Chapter 2

Mathematical models of systems

Before doing any controller design, we must first find a "good mathematical model" (i.e. simple but accurate) of the system we want to control. The system model serves as the basis on which a controller is designed. Modeling is an important part of the control engineering design process discussed in Chapter 1.

In principle, there are two different ways in which mathematical models can be obtained: from prior knowledge — e.g., in terms of physical laws — or by experimentation on a process. When attempting to obtain a specific model, it is often necessary and beneficial to combine both approaches.

In this chapter we focus on the first approach to modeling, i.e., modeling based on physical principles. We also introduce state models, and discuss linearization of nonlinear systems.

2.1 Notation

Generally, signals are written lower case: e.g., $x(t)$. Their transforms are capitalized: $X(s)$ or $X(j\omega)$. Resistance values, masses, etc. are capital: R, M, etc. The impulse is $\delta(t)$. In signals and systems the unit step is denoted $u(t)$, but in control $u(t)$ denotes a plant input so we'll denote the unit step by $\mathbf{1}(t)$. We will use the following notation for parts of the complex plane

$$\begin{aligned}
\mathbb{C}^- &:= \{s \in \mathbb{C} : \mathrm{Re}\,(s) < 0\} && \text{open left half-plane (OLHP)} \\
\mathbb{C}^+ &:= \{s \in \mathbb{C} : \mathrm{Re}\,(s) > 0\} && \text{open right half-plane (ORHP)} \\
j\mathbb{R} &:= \{s \in \mathbb{C} : \mathrm{Re}\,(s) = 0\} && \text{imaginary axis} \\
j\mathbb{R} \cup \mathbb{C}^+ &:= \{s \in \mathbb{C} : \mathrm{Re}\,(s) \geq 0\} && \text{closed right half-plane (CRHP).}
\end{aligned}$$

The symbol $:=$ means equal by definition.

2.2 General comments on modeling

A "model" is a set of equations used to represent a physical system. The model is never perfect because we often use experimental data that has noise and errors; real life is too complicated to model perfectly so we typically ignore some aspects of the system, e.g., we don't consider quantum mechanical effects in this course; we often purposely make approximations in our model to make the model tractable, i.e., simple enough.

There is a trade-off between model complexity and model accuracy. Typically, increasing the accuracy of a model also increases its complexity. The goal is to develop a model that is adequate for the problem at hand without being overly complex. Models are used for both control design and simulation. The models used for control design are usually simpler than simulation models.

In this course we deal with linear time-invariant systems as in the following example.

Example 2.2.1. This example demonstrates that there are various, equivalent, ways to model a linear time-invariant system. Consider an RC low pass filter in Figure 2.1[1] with input voltage $u(t)$ and output voltage $y(t)$ — the voltage across the capacitor.

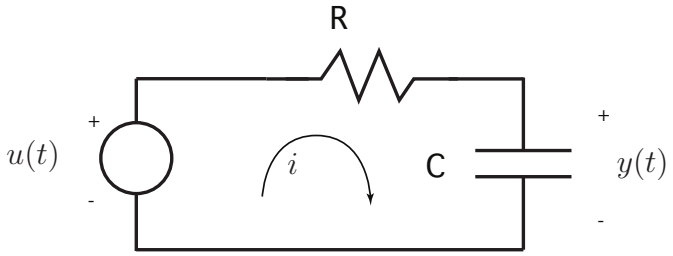

Figure 2.1: An RC filter.

In terms of Laplace transforms, i.e., using complex impedance and the voltage divider rule,

$$\frac{Y(s)}{U(s)} = \frac{\frac{1}{Cs}}{R + \frac{1}{Cs}} = \frac{1}{RCs + 1}.$$

[1] A circuit with a voltage source, a capacitor and a resistor in series.

This suffices to define a system in input-output form,

$$Y(s) = G(s)U(s),\ G(s) := \frac{1}{RCs+1}$$

where

$$\begin{aligned} U(s) &:= \text{Laplace transform of } u(t) \\ Y(s) &:= \text{Laplace transform of } y(t) \\ G(s) &:= \text{transfer function.} \end{aligned}$$

An alternative model of the same system is the linear differential equation

$$RC\dot{y} + y = u$$

with constant coefficients.

Or we could write instead the convolution integral equation model,

$$y = g * u,$$

that is,

$$y(t) = \int_{-\infty}^{t} g(t-\tau)u(\tau)\mathrm{d}\tau,$$

where $g(t)$ is the impulse response function

$$g(t) = \mathcal{L}^{-1}\left(G(s)\right) = \frac{1}{RC}\mathrm{e}^{-\frac{t}{RC}}\mathbf{1}(t).$$

We could also model this system using the Bode plot (more on this in the next chapter) of $G(s)$. Finally, we could write instead the state model (more on this later):

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= C_1 x \end{aligned}$$

where $x := y$ and

$$A := -\frac{1}{RC},\ B := \frac{1}{RC},\ C_1 := 1.$$

△

2.3 Block diagrams

The importance of block diagrams in control engineering can't be overemphasized. One could easily argue that you don't understand your system until you have a block diagram of it.

We shall take the point of view that a block diagram is a picture of a function. We can draw a picture of the function $y = f(x)$ like in Figure 2.2.

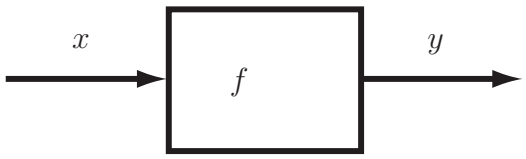

Figure 2.2: A block diagram is a picture of a function.

Thus a box represents a function and the arrows represent variables; the input is the independent variable, the output the dependent variable.

Example 2.3.1. The simplest vehicle to control is a cart on wheels illustrated in Figure 2.3.

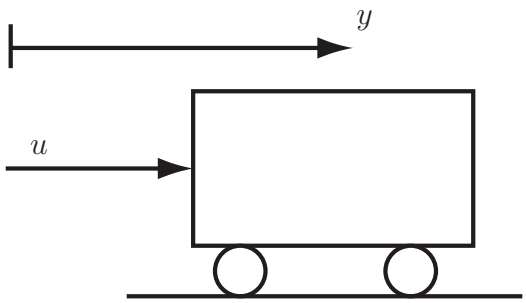

Figure 2.3: A simple cart.

Figure 2.3 is a schematic diagram, not a block diagram, because it doesn't say which of u, y causes the other. Assume the cart can move only in a straight line on a flat surface. (There may be air resistance to the motion and other friction effects.) Assume a force u is applied to the cart and let y denote the position of the cart measured from a stationary reference position. Then u and y are functions of time t and we could indicate this by $u(t)$ and $y(t)$. We regard the functions u and y as signals.

Newton's second law tells us that there's a mathematical relationship between u and y, namely, $u = M\ddot{y}$. We take the viewpoint that the force can be applied independently of anything else, that is, it's an input. Then y is an output. We represent this graphically by a block diagram.

Figure 2.4: Block diagram of a simple cart.

Suppose the cart starts at rest at the origin at time 0, i.e., $y(0) = \dot{y}(0) = 0$. Then the position depends only on the force applied. However y at time t depends on u not just at time t, but on past times as well. So we can write $y = F(u)$, i.e., y is a function of u, but we can't write $y(t) = F(u(t))$ because the position at time t doesn't depend only on the force at that same time t.

△

We may need to allow a block to have more than one input as in Figure 2.5. This means that y is a function of u and v, $y = F(u, v)$.

Figure 2.5: Block with multiple inputs.

Example 2.3.2. This example concerns a beam balanced on a fulcrum as shown in Figure 2.6. Suppose a torque τ is applied to the board. Let θ denote the angle of tilt and d the distance of roll. Then both θ and d are functions of τ. The block diagram could be as in Figure 2.7 or, as in Figure 2.8, or initial conditions, if they are not fixed, could be modeled as inputs too.

△

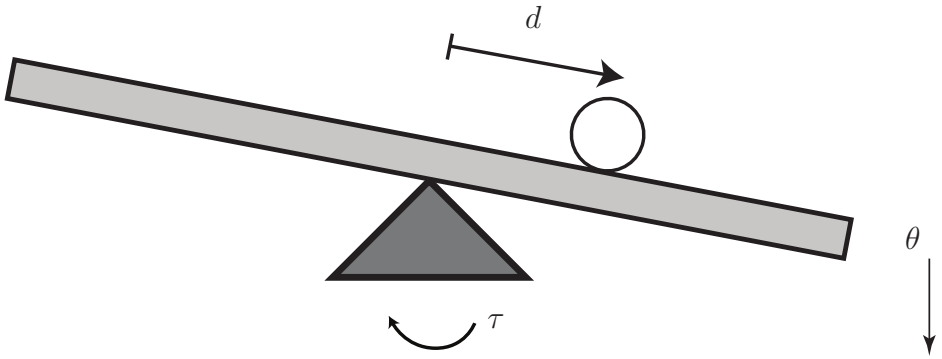

Figure 2.6: Beam balanced on a fulcrum.

