Chapter 2

System Modeling

2.1 Introduction

In this chapter we introduce the notion of a dynamical system and describe how to model system systems. Roughly speaking, a dynamical system is one in which the effects of actions do not occur immediately. For example, the velocity of a car does not change immediately when the gas pedal is pushed nor does the temperature in a room rise instantaneously when an air conditioner is switched on. Similarly, a headache does not vanish right after an aspirin is taken, requiring time to take effect. In business systems, increased funding for a development project does not increase revenues in the short term, although it may do so in the long term (if it was a good investment). All of these are examples of dynamical systems, in which the behavior of the system evolves with time.

Modeling is the method by which we deseribe a dynamical system in a precise mathematical form, for the purpose of analysis and simulation. A model of a system is a *representation* of the system dynamics and it is used to answer questions about that system. The model we choose depends on the questions that we wish to answer, and so there may be multiple models for a single physical system, with different levels of fidelity depending on the phenomena of interest. In this chapter we provide an introduction to the concept of modeling, and provide some basic material on two specific methods that are commonly used in feedback and control systems: differential equations and different equations.

2.2 Two Views on Dynamics

Dynamical systems can be viewed from two different ways: the internal view or the external view. The internal view which attempts to describe the internal workings of the system originates from classical mechanics. The prototype problem was the problem to describe the motion of the planets. For this problem it was natural to give a complete characterization of the motion of all planets. This involves careful analysis of the effects of gravitational pull and the relative positions of the planets in a system.

The other view on dynamics originated in electrical engineering. The prototype problem was to describe electronic amplifiers. It was natural to view an amplifier as a device that transforms input voltages to output voltages and disregard the internal detail of the amplifier. This resulted in the input-output view of systems. The two different views have been amalgamated in control theory. Models based on the internal view are called internal descriptions, state models, or white box models. The external view is associated with names such as external descriptions, input-output models or black box models. In this book we will mostly use the words state models and input-output models.

The Heritage of Mechanics

Dynamics originated in the attempts to describe planetary motion. The basis was detailed observations of the planets by Tycho Brahe and the results of Kepler who found empirically that the orbits could be well described by ellipses. Newton embarked on an ambitious program to try to explain why the planets move in ellipses and he found that the motion could be explained by his law of gravitation and the formula that force equals mass times acceleration. In the process he also invented calculus and differential equations. Newtons results was the first example of the idea of reductionism, i.e. that seemingly complicated natural phenomena can be explained by simple physical laws. This became the paradigm of natural science for many centuries.

One of the triumphs of Newton's mechanics was the observation that the motion of the planets could be predicted based on the current positions and velocities of all planets. It was not necessary to know the past motion. The *state* of a dynamical system is a collection of variables that characterize the motion of a system completely for the purpose of predicting future motion. For a system of planets the state is simply the positions and the velocities of the planets. A mathematical model simply gives the rate of change of the

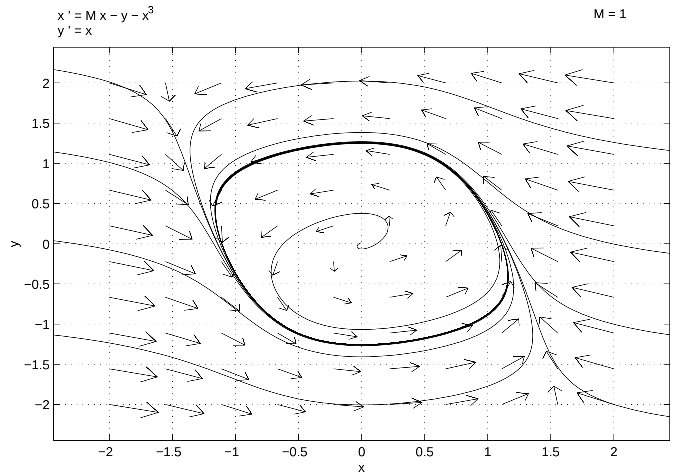

Figure 2.1: Illustration of a state model. A state model gives the rate of change of the state as a function of the state. The velocity of the state are denoted by arrows.

state as a function of the state itself, i.e. a differential equation.

$$\frac{dx}{dt} = f(x) \tag{2.1}$$

This is illustrated in Figure 2.1 for a system with two state variables. The particular system represented in the figure is the van der Pol equation:

$$\frac{dx_1}{dt} = x_1 - x_1^3 - x_2$$
$$\frac{dx_2}{dt} = x_1,$$

which is a model of an electronic oscillator. The model (2.1) gives the velocity of the state vector for each value of the state. These are represented by the arrows in the figure. The figure gives a strong intuitive representation of the equation as a vector field or a flow. Systems of second order can be represented in this way. It is unfortunately difficult to visualize equations of higher order in this way.

