistor is often disregarded. An inductor with a constant inductance is also a model; in reality, the inductance may vary with the amount of current flowing through it. An operational amplifier is a fairly complicated device; it can be modeled, however, as shown in Figure 2.1. In mechanical engineering, an automobile suspension system may be modeled as shown in Figure 2.2. In bioengineering, a human arm may be modeled as shown in Figure 2.3(b) or, more realistically, as in Figure 2.3(c). Modeling is an extremely important problem, because the success of a design depends upon whether or not physical systems are adequately modeled.

Depending on the questions asked and depending on operational ranges, a physical system may have different models. For example, an electronic amplifier has

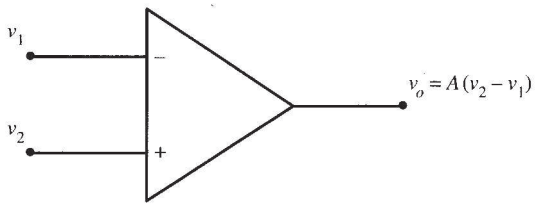


Figure 2.1 Model of operational amplifier.

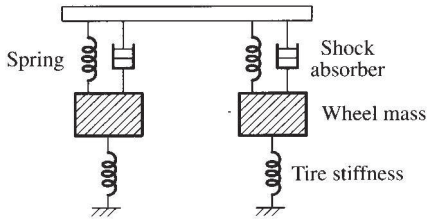


Figure 2.2 Model of automobile suspension system.

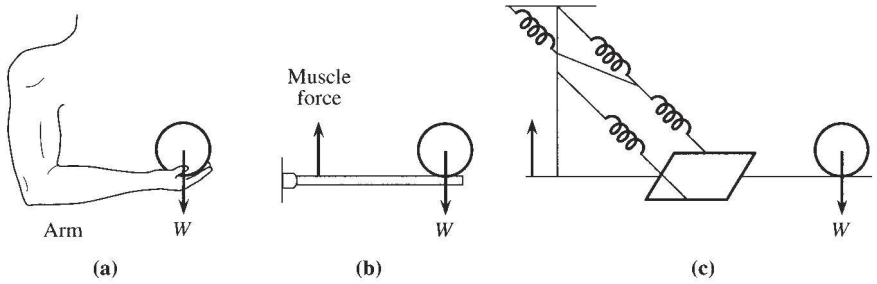


Figure 2.3 Models of arm.

different models at high and low frequencies. A spaceship may be modeled as a particle in the study of trajectory; however, it must be modeled as a rigid body in the study of maneuvering. In order to develop a suitable model for a physical system, we must understand thoroughly the physical system and its operational range. In this text, models of physical systems are also called *systems*. Hence, a *physical system* is a device or a collection of devices existing in the real world; a *system* is a model of a physical system. As shown in Figure 2.4, a system is represented by a unidirectional block with at least one input terminal and one output terminal. We remark that *terminal* does not necessarily mean a physical terminal, such as a wire sticking out of the block, but merely indicates that a signal may be applied or measured from that point. If an excitation or input signal $u(t)$ is applied to the input terminal of a

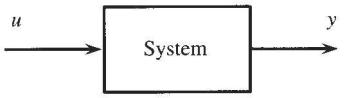


Figure 2.4 System.

system, a unique response or output signal $y(t)$ will be measurable or observable at the output terminal. This *unique* relationship between *excitation* and *response, input* and *output*, or *cause* and *effect* is implicit for every physical system and its model.

A system is called a *single-variable system* if it has only one input terminal and only one output terminal. Otherwise, it is called a *multivariable system.* A multivariable system has two or more input terminals and/or two or more output terminals. We study in this text mainly single-variable systems.

2.2 LINEAR TIME-INVARIANT LUMPED SYSTEMS

The choice of a model for a physical device depends heavily on the mathematics to be used. It is useless to choose a model that closely resembles the physical device but cannot be analyzed using existing mathematical methods. It is also useless to choose a model that can be analyzed easily but does not resemble the physical device. Therefore, the choice of models is not a simple task. It is often accomplished by a compromise between ease of analysis and resemblance to real physical systems.