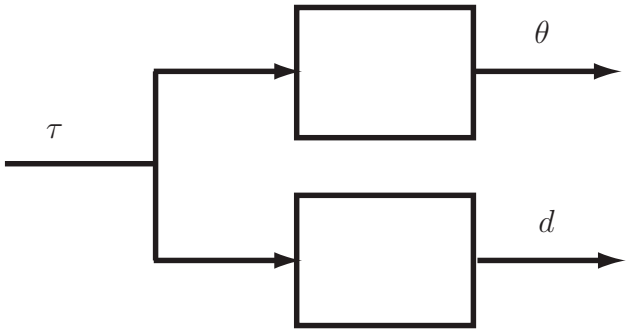

Figure 2.7: A block diagram of the beam system.

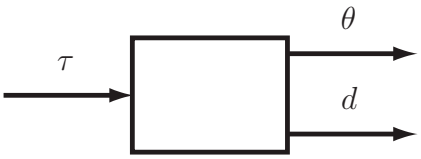

Figure 2.8: Alternate block diagram of the beam system.

Block diagrams may also have summing junctions as in Figure 2.9.

Getting a block diagram is sometimes harder than you think it will be. For example, you're riding a skateboard: The system components obviously include you and the skateboard. Anything else? You need the road to push on, and there must also be the earth to create a gravitation force on you (you can't skateboard in space). Would that be it? You, the skateboard, the road/earth. Does this system have an input? It must, if you're to move along the road following in, say, the bike lane. Try drawing a block diagram.

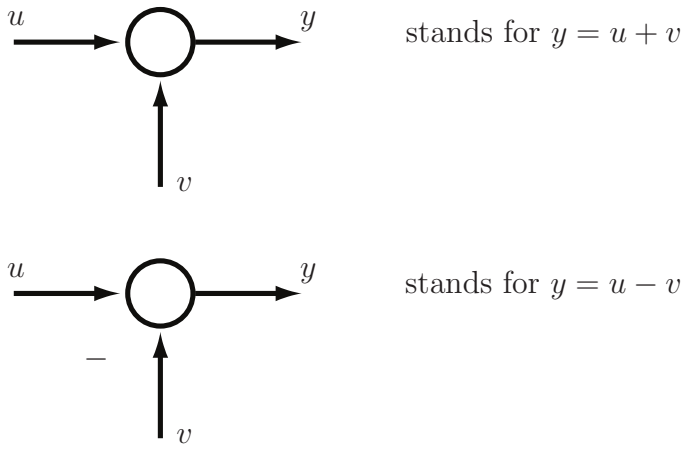

Figure 2.9: Summing junctions.

2.4 Modularity

A standard modeling technique that applies to any type of system is to use a "divide and conquer" approach. The basic idea is to

1. Partition a complex system into smaller, relatively easy pieces.
2. Model each of the smaller pieces.
3. Recombine the small models to obtain the full model.

Step 1 is accomplished by drawing a functional block diagram for the overall system. Each subsystem is drawn as a block with inputs and outputs. No detailed modeling is done at this stage.

In step 2 we come up with a mathematical model relating the inputs and outputs for each subsystem. This can be done using an analytical approach in which physical principles (e.g. Newton's laws, KVL, etc) are used. In this approach you may end up with a model with unknown parameters that will need to be estimated experimentally. Alternatively, if you don't know how to model the system analytically, you can treat the system as a black box and derive a model based on experimental input-output data. This is called system identification.

It's hard to provide a completely systematic approach for modeling so instead we'll give a few examples later on in this section.

2.4.1 Simple modeling components : mechanical

Recall that for translational motion of a mass M with position x, we have by Newton's equations, that

$$\begin{aligned} M\ddot{x} &= \sum \text{applied forces} \\ M\ddot{x} &= \sum F_i \end{aligned} \tag{2.1}$$

and similarly for the rotational motion of a body with inertia J and angular position θ

$$\begin{aligned} J\ddot{\theta} &= \sum \text{applied torques} \\ J\ddot{\theta} &= \sum \tau_i. \end{aligned} \tag{2.2}$$

The forces and torques that appear in the above summations are usually modeled by various types of simple components. These simple components form the building blocks of the mathematical model. Below we review the equations for some of them.

Spring

For linear translational springs the equations of motion are

$$F_k(x) = kx \tag{2.3}$$

where $k \in \mathbb{R}$ is the spring constant. For a linear rotational spring the equation of motion is

$$\tau_k(\theta) = k\theta. \tag{2.4}$$

Damper

A linear damper is often used to model viscous friction. For translational motion the equation is

$$F_b(\dot{x}) = b\dot{x} \tag{2.5}$$

where $b \in \mathbb{R}$ is a constant. For rotational motion the equation is

$$\tau_b(\theta) = b\dot{\theta}. \tag{2.6}$$

The above ideas are illustrated in Figure 2.10 and Figure 2.11.

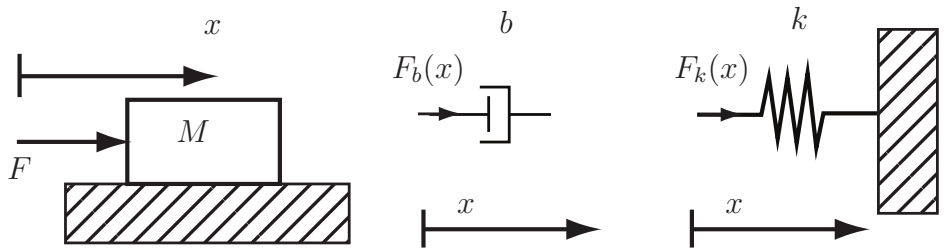

Figure 2.10: Simple translational mechanical components.

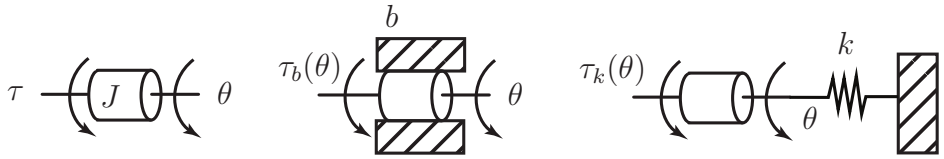

Figure 2.11: Simple rotational mechanical components.

2.4.2 Simple modeling components : electrical

Resistors

The current-voltage characteristics of a resistor R are static. Namely, the voltage $v(t)$ across a resistor at any fixed time t only depends on the current $i(t)$ at that particular time. The equations are

$$v(t) = Ri(t), \tag{2.7}$$

or, using complex impedances, in the Laplace domain,

$$V(s) = RI(s). \tag{2.8}$$

Inductors

The current-voltage characteristics of an inductor L are dynamic. Namely, the current $i(t)$ passing through an inductor at any fixed time t only depends on the voltage across the inductor $v(\tau)$ for $\tau \leq t$. The equations are

$$\begin{aligned} i(t) &= \frac{1}{L}\int_0^t v(\tau)\mathrm{d}\tau \\ v(t) &= L\frac{\mathrm{d}i(t)}{\mathrm{d}t}. \end{aligned} \tag{2.9}$$

In the Laplace domain, if we assume zero initial conditions, Equation (2.9) becomes

$$V(s) = sLI(s). \tag{2.10}$$

Capacitors

The current-voltage characteristics of a capacitor C are also dynamic. Namely, the voltage $v(t)$ across a capacitor at any fixed time t only depends on the current passing through the capacitor $i(\tau)$ for $\tau \leq t$. The equations are

$$\begin{aligned} v(t) &= \frac{1}{C}\int_0^t i(\tau)\mathrm{d}\tau \\ i(t) &= C\frac{\mathrm{d}v(t)}{\mathrm{d}t}. \end{aligned} \tag{2.11}$$

In the Laplace domain, if we assume zero initial conditions, Equation (2.11)

$$V(s) = \frac{1}{sC}I(s). \tag{2.12}$$

The electrical components are illustrated in Figure 2.12.

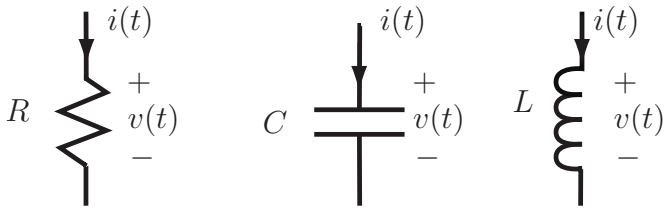

Figure 2.12: Simple electrical components.

2.4.3 Examples

Let's use the simple component models to model a system with only translational motion.