The ideas of dynamics and state have had a profound influence on philosophy where it inspired the idea of predestination. If the state of a natural system is known at some time, its future development is complete determined. The vital development of dynamics has continued in the 20th century. One of the interesting outcomes is chaos theory. It was discovered that

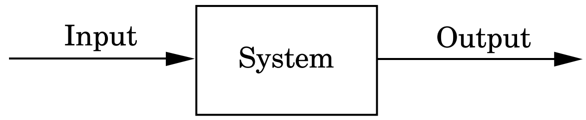

Figure 2.2: Illustration of the input-output view of a dynamical system.

there are simple dynamical systems that are extremely sensitive to initial conditions, small perturbations may lead to drastic changes in the behavior of the system. The behavior of the system could also be extremely complicated. The emergence of chaos also resolved the problem of determinism, even if the solution is uniquely determined by the initial conditions it is in practice impossible to make predictions because of the sensitivity of initial conditions.

The differential equation (2.1) is called an autonomous system because there are no external influences. Such a model is natural to use for celestial mechanics, because it is difficult to influence the motion of the planets. The situation in control is quite different because the external influences are quite important. One way to capture this is to replace equation (2.1) by

$$\frac{dx}{dt} = f(x, u) \tag{2.2}$$

where u represents the effect of external influences. The model (2.2) is called a controlled differential equation. The model implies that the velocity of the state can be influenced by the *input* u. Adding the input makes the model richer. New questions arises, for example, what influence can the control variable have on the trajectories of the system? Is it possible to reach all points in the state space by proper choices of the control?

The Heritage of Electrical Engineering

A very different view of dynamics emerged from electrical engineering. The prototype problem was design of electronic amplifiers. Since an amplifier is a device for amplification of signals it is natural to focus on the input-output behavior. A system was considered as a device that transformed inputs to outputs, as illustrated in Figure Figure 2.2. Conceptually an input-output model can be viewed as a giant table of inputs and outputs. The input-output view is particularly useful for the special class of *linear* systems. To define linearity we let $(u_1,\ y_1)$ and $(u_2,\ y_2)$ denote two input-output pairs, and a and b be real numbers. A system is linear if $(au_1 + bu_2,\ ay_1 + by_2)$

is also an input-output pair (superposition). A nice property of control problems is that they can often be modeled by linear, time-invariant systems. Chapter **??** provides a much more detailed analysis of linear systems.

Time invariance is another concept. It means that the behavior of the system at one time is equivalent to the behavior at another time. It can be expressed as follows. Let $(u,\ y)$ be an input-output pair and let u_t denote the signal obtained by shifting the signal u, t units forward. A system is called time-invariant if $(u_t,\ y_t)$ is also an input-output pair. This view point has been very useful, particularly for linear, time-invariant systems, whose input output relation can be described by

$$y(t) = \int_0^t g(t-\tau)u(\tau)d\tau. \tag{2.3}$$

where g is the impulse response of the system. If the input u is a unit step the output becomes

$$y(t) = h(t) = \int_0^t g(t-\tau)d\tau = \int_0^t g(\tau)u(\tau)d\tau \tag{2.4}$$

The function h is called the step response of the system. Notice that the impulse response is the derivative of the step response.

Another possibility to describe a linear, time-invariant system is to represent a system by its response to sinusoidal signals, this is called frequency response. A rich powerful theory with many concepts and strong, useful results have emerged. The results are based on the theory of complex variables and Laplace transforms. The input-output view lends it naturally to experimental determination of system dynamics, where a system is characterized by recording its response to a particular input, e.g. a step.

The words input-output models, external descriptions, black boxes are synonyms for input-output descriptions.

The Control View

When control emerged in the 1940s the approach to dynamics was strongly influenced by the Electrical Engineering view. The second wave of developments starting in the late 1950s was inspired by the mechanics and the two different views were merged. Systems like planets are autonomous and cannot easily be influenced from the outside. Much of the classical development of dynamical systems therefore focused on autonomous systems. In control it is of course essential that systems can have external influences.