The systems to be used in this text will be limited to those that can be described by ordinary linear differential equations with constant real coefficients such as

$$3\frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + y(t) = 2\frac{du(t)}{dt} - 3u(t)$$

or, more generally,

$$\begin{aligned} &a_n\frac{d^ny(t)}{dt^n} + a_{n-1}\frac{d^{n-1}y(t)}{dt^{n-1}} + \cdots + a_1\frac{dy(t)}{dt} + a_0y(t) \\ &= b_m\frac{d^mu(t)}{dt^m} + b_{m-1}\frac{d^{m-1}u(t)}{dt^{m-1}} + \cdots + b_1\frac{du(t)}{dt} + b_0u(t) \end{aligned} \tag{2.1}$$

where a_i and b_i are real constants, and $n \geq m$. Such equations are called *nth order linear time-invariant lumped (LTIL) differential equations.* In order to be describable by such an equation, the system must be linear, time-invariant, and lumped. Roughly speaking, a system is linear if it meets the additivity property [that is, the response of $u_1(t) + u_2(t)$ equals the sum of the response of $u_1(t)$ and the response of $u_2(t)$], and the homogeneity property [the response of $\alpha u(t)$ equals α times the response of $u(t)$]. A system is time-invariant if its characteristics—such as mass or moment of inertia for mechanical systems, or resistance, inductance or capacitance for electrical systems—do not change with time. A system is lumped if the effect of any past input $u(t)$, for $t \leq t_0$, on future output $y(t)$, for $t \geq t_0$, can be summarized by a *finite* number of initial conditions at $t = t_0$. For a detailed discussion of these concepts,

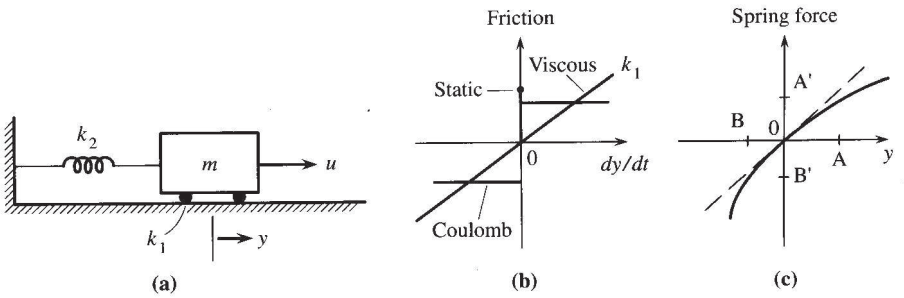


Figure 2.5 Mechanical system.

see References [15, 18]. We now discuss how these equations are developed to describe physical systems.

2.2.1 Mechanical Systems

Consider the system shown in Figure 2.5(a). It consists of a block with mass m connected to a wall by a spring. The input is the applied force $u(t)$, and the output is the displacement $y(t)$ measured from the equilibrium position. Before developing an equation to describe the system, we first discuss the characteristics of the friction and spring. The friction between the block and the floor is very complex. It generally consists of three parts—static, Coulomb, and viscous frictions—as shown in Figure 2.5(b). Note that the coordinates are friction versus velocity. When the mass is stationary or its velocity is zero, we need a certain amount of force to overcome the static friction to start its movement. Once the mass is moving, there is a constant friction, called the *Coulomb friction*, which is independent of velocity. The viscous friction is generally modeled as

$$\text{Viscous friction} = k_1 \times \text{Velocity} \tag{2.2}$$

where k_1 is called the *viscous friction coefficient*. This is a linear equation. Most texts on general physics discuss only static and Coulomb frictions. In this text, however, we consider only viscous friction; static and Coulomb frictions will be disregarded. By so doing, we can model the friction as a linear phenomenon.