Example 2.4.1. Find a mathematical model for the simplified model of a car suspension system shown in Figure 2.13. In Figure 2.13 M_1 represents 1/4 of the mass of the car chassis while M_2 represents half the axle mass. We assume that

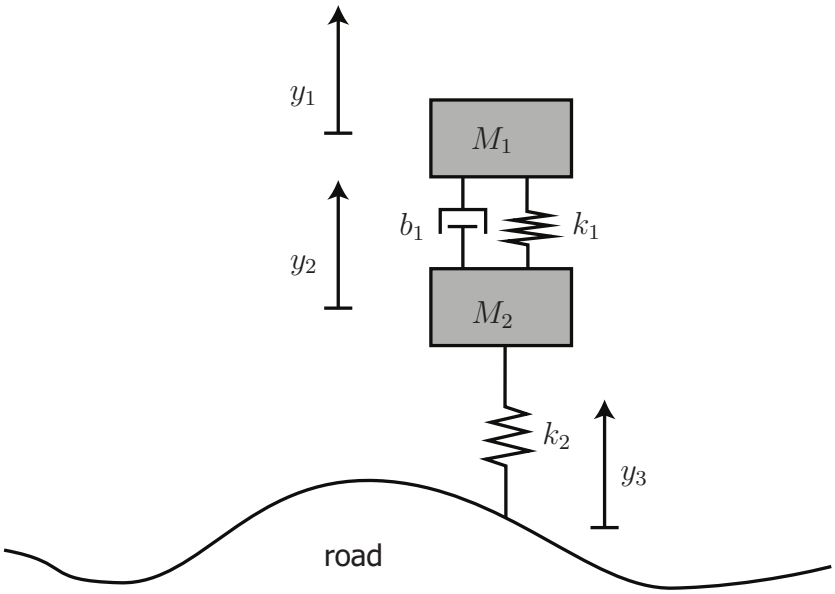

Figure 2.13: Simplified model of a car suspension system.

- y_3 is determined by the road surface, it is an exogenous signal, i.e., it comes from the "outside world".
- We take up to be the positive direction.
- We take the origin of our coordinates, i.e. the zero position, to be the equilibrium position so that we can ignore gravity. We can do this when we have linear springs.

The first step is to draw a functional block diagram for this system. In this example, because the system is relatively simple, a functional block diagram is easy to identify as shown in Figure 2.14. Figure 2.15 is the free body diagram for mass M_1 and shows the forces acting on it. The equation of

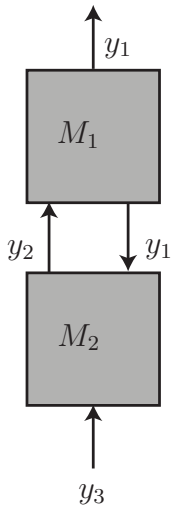

Figure 2.14: Functional block diagram of the car suspension model in Figure 2.13.

motion is

$$M_1\ddot{y}_1 = \sum \text{applied forces}$$
$$M_1\ddot{y}_1 = -k_1(y_1 - y_2) - b_1(\dot{y}_1 - \dot{y}_2)$$
$$\Rightarrow M_1\ddot{y}_1 + b_1\dot{y}_1 + k_1 y_1 = b_1\dot{y}_2 + k_1 y_2.$$

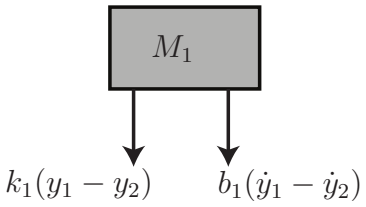

Figure 2.15: Free body diagram for mass M_1.

Figure 2.16 is the free body diagram for mass M_2. The equation of motion is

$$M_2\ddot{y}_2 = \sum \text{applied forces}$$
$$M_2\ddot{y}_2 = k_1(y_1 - y_2) + b_1(\dot{y}_1 - \dot{y}_2) - k_2(y_2 - y_3)$$
$$\Rightarrow M_2\ddot{y}_2 + b_1\dot{y}_2 + (k_1 + k_2)y_2 = k_1 y_1 + b_1\dot{y}_1 + k_2 y_3.$$

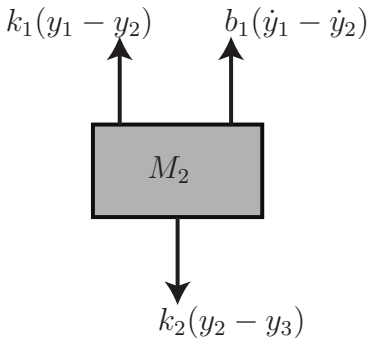

Figure 2.16: Free body diagram for mass M_2.

△

Example 2.4.2. A common actuator used in control systems is the DC motor to provide rotary motion. We will find the transfer function relating the applied armature voltage v to the rotor angle θ.

The "motor equations" give the torque τ on the rotor in terms of the armature current i and express the back emf voltage[2] e in terms of the shaft's rotational velocity $\dot{\theta}$

$$\begin{aligned} \tau &= K_\tau i \\ e &= K_e \dot{\theta} \end{aligned} \tag{2.13}$$

where K_τ and K_e are constants. These equations provide the link between the electrical and mechanical subsystems of the DC motor.

In order to model the DC motor we first draw a functional block diagram in Figure 2.17.

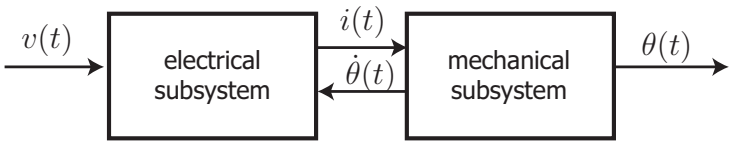

Figure 2.17: Functional block diagram of a DC motor.

[2]Since the generated electromotive force (emf) works against the applied armature voltage, it is called back emf.

Next model each block in the functional diagram. Start by drawing the equivalent circuit for the electrical subsystem and the free body diagram for the mechanical subsystem, the rotor, as shown in Figure 2.18. We assume that the rotor shaft has inertia J and viscous friction, which we model by a rotational damper with coefficient b.

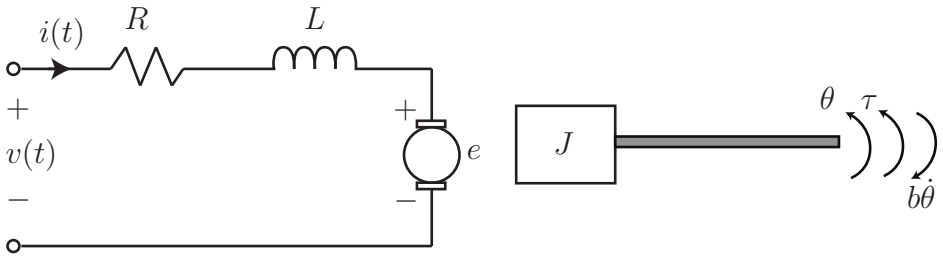

Figure 2.18: Electrical circuit and free body diagram for the DC motor.

Application of Newton's laws to the rotor, along with the motor equations (2.13), yields

$$\begin{aligned} J\ddot{\theta} &= -b\dot{\theta} + \tau \\ &= -b\dot{\theta} + K_\tau i \end{aligned} \tag{2.14}$$

The electrical equation is found using KVL and is given by

$$\begin{aligned} &-v + Ri + L\frac{di}{dt} + e = 0 \\ &\Rightarrow -v + Ri + L\frac{di}{dt} + K_e\dot{\theta} = 0. \end{aligned} \tag{2.15}$$

△

2.5 State models

Example 2.5.1. Consider a cart on wheels, driven by a force F and subject to air resistance as in Figure 2.19 Typically air resistance creates a force depending on the velocity $\dot{y}$; let's say this force is a possibly nonlinear function $D(\dot{y})$. Assuming M is constant, Newton's second law gives

$$M\ddot{y} = F - D(\dot{y}).$$

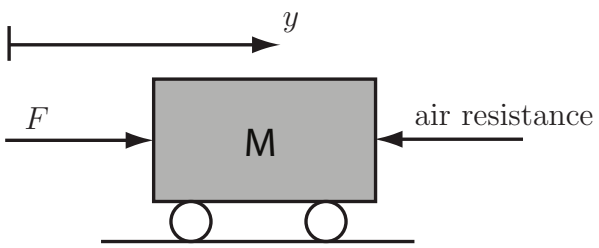

Figure 2.19: A simple cart with air resistance.

We are going to put this in a standard form by defining two so-called state variables, in this example position and velocity:

$$x_1 := y, \quad x_2 := \dot{y}.$$

Then

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{M}F - \frac{1}{M}D(x_2) \\ y &= x_1. \end{aligned}$$

These equations have the form

$$\begin{aligned} \dot{x} &= f(x,u) \qquad \text{(state equation)} \\ y &= h(x) \qquad \text{(output equation)} \end{aligned} \tag{2.16}$$

where

$$x := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \; u := F$$

$$f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2, \; f(x_1, x_2, u) = \begin{bmatrix} x_2 \\ \frac{1}{M}u - \frac{1}{M}D(x_2) \end{bmatrix}$$

$$h : \mathbb{R}^2 \to \mathbb{R}, \; h(x_1, x_2) = x_1.$$

The function f is nonlinear if D is; h is linear. Equation (2.16) constitutes a state model of the system, and x is called the **state** or state vector. Here the plant is a possibly nonlinear system, u (applied force) is the input, y (cart position) is the output, and

$$x = \begin{bmatrix} \text{cart position} \\ \text{cart velocity} \end{bmatrix}$$

is the state of the system. (We'll define state later.)