The emergence of space flight is a typical example where precise control of the orbit is essential. Information also plays an important role in control because it is essential to know the information about a system that is provided by available sensors. The models from mechanics were thus modified to include external control forces and sensors. In control the model given by (2.5) is thus replaced by

$$\begin{aligned}\frac{dx}{dt} &= f(x,u)\\ y &= g(x,u)\end{aligned} \tag{2.5}$$

where u is a vector of control signal and y a vector of measurements. This viewpoint has added to the richness of the classical problems and led to new important concepts. For example it is natural to ask if all points in the state space can be reached (reachability) and if the measurement contains enough information to reconstruct the state.

The input-output approach was also strengthened by using ideas from functional analysis to deal with nonlinear systems. Relations between the state view and the input output view were also established. Current control theory presents a rich view of dynamics based on good classical traditions.

The importance of disturbances and model uncertainty are critical elements of control because these are the main reasons for using feedback. To model disturbances and model uncertainty is therefore essential. One approach is to describe a model by a nominal system and some characterization of the model uncertainty. The dual views on dynamics is essential in this context. State models are very convenient to describe a nominal model but uncertainties are easier to describe using frequency response.

2.3 Linear Differential Equations

In this section we provide a brief review of linear differential equations, which should be familiar to most readers. Chapter **??** provides a more detailed introduction to linear differential equations in so-called state-space form.

Consider the following description of a linear time-invariant dynamical system

$$\frac{d^n y}{dt^n} + a_1\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_n y = b_1\frac{d^{n-1}u}{dt^{n-1}} + b_2\frac{d^{n-2}u}{dt^{n-2}} + \ldots + b_n u, \tag{2.6}$$

where u is the input and y the output. The system is of order n order, where n is the highest derivative of y. The ordinary differential equations is a standard topic in mathematics. In mathematics it is common practice to have $b_n = 1$ and $b_1 = b_2 = \ldots = b_{n-1} = 0$ in (2.6). The form (2.6) adds richness and is much more relevant to control. The equation is sometimes called a controlled differential equation.

It follows from the rules for differentiation that

$$\frac{d^k}{dt^k}(\alpha y_1 + \beta y_2) = \alpha \frac{dy_1^k}{dt^k} + \beta \frac{dy_1^k}{dt^k}$$

If $(u,\ y)$ is a pair of inputs and outputs it follows that $(u',\ y')$ is also an input output pair. Simlarly, if the $(u_1,\ y_1)$ and $(u_2,\ y_2)$ are pairs of inputs and outputs that satisfy the Equation (2.6) $\alpha u_1 + \beta u_2,\ \alpha y_1 + \beta y_2)$ is also an input output pair, which is the principle of superposition.

The Homogeneous Equation

If the input u to the system (2.6) is zero, we obtain the equation

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = 0, \tag{2.7}$$

which is called the homogeneous equation associated with equation (2.6). The characteristic polynomial of Equations (2.6) and (2.7) is

$$a(s) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n \tag{2.8}$$

The roots of the characteristic equation determine the properties of the solution. If $a(\alpha) = 0$, then $y(t) = Ce^{\alpha t}$ is a solution to Equation (2.7).

If the characteristic equation has distinct roots α_k the solution is

$$y(t) = \sum_{k=1}^{n} C_k e^{\alpha_k t}, \tag{2.9}$$

where C_k are arbitrary constants. The Equation (2.7) thus has n free parameters.

Roots of the Characteristic Equation give Insight

A real root $s = \alpha$ correspond to ordinary exponential functions $e^{\alpha t}$. These are monotone functions that decrease if α is negative and increase if α is positive as is shown in Figure 2.3. Notice that the linear approximations shown in dashed lines change by one unit for one unit of αt. Complex roots

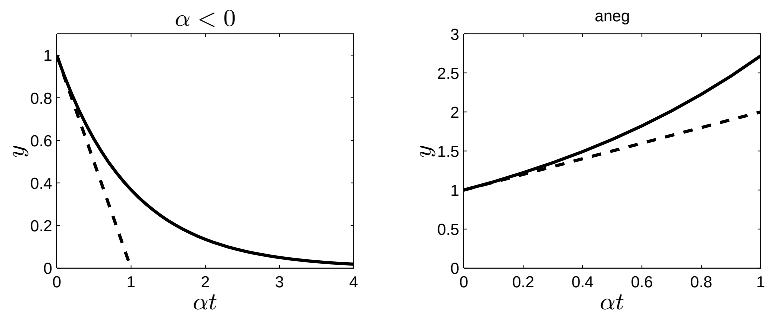

Figure 2.3: The exponential function $y(t) = e^{\alpha t}$. The linear approximations of of the functions for small αt are shown in dashed lines. The parameter $T = 1/\alpha$ is the time constant of the system.