In general physics, Hooke's law states that the displacement of a spring is proportional to the applied force, that is

$$\text{Spring force} = k_2 \times \text{Displacement} \tag{2.3}$$

where k_2 is called the *spring constant*. This equation is plotted in Figure 2.5(c) with the dotted line. It implies that no matter how large the applied force is, the displacement equals force/k_2. This certainly cannot be true in reality; if the applied force is larger than the elastic limit, the spring will break. In general, the characteristic of a physical spring has the form of the solid line shown in Figure 2.5(c).[1] We see that

[1]This is obtained by measurements under the assumption that the mass of the spring is zero and that the spring has no drifting and no hysteresis. See Reference [18].

if the applied force is outside the range [A′, B′], the characteristic is quite different from the dotted line. However, if the applied force lies inside the range [A′, B′], called the *linear operational range*, then the characteristic can very well be represented by (2.3). We shall use (2.3) as a model for the spring.

We now develop an equation to describe the system by using (2.3) and considering only the viscous friction in (2.2). The applied force $u(t)$ must overcome the friction and the spring force, and the remainder is used to accelerate the mass. Thus we have

$$u(t) - k_1 \frac{dy(t)}{dt} - k_2 y(t) = m \frac{d^2y(t)}{dt^2}$$

or

$$m \frac{d^2y(t)}{dt^2} + k_1 \frac{dy(t)}{dt} + k_2 y(t) = u(t) \tag{2.4}$$

This is an ordinary linear differential equation with constant coefficients. It is important to remember that this equation is obtained by using the linearized relation in (2.3) and considering only the viscous friction in (2.2). Therefore, it is applicable only for a limited operational range.

Consider now the rotational system shown in Figure 2.6(a). The input is the applied torque $T(t)$ and the output is the angular displacement $\theta(t)$ of the load. The shaft is not rigid and is modeled by a torsional spring. Let J be the moment of inertia of the load and the shaft. The friction between the shaft and bearing may consist of static, Coulomb, and viscous frictions. As in Figure 2.5, we consider only the viscous friction. Let k_1 be the viscous friction coefficient and k_2 the torsional spring constant. Then the torque generated by the friction equals $k_1 d\theta(t)/dt$ and the torque generated by the spring is $k_2\theta(t)$. The applied torque $T(t)$ must overcome the friction and spring torques; the remainder is used to accelerate the load. Thus we have

$$T(t) - k_1 \frac{d\theta(t)}{dt} - k_2\theta(t) = J \frac{d^2\theta(t)}{dt^2}$$

or

$$J \frac{d^2\theta(t)}{dt^2} + k_1 \frac{d\theta(t)}{dt} + k_2\theta(t) = T(t) \tag{2.5a}$$

This differential equation describes the system in Figure 2.6(a).

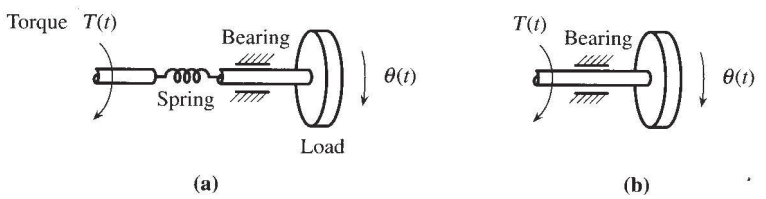


Figure 2.6 Rotational mechanical system.

If we identify the following equivalences:

Translational movement		Rotational movement
Linear displacement y	$\longleftrightarrow$	Angular displacement θ
Force u	$\longleftrightarrow$	Torque T
Mass m	$\longleftrightarrow$	Moment of inertia J

then Equation (2.5a) is identical to (2.4). The former describes a rotational movement, the latter, a linear or translational movement.

Exercise 2.2.1

The suspension system of an automobile can be modeled as shown in Figure 2.7. This model is simpler than the one in Figure 2.2, because it neglects the wheel mass and combines the tire stiffness with the spring. The model consists of one spring with spring constant k_2 and one dashpot or shock absorber. The dashpot is a device that provides viscous frictional force. Let k_1 be its viscous friction coefficient and let m be the mass of the car. A vertical force $u(t)$ is applied to the mass when the wheel hits a pothole. Develop a differential equation to describe the system.

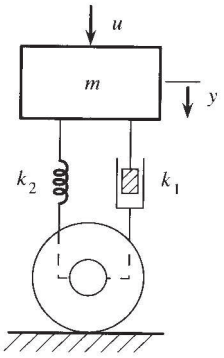


Figure 2.7 Suspension system of automobile.