As a special case, suppose the air resistance is a linear function of velocity

$$D(x_2) = dx_2, \qquad d \text{ is a real constant.}$$

Then f is linear:

$$f(x,u) = Ax + Bu, \quad A := \begin{bmatrix} 0 & 1 \\ 0 & -d/M \end{bmatrix}, \; B := \begin{bmatrix} 0 \\ 1/M \end{bmatrix}.$$

Defining $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$, we get the state model

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & -\frac{d}{M} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned} \tag{2.17}$$

In other words, when $D(\dot{y})$ is linear, the state model has the structure

$$\dot{x} = Ax + Bu, \quad y = Cx.$$

This model is of a linear, time-invariant (LTI) system.

---△

It is convenient to write vectors sometimes as column vectors and sometimes as n-tuples, i.e., ordered lists. For example

$$x := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad x = (x_1, x_2).$$

We'll use both.

Generalizing the example, we can say that an important class of models is

$$\begin{aligned} \dot{x} &= f(x,u), \; f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \\ y &= h(x,u), \; h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p. \end{aligned} \tag{2.18}$$

This model is nonlinear, time-invariant. The input u has dimension m, the output y dimension p, and the state x dimension n. An example where $m = 2, p = 2, n = 4$ is shown in Figure 2.20. In Figure 2.20 we have

$$u = (u_1, u_2), \; y = (y_1, y_2), \; x = (y_1, \dot{y}_1, y_2, \dot{y}_2).$$

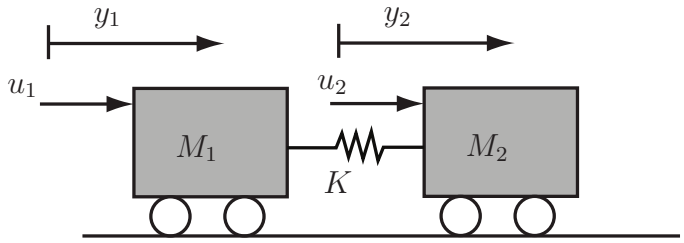

Figure 2.20: A multiple-input multiple-output example.

The LTI special case is

$$\begin{aligned} \dot{x} &= Ax + Bu,\ A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m} \\ y &= Cx + Du,\ C \in \mathbb{R}^{p\times n},\ D \in \mathbb{R}^{p\times m}. \end{aligned} \tag{2.19}$$

In this course we deal exclusively with single-input single-output systems. Therefore we restrict ourselves to the case when $m = 1$ and $p = 1$.

Now we turn to the concept of the state of a system. Roughly speaking, $x(t_0)$ encapsulates all the system dynamics up to time t_0, that is, no additional prior information is required. More precisely, the concept is this: For any t_0 and t_1, with $t_0 < t_1$, knowing $x(t_0)$ and knowing $\{u(t) : t_0 \le t \le t_1\}$, we can compute $x(t_1)$, and hence $y(t_1)$.

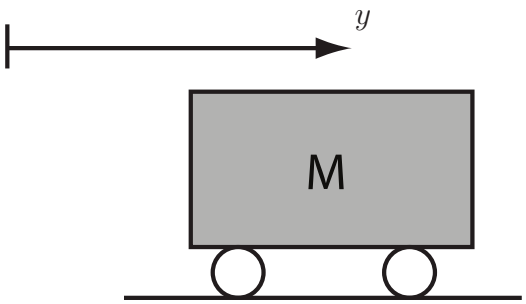

Figure 2.21: A simple cart with no air resistance, no applied force.

Example 2.5.2. Consider the cart system in Figure 2.21 with no applied forces and no air resistance. If we were to try simply $x = y$, then knowing $x(t_0)$ without $\dot{y}(t_0)$, we could not solve the initial value problem for the future cart position. Similarly $x = \dot{y}$ won't work. Since the equation of motion,

$M\ddot{y} = 0$, is second order, we need two initial conditions at $t = t_0$, implying we need a 2-dimensional state vector. In general for mechanical systems it is customary to take x to consist of positions and velocities of all masses.

△

Why are state models useful? First, there's a rich control theory for them. This is covered in ECE488. You may have heard of the Kalman filter; it is based on a state model. Secondly, it gives a convenient data type to store a model, namely, the matrices A, B, C, D. And third, it gives a simple way to linearize, namely, compute Jacobians. More on this later.

Example 2.5.3. A mass-spring-damper

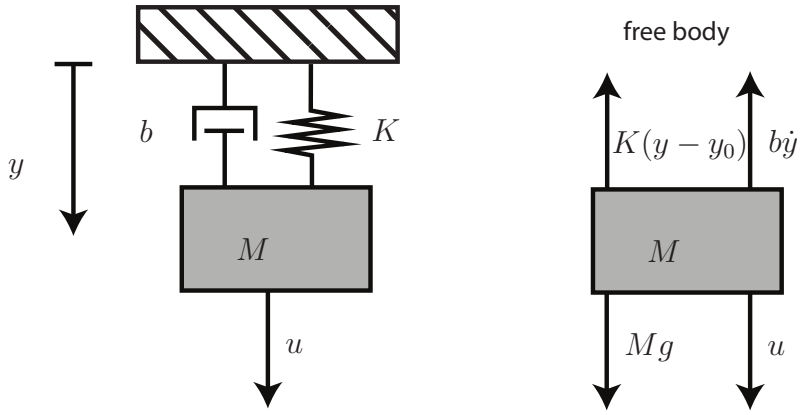

Figure 2.22: A mass spring damper.

The dynamic equation is[3]

$$M\ddot{y} = u + Mg - K(y - y_0) - b\dot{y}.$$

It's appropriate to take

$$x = (x_1, x_2), \quad x_1 = y, \ x_2 = \dot{y}$$

and then we get the equations

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{M}u + g - \frac{K}{M}x_1 + \frac{K}{M}y_0 - \frac{b}{M}x_2 \\ y &= x_1. \end{aligned}$$

[3]Here we are assuming that the spring is unstretched / uncompressed when $y = y_0$.

This has the form

$$\begin{aligned} \dot{x} &= Ax + Bu + c \\ y &= Cx, \end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{K}{M} & -\frac{b}{M} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ g + \frac{K}{M} y_0 \end{bmatrix}, \quad C = [1 \quad 0].$$

The constant vector c is known, and hence is taken as part of the system rather than as a signal. This system is nonlinear. Can you see why?

△

Example 2.5.4. Consider the RLC circuit in Figure 2.23. Apply KVL to

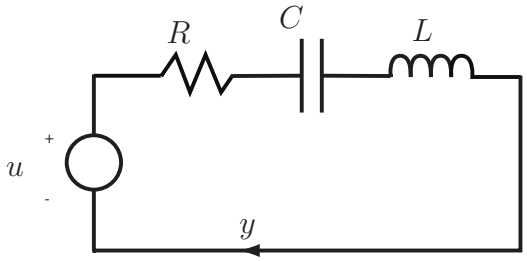

Figure 2.23: RLC circuit.

this circuit to get

$$-u + V_R + V_L + V_c = 0.$$

We substitute the device equations for the resistor, inductor and capacitor to get

$$-u + Ry + L\dot{y} + \frac{1}{C}\int_0^t y(\tau)\mathrm{d}\tau.$$

This is not a differential equation because of the integral. Therefore, take the time derivative of this equation to obtain the ODE model

$$-\dot{u} + R\dot{y} + L\ddot{y} + \frac{1}{C}y = 0.$$

For the state model, we observe that there are two energy storage elements, the inductor and the capacitor. It is natural to take the state variables to be voltage drop across C and current through L. Let

$$\begin{aligned} x_1 &:= \text{capacitor voltage} = \frac{1}{C}\int_0^t y(\tau)\mathrm{d}\tau \\ x_2 &:= \text{inductor current} = y. \end{aligned}$$

Then the KVL equation becomes

$$-u + Rx_2 + L\dot{x}_2 + x_1 = 0$$

and the capacitor equation becomes

$$\dot{x}_1 = \frac{1}{C}x_2.$$

Thus

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} u.$$

Since this is an LTI system, the state model of this system has the structure

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad A = \begin{bmatrix} 0 & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} \\ y &= Cx + Du, \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad D = 0. \end{aligned}$$