$s = \sigma \pm i\omega$ correspond to the time functions.

$$e^{\sigma t} \sin \omega t, \qquad e^{\sigma t} \cos \omega t$$

which have oscillatory behavior, see Figure 2.4. The distance between zero crossings is π/ω and corresponding amplitude change is $e^{\sigma\pi/\omega}$.

Multiple Roots

When there are multiple roots the solution to Equation (2.7) has the form

$$y(t) = \sum_{k=1}^{n} C_k(t) e^{\alpha_k t}, \tag{2.10}$$

Where $C_k(t)$ is a polynomial with degree less than the multiplicity of the root α_k. The solution (2.10) thus has n free parameters.

The Inhomogeneous Equation – A Special Case

The equation

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = u(t) \tag{2.11}$$

has the solution

$$y(t) = \sum_{k=1}^{n} C_{k-1}(t) e^{\alpha_k t} + \int_0^t h(t-\tau) u(\tau) d\tau, \tag{2.12}$$

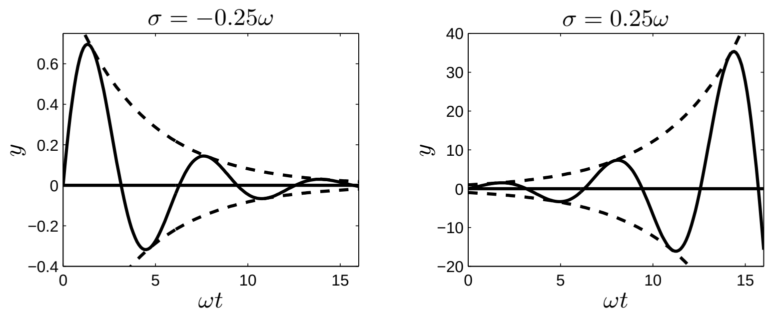

Figure 2.4: The exponential function $y(t) = e^{\sigma t}\sin \omega t$. The linear approximations of of the functions for small αt are shown in dashed lines. The dashed line corresponds to a first order system with time constant $T = 1/\sigma$. The distance between zero crossings is π/ω.

where h is the solution to the homogeneous equation (2.7), i.e.

$$\frac{d^n h}{dt^n} + a_1 \frac{d^{n-1} h}{dt^{n-1}} + \ldots + a_n h = 0 \tag{2.13}$$

with initial conditions

$$h(0) = 0, \quad h'(0) = 0, \ldots, \; h^{(n-2)}(0) = 0, \quad h^{(n-1)}(0) = 1. \tag{2.14}$$

The solution (2.12) is thus a sum of two terms, the general solution to the homogeneous equation and a particular solution which depends on the input u. The solution has n free parameters which can be determined from initial conditions.

To show that (2.12) satisfies (2.11) we first observe that the sum in (2.12) satisfies the homogeneous equation (2.7). Consider

$$v(t) = \int_0^t h(t-\tau)u(\tau)d\tau,$$

It follows from (2.14) that v(0)=0. Taking derivatives we find that

$$
\begin{aligned}
v'(t) &= \int_0^t h'(t-\tau)u(\tau)d\tau + h(0)u(t) \\
v''(t) &= \int_0^t h''(t-\tau)u(\tau)d\tau + h'(0)u(t) \\
&\vdots \\
v^{(n)}(t) &= \int_0^t h^{(n)}(t-\tau)u(\tau)d\tau + h^{(n-1)}(0)u(t)
\end{aligned}
$$

It follows from (2.13) and (2.14) that v satisfies the differential equation (2.11).