[**Answer:** Same as (2.4).]

Exercise 2.2.2

Show that the system shown in Figure 2.6(b) where the shaft is assumed to be rigid is described by

$$J\frac{d^2\theta(t)}{dt^2} + k_1\frac{d\theta(t)}{dt} = T(t) \qquad (2.5b)$$

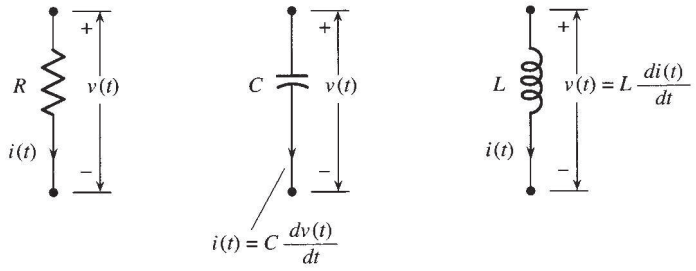


Figure 2.8 Electrical components.

2.2.2 RLC Networks

We discuss in this section circuits that are built by interconnecting resistors, capacitors, inductors, and current and voltage sources, beginning with the three basic elements shown in Figure 2.8. A resistor is generally modeled as $v = Ri$, where R is the resistance, v is the applied voltage, and i is current flowing through it with polarity chosen as shown. In the model $v = Ri$, nothing is said regarding power consumption. In reality, if v is larger than a certain value, the resistor will burn out. Therefore, the model is valid only within the specified power limitation. A capacitor is generally modeled as $Q = Cv$, where C is the capacitance, v is the applied voltage, and Q is the charge stored in the capacitor. The model implies that as the voltage increases to infinity, so does the stored charge. Physically this is not possible. As v increases, the stored charge will saturate and cease to increase, as shown in Figure 2.9. However, for v in a limited range, the model $Q = Cv$ does represent the physical capacitor satisfactorily. An inductor is generally modeled as $\phi = Li$ where L is the inductance, ϕ is the flux, and i is the current. In reality, the flux generated in an inductor will saturate as i increases. Therefore the relationship $\phi = Li$ is again applicable only within a limited range of i. Now if R, L, and C change with time, they are time-varying elements. If R, L, and C are constants, independent of time, then they are time-invariant elements. Using these linear time-invariant models, we can express their voltages and currents as

$$v(t) = Ri(t) \tag{2.6a}$$

$$i(t) = C\frac{dv(t)}{dt} \tag{2.6b}$$

$$v(t) = L\frac{di(t)}{dt} \tag{2.6c}$$

Now we shall use (2.6) to develop differential equations to describe RLC networks. Consider the network shown in Figure 2.10. The input is a current source $u(t)$ and the output $y(t)$ is the voltage across the capacitor as shown. The current of the capacitor, using (2.6b), is

$$i_c(t) = C\frac{dy(t)}{dt} = 2\frac{dy(t)}{dt}$$

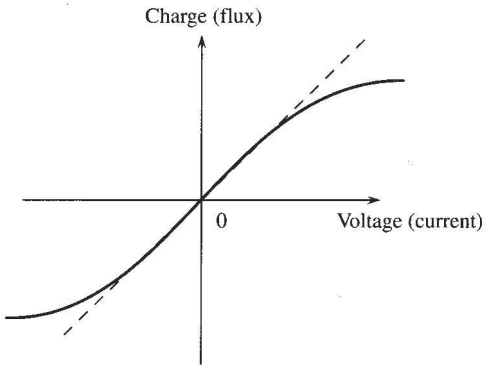


Figure 2.9 Characteristics of capacitor and inductor.

This current also passes through the 1-Ω resistor. Thus the voltage drop across A and B is

$$v_{AB} = i_c(t) \cdot 1 + y(t) = 2\frac{dy(t)}{dt} + y(t)$$

The current $i_1(t)$ passing through the 0.5-Ω resistor is

$$i_1(t) = \frac{v_{AB}}{0.5} = 4\frac{dy(t)}{dt} + 2y(t)$$

Thus we have

$$u(t) = i_1(t) + i_c(t) = 4\frac{dy(t)}{dt} + 2y(t) + 2\frac{dy(t)}{dt}$$

or

$$6\frac{dy(t)}{dt} + 2y(t) = u(t) \tag{2.7}$$

This first-order differential equation describes the network in Figure 2.10.