△

Example 2.5.5. Recall Example 2.4.1 where we used modularity to model a simplified car suspension system. The dynamic equations were found to be

$$\begin{aligned} M_1\ddot{y}_1 + b_1\dot{y}_1 + k_1y_1 &= b_1\dot{y}_2 + k_1y_2 \\ M_2\ddot{y}_2 + b_1\dot{y}_2 + (k_1+k_2)y_2 &= k_1y_1 + b_1\dot{y}_1 + k_2y_3. \end{aligned}$$

For mechanical systems, we take the position and velocities of the masses as our state variables. Namely, let

$$x := (x_1, x_2, x_3, x_4), \quad x_1 := y_1, \ x_2 := \dot{y}_1, \ x_3 := y_2, \ x_4 := \dot{y}_2.$$

Let the input to the system be the road height $u := y_3$. Let the output of the system be the position of the chassis $y := y_1$. We get the equations

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= \frac{1}{M_1}\left(-k_1 x_1 - b_1 x_2 + k_1 x_3 + b_1 x_4\right) \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= \frac{1}{M_2}\left(k_1 x_1 + b_1 x_2 - (k_1 + k_2) x_3 - b_1 x_4 + k_2 u\right) \\
y &= x_1.
\end{aligned}$$

This has the form

$$\begin{aligned}
\dot{x} &= Ax + Bu \\
y &= Cx,
\end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1}{M_1} & -\frac{b_1}{M_1} & \frac{k_1}{M_1} & \frac{b_1}{M_1} \\ 0 & 0 & 0 & 1 \\ \frac{k_1}{M_2} & \frac{b_1}{M_2} & -\frac{k_1+k_2}{M_2} & \frac{b_1}{M_2} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_2}{M_2} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}.$$

△

Example 2.5.6. We return to the DC motor from Example 2.4.2. The equations of motion are

$$\begin{aligned}
J\ddot{\theta} &= -b\dot{\theta} + K_\tau i \\
-v + Ri + L\frac{di}{dt} &+ K_e\dot{\theta} = 0.
\end{aligned}$$

For state variables we take the position and velocities for the mechanical subsystem and the inductor current for the electrical subsystem

$$x := (x_1, x_2, x_3), \quad x_1 := \theta, \ x_2 := \dot{\theta}, \ x_3 := i.$$

Let the input to the system be the applied voltage $u := v$. Let the output of the system be the position of the motor $y := \theta$. We get the equations

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{J}\left(-bx_2 + K_\tau x_3\right) \\ \dot{x}_3 &= \frac{1}{L}\left(-K_e x_2 - Rx_3 + u\right) \\ y &= x_1. \end{aligned}$$

This has the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K_\tau}{J} \\ 0 & -\frac{K_e}{L} & -\frac{R}{L} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

△

2.6 Linearization

Virtually *all* physical systems are nonlinear in nature. Sometimes it is possible to approximate the operation of a physical system by a linear model. This is the case, for example, if the mode of operation of the physical system does not deviate too much from the "nominal" set of operating conditions. Thus the analysis of linear systems occupies an important place in system theory and forms the basis of almost all undergraduate control education.

So far we have seen that many systems can be modeled by nonlinear state equations of the form

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x, u). \end{aligned}$$

There might be disturbance inputs present, but for now we suppose they are lumped into u. This state-space model includes the linear version (2.19) as a special case, i.e., $f(x, u) = Ax + Bu$, $h(x, u) = Cx + Du$, but also includes a much larger class of systems, like, for instance

- Robot manipulators
- Humanoid robots
- Population dynamics
- Unmanned ground / air vehicles
- Economic systems
- Chemical processes.

There are techniques for controlling nonlinear systems, but that's an advanced subject. However, many systems can be linearized about an equilibrium point. In other words, although almost every physical system contains nonlinearities, often times its behaviour *within a certain operating range of an equilibrium point* can be reasonably approximated by that of a linear model. The linear approximation is then used as the basis for control design. One reason for approximating the nonlinear system (2.18) by a linear model of the form (2.19) is that, by doing so, one can apply rather simple and systematic linear control design techniques. Keep in mind, however, that a linearized model will only be good approximation to the nonlinear model when the system operates in a sufficiently small range around an equilibrium point.

In this section we see how to do this. The idea is to use Taylor series.

Example 2.6.1. To review Taylor series, let's linearize the function $y = f(x) = x^3$ about the point $x_0 = 1$. The Taylor series expansion is

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n (x - x_0)^n, \;\; c_n = \frac{f^{(n)}(x_0)}{n!} \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \dots \; . \end{aligned}$$

Taking only terms $n = 0, 1$ gives

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0),$$

that is

$$y - y_0 \approx f'(x_0)(x - x_0).$$

Defining $\delta y = y - y_0,\ \delta x = x - x_0$, we have the linearized function $\delta y = f'(x_0)\delta x$, or $\delta y = 3\delta x$ in this case. Obviously, this approximation gets better and better as $|\delta x|$ gets smaller and smaller.

△

Taylor series extends to functions $f : \mathbb{R}^n \to \mathbb{R}^m$.

Example 2.6.2.

$$f : \mathbb{R}^3 \to \mathbb{R}^2,\ f(x_1, x_2, x_3) = (x_1x_2 - 1, x_3^2 - 2x_1x_3)$$

Suppose we want to linearize f at the point $x_0 = (1, -1, 2)$. Terms $n = 0, 1$ in the expansion are

$$f(x) \approx f(x_0) + \frac{\partial f}{\partial x}(x_0)(x - x_0),$$

where

$$\begin{aligned}
\frac{\partial f}{\partial x}(x_0) &= \text{Jacobian of } f \text{ at } x_0 \\
&= \left(\frac{\partial f_i}{\partial x_j}(x_0)\right) \\
&= \left[\begin{array}{ccc} x_2 & x_1 & 0 \\ -2x_3 & 0 & 2x_3 - 2x_1 \end{array}\right]\Bigg|_{x_0} \\
&= \left[\begin{array}{ccc} -1 & 1 & 0 \\ -4 & 0 & 2 \end{array}\right].
\end{aligned}$$

Thus the linearization of $y = f(x)$ at x_0 is $\delta y = A\delta x$, where

$$\begin{aligned}
A = \frac{\partial f}{\partial x}(x_0) &= \left[\begin{array}{ccc} -1 & 1 & 0 \\ -4 & 0 & 2 \end{array}\right] \\
\delta y = y - y_0 &= f(x) - f(x_0) \\
\delta x &= x - x_0.
\end{aligned}$$

△

By direct extension, if $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, then

$$f(x, u) \approx f(x_0, u_0) + \frac{\partial f}{\partial x}(x_0, u_0)\delta x + \frac{\partial f}{\partial u}(x_0, u_0)\delta u.$$

Now we turn to linearizing the differential equation

$$\dot{x} = f(x, u).$$

First, **assume** there is an equilibrium point, that is, a constant solution $x(t) \equiv x_0, u(t) \equiv u_0$. This is equivalent to saying that $0 = f(x_0, u_0)$. Now consider a nearby solution:

$$x(t) = x_0 + \delta x(t), u(t) = u_0 + \delta u(t), \quad \delta x(t), \delta u(t) \text{ small.}$$

We have

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ &= f(x_0, u_0) + A\delta x(t) + B\delta u(t) + \text{higher order terms} \end{aligned}$$

where

$$A := \frac{\partial f}{\partial x}(x_0, u_0), \ B := \frac{\partial f}{\partial u}(x_0, u_0).$$

Since $\dot{x} = \delta\dot{x}$ and $f(x_0, u_0) = 0$, we have the linearized equation to be

$$\dot{\delta x} = A\delta x + B\delta u.$$

Similarly, the output equation $y = h(x, u)$ linearizes to

$$\delta y = C\delta x + D\delta u,$$

where

$$C = \frac{\partial h}{\partial x}(x_0, u_0), \ D = \frac{\partial h}{\partial u}(x_0, u_0).$$

Summary

Linearizing $\dot{x} = f(x, u), \ y = h(x, u)$: Select, if one exists, an equilibrium point. Compute the four Jacobians, A, B, C, D, of f and h at the equilibrium point. Then the linearized system is

$$\dot{\delta x} = A\delta x + B\delta u, \ \delta y = C\delta x + D\delta u.$$

Under mild conditions (sufficient smoothness of f and h), this linearized system is a valid approximation of the nonlinear one in a sufficiently small neighborhood of the equilibrium point.

2.7 Simulation

Concerning the model

$$\dot{x} = f(x,u), y = h(x,u),$$

simulation involves numerically computing $x(t)$ and $y(t)$ given an initial state $x(0)$ and an input $u(t)$. If the model is nonlinear, simulation requires an ODE solver, based on, for example, the Runge-Kutta method. Scilab and MATLAB have ODE solvers and also very nice simulation GUIs, Scicos and SIMULINK, respectively.

2.8 The Laplace transform

The Laplace transform is the fundamental tool used in control systems to get to the frequency domain.[4] You have already met Laplace transforms in Math 213. So we only have to give a brief review here.