The Inhomogeneous Equation - The General Case

Having found a solution to (2.11) it is straightforward to find a solution to the general equation (2.6). If y is a solution to the (2.11) it follows that dy/dt is a solution to the differential equation.

$$
\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = \frac{du}{dt}
$$

Repeating this argument for higher derivatives we find that the Equation (2.6) has the solution

$$
y(t) = \sum_{k=1}^{n} C_{k-1}(t) e^{\alpha_k t} + \int_0^t g(t-\tau)u(\tau)d\tau, \tag{2.15}
$$

where the function g is given by

$$
g(t) = b_1 h^{(n-1)}(t) + b_2 h^{(n-2)}(t) + \ \ldots \ + b_n h(t). \tag{2.16}
$$

The solution is thus the sum of two terms, the general solution to the homogeneous equation and a particular solution. The general solution to the homogeneous equation does not depend on the input and the particular solution which depends on the input. The particular solution is given by

$$
y(t) = \int_0^t g(t-\tau)u(\tau)d\tau
$$

where g is called the *impulse response*, Notice that the impulse response has the form

$$g(t) = \sum_{k=1}^{n} c_k(t) e^{\alpha_k t}. \tag{2.17}$$

It thus has the same form as the general solution to the homogeneous equation (2.10). The coefficients c_k are given by the conditions (2.14). If the characteristic equation has distinct roots $c_k(t)$ are constants. If α_k is a root of multiplicity m then $c_k(t)$ is a polynomial of degree $m - 1$.

The impulse response is also called the weighting function because the second term of (2.15) can be interpreted as a weighted sum of past inputs.

The Step Response

Consider (2.15) and assume that all initial conditions are zero. The output is then given by

$$y(t) = \int_0^t g(t - \tau) u(\tau) d\tau, \tag{2.18}$$

If the input is constant $u(t) = 1$ we get

$$y(t) = \int_0^t g(t - \tau) d\tau = \int_0^t g(\tau) d\tau = H(t), \tag{2.19}$$

The function H is called the unit step response or the step response for short. It follows from the above equation that

$$g(t) = \frac{dh(t)}{dt} \tag{2.20}$$

The step response can easily be determined experimentally by waiting for the system to come to rest and applying a constant input. In process engineering the experiment is called a bump test. The impulse response can then be determined by differentiating the step response.

The Convolution Integral

The relation between the input and the output for a system which is initially at rest is given by Equation (2.18), i.e.

$$y(t) = \int_0^t g(t - \tau) u(\tau) d\tau.$$

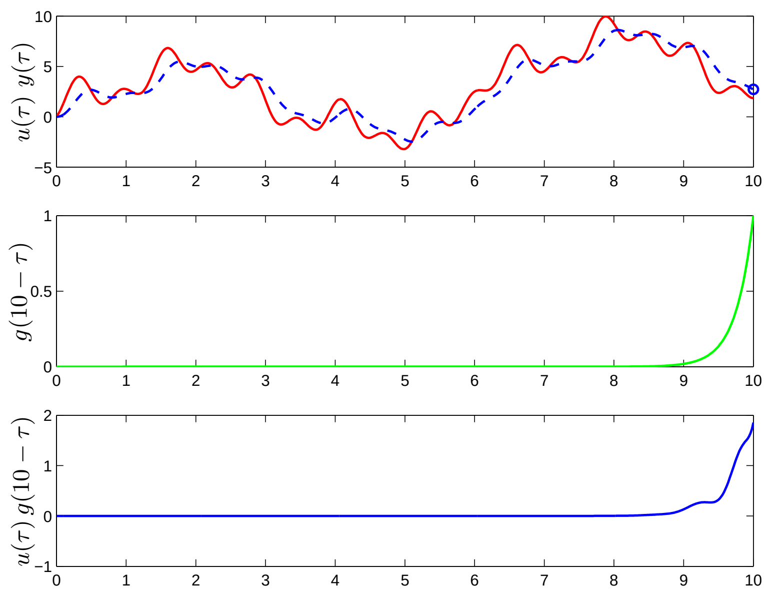

Figure 2.5: Illustration of the convolution integral for the impulse response $g(t) = e^{-4t}$. The top shows the input u in full lines and the output y in dashed lines. The lower graphs illustrates how $y(10)$ is obtained.

Mathematically the output is called a convolution of the input u and the impulse response g. This integral has a nice interpretation which is illustrated in Figure Figure 2.5. The figure illustrates that the output is obtained as a weighted average of the input. The top plot shows the input u in full lines and the output y in dashed lines. The lower graphs illustrates how the value $y(10)$ is obtained. The middle curve shows the impulse response $g(10-\tau)$ and the lower plot shows the product $u(\tau)g(10-\tau)$. The value $y(10)$ is simply the integral of $u(\tau)g(10-\tau)$. By understanding of the interpretation of the convolution integral it is easy to develop an intuitive understanding of the qualitative behavior of a system from the impulse response.