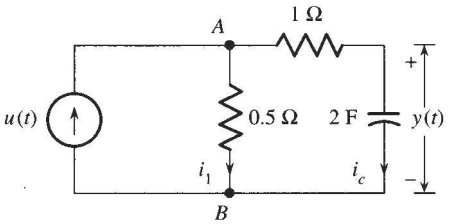


Figure 2.10 RC network.

Exercise 2.2.3

Find differential equations to describe the networks in Figure 2.11. The network in Figure 2.11(b) is called a *phase-lag network.*

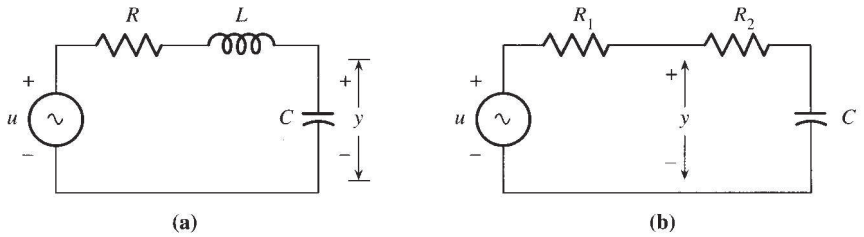


Figure 2.11 Networks.

[**Answers:** (a) $LC\, d^2y(t)/dt^2 + RC\, dy(t)/dt + y(t) = u(t)$.
(b) $C(R_1 + R_2)dy(t)/dt + y(t) = CR_2\, du(t)/dt + u(t)$.]

2.2.3 Industrial Process—Hydraulic Tanks

In chemical plants, it is often necessary to maintain the levels of liquids. A simplified model of such a system is shown in Figure 2.12, in which

q_i, q_1, q_2 = rates of the flow of liquid

A_1, A_2 = areas of the cross section of tanks

h_1, h_2 = liquid levels

R_1, R_2 = flow resistance, controlled by valves

It is assumed that q_1 and q_2 are governed by

$$q_1 = \frac{h_1 - h_2}{R_1} \quad \text{and} \quad q_2 = \frac{h_2}{R_2} \tag{2.8}$$

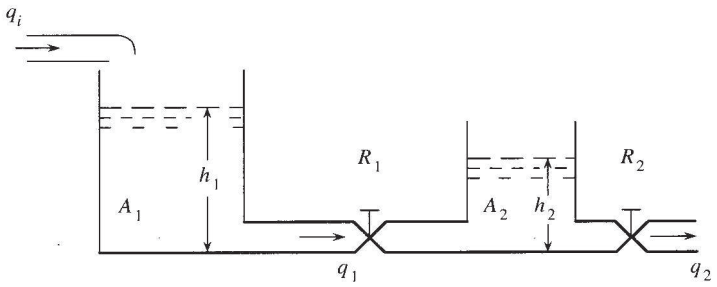


Figure 2.12 Control of liquid levels.

They are proportional to relative liquid levels and inversely proportional to flow resistances. The changes of liquid levels are governed by

$$A_1 dh_1 = (q_i - q_1)dt$$

and

$$A_2 dh_2 = (q_1 - q_2)dt$$

which imply

$$A_1 \frac{dh_1}{dt} = q_i - q_1 \tag{2.9a}$$

and

$$A_2 \frac{dh_2}{dt} = q_1 - q_2 \tag{2.9b}$$

These equations are obtained by linearization and approximation. In reality, the flow of liquid is very complex; it may involve turbulent flow, which cannot be described by linear differential equations. To simplify analysis, turbulent flow is disregarded in developing (2.9). Let q_i and q_2 be the input and output of the system. Now we shall develop a differential equation to describe them. The differentiation of (2.8) yields

$$\frac{dh_2}{dt} = R_2 \frac{dq_2}{dt} \quad \text{and} \quad \frac{dh_1}{dt} = R_1 \frac{dq_1}{dt} + \frac{dh_2}{dt} = R_1 \frac{dq_1}{dt} + R_2 \frac{dq_2}{dt}$$