In signal processing, the two-sided Laplace transform (LT) is used, but in control only the one-sided LT is used. Let $f(t)$ be a real-valued function defined for all $t \geq 0$; it could of course be defined for all $-\infty < t < \infty$. Its Laplace transform is the integral

$$F(s) = \int_0^\infty f(t)\mathrm{e}^{-st}\mathrm{d}t =: \mathcal{L}\{f\}.$$

Here s is a complex variable. Normally, the integral converges for some values of s and not for others. That is, there is a region of convergence (ROC). It turns out that the ROC is always an open right half-plane, of the form $\{s \in \mathbb{C} : \mathrm{Re}(s) > a\}$. Then $F(s)$ is a complex-valued function of s.

Roughly speaking, the Laplace transform "decomposes" the signal $f(t)$ into a "weighted sum" of complex exponentials e^{st} where $s = \sigma + j\omega$ and $\mathrm{e}^{st} = \mathrm{e}^{(\sigma+j\omega)t} = \mathrm{e}^{\sigma t}\mathrm{e}^{j\omega t} = \mathrm{e}^{\sigma t}(\cos(\omega t) + j\sin(\omega t))$. The utility of the Laplace transform is that it converts linear, constant coefficient ordinary differential

[4] This is in contrast to communication theory where the Fourier transform dominates. Control problems frequently involve unstable systems, unbounded signals, and transient response requirements–all absent in communication problems.

equations into algebraic equations. This is because we can replace time-domain convolution with Laplace domain multiplication.

Example 2.8.1. The unit step:

$$f(t) = \mathbf{1}(t) = \begin{cases} 1 & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$$

Actually, the precise value at $t = 0$ doesn't matter. The LT is

$$F(s) = \int_0^\infty \mathrm{e}^{-st} \mathrm{d}t = -\frac{\mathrm{e}^{-st}}{s}\bigg|_0^\infty = \frac{1}{s}$$

and the ROC is

$$\mathrm{ROC} : \mathrm{Re}\,(s) > 0.$$

The same $F(s)$ is obtained if $f(t) = 1$ for all t, even $t < 0$. This is because the LT is oblivious to negative time. Notice that $F(s)$ has a pole at $s = 0$ on the western boundary of the ROC.

---△

The LT exists provided $f(t)$ satisfies two conditions. The first is that it is piecewise continuous on $t \geq 0$. This means that, on any time interval (t_1, t_2), $f(t)$ has at most a finite number of jumps, and between these jumps $f(t)$ is continuous. A square wave has this property for example. The second condition is that it is of exponential order, meaning there exist constants M, c such that $|f(t)| \leq M\mathrm{e}^{ct}$ for all $t \geq 0$. This means that if $f(t)$ blows up, at least there is some exponential that blows up faster. For example, $\exp(t^2)$ blows up too fast.

Example 2.8.2. Some other examples: An exponential:

$$f(t) = \mathrm{e}^{at}, \quad F(s) = \frac{1}{s-a}, \quad \mathrm{ROC} : \mathrm{Re}\,(s) > a.$$

A sinusoid:

$$f(t) = \cos \omega t = \frac{1}{2}\left(\mathrm{e}^{j\omega t} + \mathrm{e}^{-j\omega t}\right)$$

$$F(s) = \frac{s}{s^2 + w^2}, \ \mathrm{ROC} : \mathrm{Re}\,(s) > 0.$$

---△

The LT thus maps a class of time-domain functions $f(t)$ into a class of complex-valued functions $F(s)$. The mapping $f(t) \mapsto F(s)$ is linear.

Example 2.8.3. Let's use linearity to find the LT of the signal in Figure 2.24.

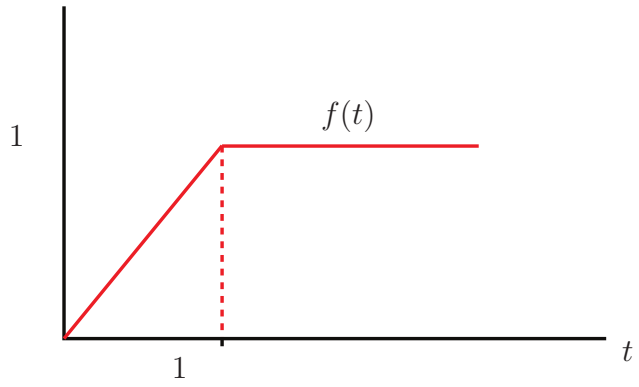

Figure 2.24: Time domain signal for Example 2.8.3.

Thus $f = f_1 + f_2$, where f_1 is the unit ramp starting at time 0 and f_2 the ramp of slope -1 starting at time 1. By linearity, $F(s) = F_1(s) + F_s(s)$. We compute that

$$F_1(s) = \frac{1}{s^2}, \ \mathrm{Re}\,(s) > 0$$
$$F_2(s) = -\frac{e^{-s}}{s^2}, \ \mathrm{Re}\,(s) > 0.$$

Thus

$$F(s) = \frac{1 - e^{-s}}{s^2}, \ \mathrm{Re}\,(s) > 0.$$

---△

There are tables of LTs. So in practice, if you have $F(s)$, you can get $f(t)$ using a table. In some cases, given $F(s)$, you can find $f(t)$ using partial fraction expansions. If

$$F(s) = \frac{N(s)}{D(s)}$$

where N and D are polynomials and the degree of N is strictly less than the degree of D, then we can use partial fraction expansions and the linearity of the LT to obtain $f(t)$.

Example 2.8.4. Given $F(s) = \frac{3s+17}{s^2-4}$, let us find $f(t)$. We don't need the ROC to find $f(t)$, but actually we know what it is. Because we're using the one-sided LT the ROC must be a right half-plane, and because $F(s)$ must be analytic within its ROC, the ROC of $F(s)$ must be $\mathrm{Re}\,(s) > 2$. We have

$$F(s) = \frac{3s+17}{s^2-4} = \frac{c_1}{s-2} + \frac{c_2}{s+2},\ c_1 = \frac{23}{4},\ c_2 = -\frac{11}{4}$$

and therefore

$$f(t) = c_1 \mathrm{e}^{2t} + c_2 \mathrm{e}^{-2t},\ t \geq 0.$$

We do not know if $f(t)$ equals zero for $t < 0$.

---△

Example 2.8.5. Given

$$F(s) = \frac{s-2}{(s-1)(s+2)},$$

let us find $f(t)$. Once again, we don't need the ROC to find $f(t)$, but we can deduce that it is the right half-plane $\mathrm{Re}\,(s) > 2$. We have

$$F(s) = \frac{s-2}{(s-1)(s+2)} = \frac{c_1}{s-1} + \frac{c_2}{s+2},\ c_1 = -\frac{1}{3},\ c_2 = \frac{4}{3}$$

and therefore

$$f(t) = -\frac{1}{3}\mathrm{e}^{2t} + \frac{4}{3}\mathrm{e}^{-2t},\ t \geq 0.$$

---△

Example 2.8.6. Given

$$F(s) = \frac{s+1}{s(s+2)^2}$$

we write

$$F(s) = \frac{s+1}{s(s+2)^2} = \frac{c_1}{s} + \frac{c_2}{s+2} + \frac{c_3}{(s+2)^2}.$$

This implies

$$s + 1 = c_1(s+s)^2 + c_2 s(s+2) + c_3 s.$$

Comparing coefficients we get three linear equations. You can verify that

$$c_1 = \frac{1}{4},\ c_2 = -\frac{1}{4},\ c_3 = \frac{1}{2}.$$

Therefore

$$f(t) = \frac{1}{4} - \frac{1}{4}\mathrm{e}^{-2t} + \frac{1}{2}t\mathrm{e}^{-2t},\ t \geq 0.$$

△

One use of the LT is in solving initial-value problems involving linear, constant-coefficient differential equations. In control engineering we do this by simulation. But let us look briefly at the LT method.

It is useful to note that if we have the LT pair

$$f(t) \leftrightarrow F(s)$$

and f is continuously differentiable at $t = 0$, then

$$\dot{f}(t) \leftrightarrow sF(s) - f(0). \tag{2.20}$$

To prove this formula, start with the LT of $\dot{f}(t)$ and integrate by parts:

$$\int_0^\infty \mathrm{e}^{-st}\dot{f}(t)\mathrm{d}t = \mathrm{e}^{-st}f(t)\big|_0^\infty + s\int_0^\infty \mathrm{e}^{-st}f(t)\mathrm{d}t.$$

Now s is such that $\mathrm{e}^{-st}f(t)$ converges to 0 as t goes to ∞. Thus the right-hand side of the preceding equation becomes

$$-f(0) + sF(s).$$

This proves (2.20).

Example 2.8.7. Consider the initial-value problem

$$\dot{y} - 2y = t,\ y(0) = 1.$$

The range of t for the differential equation isn't stated. Let us first assume that $y(t)$ is continuously differentiable at $t = 0$. This implies the differential equation holds at least for $-\varepsilon < t < \infty$ for some positive ε, that is, it holds for a little time before $t = 0$. Then we can apply (2.20) to get

$$sY(s) - 1 - 2Y(s) = \frac{1}{s^2}.$$

Solving for $Y(s)$ we get

$$Y(s) = \frac{s^2+1}{s^2(s-2)} = -\frac{1}{4s} - \frac{1}{2s^2} + \frac{5}{4(s-2)}.$$

Therefore

$$y(t) = -\frac{1}{4} - \frac{1}{2}t + \frac{5}{4}\mathrm{e}^{2t}. \tag{2.21}$$

Since we used the one-sided LT, which is oblivious to negative time, we can assert that (2.21) holds at least for $t \geq 0$. Note that it satisfies $y(0) = 1$.

On the other hand, suppose instead that the problem is posed as follows:

$$\dot{y} - 2y = t, \quad t > 0; \quad y(0) = 1.$$

That is, the differential equation holds for positive time and the initial value of y is 1. We aren't told explicitly that $y(t)$ is continuously differentiable at $t = 0$, but we are justified in making that assumption, since any other solution, for example the one satisfying $y(t) = 0$ for $t < 0$, satisfies (2.21) for $t \geq 0$.

△

If $f(t)$ is twice continuously differentiable at $t = 0$, then

$$\ddot{f}(t) \longleftrightarrow s^2F(s) - sf(0) - \dot{f}(0).$$

Example 2.8.8. The equation

$$\ddot{y} + 4\dot{y} + 3y = \mathrm{e}^t,\ y(0) = 0,\ \dot{y}(0) = 2$$

can be solved as follows. We assume $y(t)$ is twice continuously differentiable at $t = 0$. Then

$$s^2Y(s) - 2 + 4sY(s) + 3Y(s) = \frac{1}{s-1}.$$

So

$$\begin{aligned} Y(s) &= \frac{2s-1}{(s-1)(s+1)(s+3)} \\ &= \frac{1}{8}\frac{1}{s-1} + \frac{3}{4}\frac{1}{s+1} - \frac{7}{8}\frac{1}{s+3} \\ y(t) &= \frac{1}{8}\mathrm{e}^{t} + \frac{3}{4}\mathrm{e}^{-t} - \frac{7}{8}\mathrm{e}^{-3t}. \end{aligned}$$

The same solution would be arrived at for $t > 0$ if, instead of assuming $y(t)$ is twice continuously differentiable at $t = 0$, we were to allow jumps at $t = 0$.
△

The LT of the product $f(t)g(t)$ of two functions is not equal to $F(s)G(s)$, the product of the two transforms. Then what operation in the time domain does correspond to multiplication of the transforms? The answer is convolution. Let $f(t), g(t)$ be defined on $t \geq 0$. Define a new function

$$h(t) = \int_0^t f(t-\tau)g(\tau)\mathrm{d}\tau,\ t \geq 0.$$

We say h is the convolution of f and g. Note that another equivalent way of writing h is

$$h(t) = \int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau,\ t \geq 0.$$

We also frequently use the star notation $h = f * g$ or $h(t) = f(t) * g(t)$.

Theorem 2.8.1. *The LT of $f * g$ is $F(s)G(s)$.*

Proof. Let $h := f * g$. Then the LT of h is

$$H(s) = \int_0^\infty h(t)\mathrm{e}^{-st}\mathrm{d}t.$$

Substituting for h we have

$$H(s) = \int_0^\infty \int_0^t f(t-\tau)g(\tau)\mathrm{e}^{-st}\mathrm{d}\tau\mathrm{d}t.$$

Now change the order of integration:

$$H(s) = \int_0^\infty \int_\tau^\infty f(t-\tau)\mathrm{e}^{-st}\mathrm{d}t g(\tau)\mathrm{d}\tau.$$

In the inner integral change variables, $r = t - \tau$:

$$H(s) = \int_0^\infty \left(\int_0^\infty f(r)\mathrm{e}^{-sr}\mathrm{d}r \right) e^{-s\tau} g(\tau)\mathrm{d}\tau.$$

Thus

$$H(s) = \int_0^\infty F(s)\mathrm{e}^{-s\tau} g(\tau)\mathrm{d}\tau.$$

Pull $F(s)$ out and you get $H(s) = F(s)G(s)$. □

Let's summarize the properties of the LT

Properties of Laplace transforms Let $f(t)$ and $g(t)$ be real-valued functions and let a be a real constant.

(i) $\mathcal{L}\{f+g\} = \mathcal{L}\{f\} + \mathcal{L}\{g\}$

(ii) $\mathcal{L}\{af\} = a\mathcal{L}\{f\}$

(iii) $\mathcal{L}\left\{\frac{df}{dt}\right\} = s\mathcal{L}\{f\} - f(0)$

(iv) $\mathcal{L}\{f * g\} = \mathcal{L}\{f\}\mathcal{L}\{g\}$

(v) $\mathcal{L}\left\{\int_0^t f(\tau)d\tau\right\} = \frac{1}{s}\mathcal{L}\{f\}$

2.9 Transfer functions

Linear time-invariant (LTI) systems, and only LTI systems, have transfer functions. The transfer function (TF) of an LTI system is defined to be the ratio $Y(s)/U(s)$ where the LTs are taken with zero initial conditions.

Example 2.9.1. Consider the differential equation

$$\dot{y} + (\sin t)y = u.$$

This does not have a transfer function—the system is not time invariant.

△

Example 2.9.2. Consider an RC filter as in Figure 2.25.

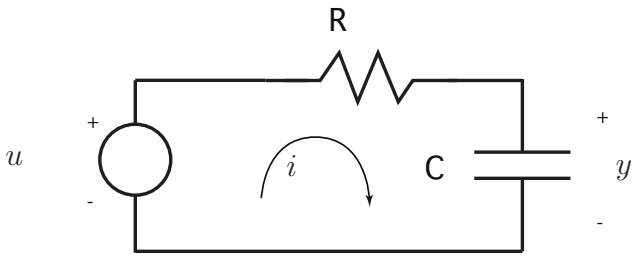

Figure 2.25: An RC filter.

The circuit equations are

$$-u + Ri + y = 0, \quad i = C\frac{dy}{dt}.$$

Therefore

$$RC\dot{y} + y = u.$$

Apply Laplace transforms with zero initial conditions:

$$sRCY(s) + Y(s) = U(s).$$

Therefore the TF is

$$\frac{Y(s)}{U(s)} = \frac{1}{RCs + 1}.$$

Or, by the voltage-divider rule using impedances

$$\frac{Y(s)}{U(s)} = \frac{\frac{1}{Cs}}{R + \frac{1}{Cs}} = \frac{1}{\mathrm{RC}s + 1}.$$

This transfer function is rational, a ratio of polynomials.

△

Example 2.9.3. A mass-spring-damper:

$$M\ddot{y} = u - Ky - D\dot{y}.$$

We get

$$\frac{Y(s)}{U(s)} = \frac{1}{Ms^2 + Ds + K}.$$

This transfer function also is rational.

△

Let's look at some other transfer functions: $G(s) = 1$ represents a pure gain; $G(s) = 1/s$ is the ideal integrator; $G(s) = 1/s^2$ is the double integrator; $G(s) = s$ is the differentiator differentiator; $G(s) = \mathrm{e}^{-sT}$ with $T > 0$ is a time delay system—note that the TF is not rational;

$$G(s) = \frac{K\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

is the standard second order TF, where $K > 0$, $\omega_n > 0$ and $\zeta \geq 0$; and

$$G(s) = K_1 + \frac{K_2}{s} + K_3 s$$

is the TF of the proportional-integral-derivative (PID) controller.

A transfer function

$$G(s) = \frac{N(s)}{D(s)}$$

is rational if it is the ratio of polynomials. We say a transfer function $G(s)$ is proper if the degree of the denominator is at least that of the numerator. The transfer functions $G(s) = 1$, $\frac{1}{s+1}$ are proper, $G(s) = s$ is not. We say $G(s)$ is strictly proper if the degree of the denominator is greater than that of the

numerator. Note that if $G(s)$ is proper then $\lim_{|s|\to\infty} G(s)$ exists; if strictly proper then $\lim_{|s|\to\infty} G(s) = 0$. If the degree of the numerator is greater than the degree of the denominator then we say $G(s)$ is improper.

If the polynomials $N(s)$ and $D(s)$ are coprime[5] then the roots of $D(s)$ are called the poles of $G(s)$. The roots of $N(s)$ are called the zeros of $G(s)$.