The substitution of these into (2.9) yields

$$q_i - q_1 = A_1 \left(R_1 \frac{dq_1}{dt} + R_2 \frac{dq_2}{dt} \right) \tag{2.10a}$$

$$q_1 - q_2 = A_2 R_2 \frac{dq_2}{dt} \tag{2.10b}$$

Now we eliminate q_1 from (2.10a) by using (2.10b) and its derivative:

$$q_i - \left(q_2 + A_2 R_2 \frac{dq_2}{dt} \right) = A_1 R_1 \left(\frac{dq_2}{dt} + A_2 R_2 \frac{d^2 q_2}{dt^2} \right) + A_1 R_2 \frac{dq_2}{dt}$$

which can be simplified as

$$A_1 A_2 R_1 R_2 \frac{d^2 q_2}{dt^2} + (A_1(R_1 + R_2) + A_2 R_2) \frac{dq_2}{dt} + q_2(t) = q_i(t) \tag{2.11}$$

This second-order differential equation describes the input q_i and output q_2 of the system in Figure 2.12.

To conclude this section, we mention that a large number of physical systems can be modeled, after simplification and approximation, as linear time-invariant lumped (LTIL) systems over limited operational ranges. These systems can then be

described by LTIL differential equations. In this text, we study only this class of systems.

2.3 ZERO-INPUT RESPONSE AND ZERO-STATE RESPONSE

The response of linear, in particular LTIL, systems can always be decomposed into the zero-input response and zero-state response. In this section we shall use a simple example to illustrate this fact and then discuss some general properties of the zero-input response. The Laplace transform in Appendix A is needed for the following discussion.

Consider the differential equation

$$\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = 3\frac{du(t)}{dt} - u(t) \tag{2.12}$$

Many methods are available to solve this equation. The simplest method is to use the Laplace transform. The application of the Laplace transform to (2.12) yields, using (A.9),

$$\begin{aligned} s^2Y(s) - sy(0^-) - \dot{y}(0^-) + 3[sY(s) - y(0^-)] + 2Y(s) \\ = 3[sU(s) - u(0^-)] - U(s) \end{aligned} \tag{2.13}$$

where $\dot{y}(t) := dy(t)/dt$ and capital letters denote the Laplace transforms of the corresponding lowercase letters.[2] Equation (2.13) is an algebraic equation and can be manipulated using addition, subtraction, multiplication, and division. The grouping of $Y(s)$ and $U(s)$ in (2.13) yields

$$(s^2 + 3s + 2)Y(s) = sy(0^-) + \dot{y}(0^-) + 3y(0^-) - 3u(0^-) + (3s - 1)U(s)$$

which implies

$$Y(s) = \underbrace{\frac{(s + 3)y(0^-) + \dot{y}(0^-) - 3u(0^-)}{s^2 + 3s + 2}}_{\text{Zero-Input Response}} + \underbrace{\frac{3s - 1}{s^2 + 3s + 2}U(s)}_{\text{Zero-State Response}} \tag{2.14}$$

This equation reveals that the solution of (2.12) is partly excited by the input $u(t)$, $t \geq 0$, and partly excited by the initial conditions $y(0^-)$, $\dot{y}(0^-)$, and $u(0^-)$. These initial conditions will be called *the initial state*. The initial state is excited by the input applied before $t = 0$. In some sense, the initial state summarizes the effect of the past input $u(t)$, $t < 0$, on the future output $y(t)$, for $t \geq 0$. If different past inputs $u_1(t)$, $u_2(t)$, . . . , $t \leq 0$, excite the *same* initial state, then their effects on the future output will be identical. Therefore, how the differential equation acquires the initial state at $t = 0$ is immaterial in studying its solution $y(t)$, for $t \geq 0$. We mention that

[2]We use A := B to denote that A, by definition, equals B, and A =: B to denote that B, by definition, equals A.