2.9.1 Obtaining a transfer function from a state model

Let's find the transfer function for the LTI state model

$$\begin{aligned}\dot{x} &= Ax + Bu\\ y &= Cx + Du.\end{aligned}$$

Take Laplace transforms with zero initial conditions

$$\begin{aligned}sX(s) &= AX(s) + BU(s)\\ Y(s) &= CX(s) + DU(s).\end{aligned}$$

Eliminate $X(s)$

$$\begin{aligned}&(sI - A)X(s) = BU(s)\\ &\Rightarrow X(s) = (sI - a)^{-1}BU(s)\\ &\Rightarrow Y(s) = \underbrace{\big(C(sI - A)^{-1} + D\big)}_{\text{transfer function}} U(s).\end{aligned}$$

This leads to the realization problem: Given $G(s)$, find A, B, C, D such that

$$G(s) = C(sI - A)^{-1}B + D.$$

A solution exists if and only if $G(s)$ is rational (a ratio of polynomials) and proper. The solution is never unique.

Summary

Let us recap our procedure for getting the transfer function of a system:

1. Apply the laws of physics etc. to get differential equations governing the behaviour of the system. Put these equations in state form. In general these are nonlinear.

[5] Two polynomials p and q are coprime if they have no common roots.

2. Find an equilibrium, if there is one. If there are more than one equilibria, you have to select one. If there isn't even one, this method doesn't apply.

3. Linearize about the equilibrium point.

4. Take Laplace transform of the linearized system with zero initial state.

5. Solve for the output $Y(s)$ in terms of the input $U(s)$.

The transfer function from input to output satisfies

$$Y(s) = G(s)U(s).$$

In general $G(s)$ is a matrix. In the SISO case, $G(s)$ is a scalar-valued transfer function.

Appendix 1: Linear functions

In this course we deal only with linear systems. Or rather, whenever we get a nonlinear system, we linearize it as soon as possible. So we had all better be very certain about what a linear function is. Let us recall even what a function is. If X and Y are two sets, a function from X to Y is a rule that assigns to every element of X an unambiguous element of Y—there cannot be two possible different values for some x in X. The terms function, mapping, and transformation are synonymous. The notation

$$f : X \longrightarrow Y$$

means that f is a function from X to Y. We typically write $y = f(x)$ for a function. To repeat, for each x there must be one and only one y such that $y = f(x)$; $y_1 = f(x)$ and $y_2 = f(x)$ with $y_1 \neq y_2$ is not allowed.

Let f be a function $\mathbb{R} \longrightarrow \mathbb{R}$. This means that f takes a real variable x and assigns a real variable y, written $y = f(x)$. So f has a graph in the (x, y) plane. To say that f is linear means the graph is a straight line through the origin; there's only one straight line that is not allowed—the y-axis. Thus $y = ax$ defines a linear function for any real constant a; the equation defining the y-axis is $x = 0$. The function $y = 2x + 1$ is not linear—its graph is a straight line, but not through the origin.

In your linear algebra course you were taught that a linear function is a function f from a vector space $\mathcal{X}$ to another (or the same) vector space $\mathcal{Y}$ having the property

$$f(a_1x_1 + a_2x_2) = a_1f(x_1) + a_2f(x_2)$$

for all vectors x_1, x_2 in $\mathcal{X}$ and all real numbers[6] a_1, a_2. If the vector spaces are $\mathcal{X} = \mathbb{R}^n$, $\mathcal{Y} = \mathbb{R}^m$, and if f is linear, then it has the form $f(x) = Ax$, where A is an $m \times n$ matrix. Conversely, every function of this form is linear.

This concept extends beyond vectors to signals. For example, consider a capacitor, whose constitutive law is

$$i = C\frac{dv}{dt}.$$

Here, i and v are not constants, or vectors—they are functions of time. If we think of the current i as a function of the voltage v, then the function is linear. This is because

$$C\frac{d(a_1v_1 + a_2v_2)}{dt} = a_1C\frac{dv_1}{dt} + a_2C\frac{dv_2}{dt}.$$

On the other hand, if we try to view v as a function of i, then we have a problem, because we need in addition an initial condition $v(0)$ (or some other initial time) to uniquely define v, not just i. Let us set $v(0) = 0$. Then v is a linear function of i. You can see this from the integral form of the capacitor equation:

$$v(t) = \frac{1}{C}\int_0^t i(\tau)d\tau.$$

Appendix 2 : The impulse and convolution

Now we take a little time to discuss the problematical object, the impulse $\delta(t)$. The impulse, also called the Dirac delta function, is not really a legitimate function, because its "value" at $t = 0$ is not a real number. And you can't rigorously get δ as the limit of a sequence of ever-narrowing rectangles, because that sequence does not converge in any ordinary sense. Yet the impulse is a useful concept in system theory and so we have to make it legitimate. The French mathematician Laurent Schwartz worked out a very nice,

[6] This definition assumes that the vector spaces are over the field of real numbers.

consistent way of dealing with the impulse, and more general "functions." The impulse is an example of a distribution.

The main idea is that $\delta(t)$ is not a function, but rather it is a way of defining the linear transformation $\phi \mapsto \phi(0)$ that maps a signal $\phi(t)$ to its value at $t = 0$. This linear transformation should properly be written as $\delta(\phi) = \phi(0)$ (i.e., δ transforms ϕ to $\phi(0)$) but historically it has been written as

$$\int_{-\infty}^{\infty} \phi(t)\delta(t)\mathrm{d}t = \phi(0).$$

You know this as the "sifting formula." Let us emphasize that the expression

$$\int_{-\infty}^{\infty} \phi(t)\delta(t)\mathrm{d}t \tag{2.22}$$

is not intended to mean integration of the product $\phi(t)\delta(t)$ of functions—δ isn't a function; rather, the expression is that of an operation on ϕ whose value is defined to be $\phi(0)$. The expression is defined to be valid for all functions $\phi(t)$ that are smooth at $t = 0$; smooth means continuously differentiable up to every order.

Needless to say, we have to be careful with δ; for example, there's no way to make sense of δ^2 because the expression (2.22) is not valid for $\phi = \delta$, again, because δ isn't a function, let alone a smooth one. Because the unit step is not smooth at $t = 0$, $\mathbf{1}(t)\delta(t)$ is undefined too. However, (2.22) does apply for $\phi(t) = \mathrm{e}^{-st}$, because it is smooth:

$$\int_{-\infty}^{\infty} e^{-st}\delta(t)\mathrm{d}t = 1.$$

Thus the LT of δ equals 1.

Initial-value problems involving δ, such as

$$\dot{y} + 2y = \delta,$$

or worse,

$$\dot{y} + 2y = \dot{\delta},$$

require more advanced theory, because $y(t)$ cannot be continuously differentiable at $t = 0$. Instead of pursuing this direction, since initial-value problems are of minor concern in control we turn instead to convolution equations.

As you learned in signals and systems, linear time-invariant systems are modeled in the time domain by a convolution equation:

$$y(t) = \int_{-\infty}^{\infty} g(t-\tau)u(\tau)\mathrm{d}\tau.$$

If $g(t)$ is a smooth function for all t, then expression (2.22) applies and we get

$$g(t) = \int_{-\infty}^{\infty} g(t-\tau)\delta(\tau)\mathrm{d}\tau.$$

Thus $g(t)$ equals the output when the input is the unit impulse. We call $g(t)$ the impulse-response function, or the impulse response. The case where g itself equals δ isn't covered by what has been said so far, but distribution theory can be used to justify $\delta * \delta = \delta$.

Example 2.9.4. Consider a low pass RC filter:

$$G(s) = \frac{1}{RCs+1}.$$

The inverse LT of $G(s)$ is

$$\frac{1}{RC}\mathrm{e}^{-\frac{t}{RC}}.$$

Since the filter is causal, to get the impulse response we should take that time-function that is zero for $t < 0$:

$$g(t) = \frac{1}{RC}\mathrm{e}^{-\frac{t}{RC}}\mathbf{1}(t).$$

For the high pass filter,

$$G(s) = \frac{RCs}{RCs+1}$$

$$g(t) = \delta(t) - \frac{1}{RC}\ \mathrm{e}^{-t/RC}\mathbf{1}(t).$$

△

In general, if $G(s)$ is strictly proper, then $g(t)$ is a regular function. If $G(s)$ is proper but not strictly proper, then $g(t)$ contains an impulse.

As a final comment, some texts (Dorf and Bishop [1]) give the LT pair

$$\dot{f}(t) \longleftrightarrow sF(s) - f(0^-)$$

indicating that $f(0)$ and $f(0^-)$ may be different, that is, $f(t)$ may be discontinuous at $t = 0$. If that is the case, then $\dot{f}$ has an impulse and distribution theory needs to be invoked.

Summary

Rigorously speaking, the impulse is not a function. Instead, it is defined only indirectly via the sifting formula. However, in the subject of control systems, and other signals and systems subjects, the impulse is used just as though it were a function. Usually no harm is done, i.e., you get the right answer. But sometimes questions arise that require a rigorous way to answer them; distribution theory is that way. For example, the derivative of the impulse, $\dot{\delta}(t)$, can be defined rigorously, and its LT is s. In this way we can obtain the impulse response of the differentiator $y = \dot{u}$—it is $g(t) = \dot{\delta}(t)$. The transfer function is $G(s) = s$.