Response to Exponential Inputs

Exponential functions play an important role in linear systems. The impulse response of a linear time invariant system is for example a sum of exponentials, see (2.17). Exponential functions also appear in the general form of the solution of a linear differential equation, see (2.15). In this section we will investigate how a linear time invariant system responds to an exponential signal. Consider the system given by (2.6) and let the input be

$$u(t) = e^{\alpha t}.$$

The solution to the differential equation is a sum of the general solution to the homogeneous equation and a particular solution. We will investigate if there is a particular solution of the form

$$y(t) = y_0 e^{\alpha t}$$

Inserting this into the differential equation (2.6) we find

$$\alpha^n y_0 + a_1 \alpha^{n-1} + \cdots + a_n y_0 = b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + \cdots + b_n$$

It thus follows that there the equation has a solution of the form $y_0 e^{\alpha t}$ and that

$$y_0 = \frac{b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + \cdots + b_n}{\alpha^n + a_1 \alpha^{n-1} + \cdots + a_n} = G(\alpha)$$

where $G(s)$ is the transfer function of the system. Let λ_k be the zeros of the characteristic polynomial $a(s)$ of the system we thus find that the general solution of the differential equation is

$$y(t) = \sum_k C_k(t) e^{\lambda_k t} + G(\alpha) e^{\alpha t} \tag{2.21}$$

The particular solution corresponding to the input $e^{\alpha t}$ is thus $G(\alpha)e^{\alpha t}$. If the initial conditions are chosen as $y^k(0) = \alpha^k G(\alpha)$ the sum disappears and we get $y(t) = G(\alpha)e^{\alpha t}$. If $\mathrm{Re}\lambda_k < \alpha$ the particular solution will dominate the response for large t for arbitrary initial conditions. We thus obtain the interesting result that the number $G(\alpha)$ tells how exponential functions propagate through the system.

Equation (2.21) is valid when α is a complex number. If $\alpha = i\omega$ we find that the response to

$$u(t) = e^{i\omega t}$$

is

$$y(t) = \sum_k C_k(t) e^{\lambda_k t} + G(i\omega) e^{i\omega t}$$

We have

$$G(i\omega)e^{i\omega t} = |G(i\omega)| e^{i \arg G(i\omega)} e^{i\omega t} = |G(i\omega)| e^{i(\omega t + \arg G(i\omega))}$$

Separating the real and imaginary parts of the input and the output we find that the input $u(t) = sin\omega t$ gives the output

$$y(t) = \sum_k C_k(t)e^{\lambda_k t} + |G(i\omega)| \sin\left(\omega_t + \arg G(i\omega)\right) \tag{2.22}$$

This result is of particular interest for stable systems. For such systems we have $\lambda_k < 0$. After an initial transient the response to a sinusoidal input will thus be sinusoidal with the same frequency as the input. The output is thus amplified by the factor $|G(i\omega)|$ and the phase is shifted by $\arg G(i\omega)$ in relation to the input. This is discussed further in Section 2.4.

2.4 Frequency Response

The idea of frequency response is to characterize a linear time-invariant system by its response to sinusoidal signals. The idea goes back to Fourier, who introduced the method to investigate propagation of heat in metals. Frequency response gives an alternative way of viewing dynamics. One advantage is that it is possible to deal with systems of very high order, even infinite. This is essential when discussing sensitivity to process variations. This will be discussed in detail in Chapter **??**.

Frequency response also gives a different way to investigate stability. In Section 2.3 it was shown that a linear system is stable if the characteristic polynomial has all its roots in the left half plane. To investigate stability of a the system we have to derive the characteristic equation of the closed loop system and determine if all its roots are in the left half plane. Even if it easy to determine the roots of the equation numerically it is not easy to determine how the roots are influenced by the properties of the controller. It is for example not easy to see how to modify the controller if the closed loop system is stable. The way stability has been defined it is also a binary property, a system is either stable or unstable. In practice it is highly desirable to have a notion of the degrees of stability. All of these issues can be related to frequency response. The key is Nyquist's stability criterion which is a frequency response concept. Frequency response was one of the key ideas that formed the foundation of control.

Response to a Sinusoidal Input

The response of linear systems to sinusoids was discussed in Section 2.3, see Equation (2.22). Consider a system with the transfer function $G(s)$ which

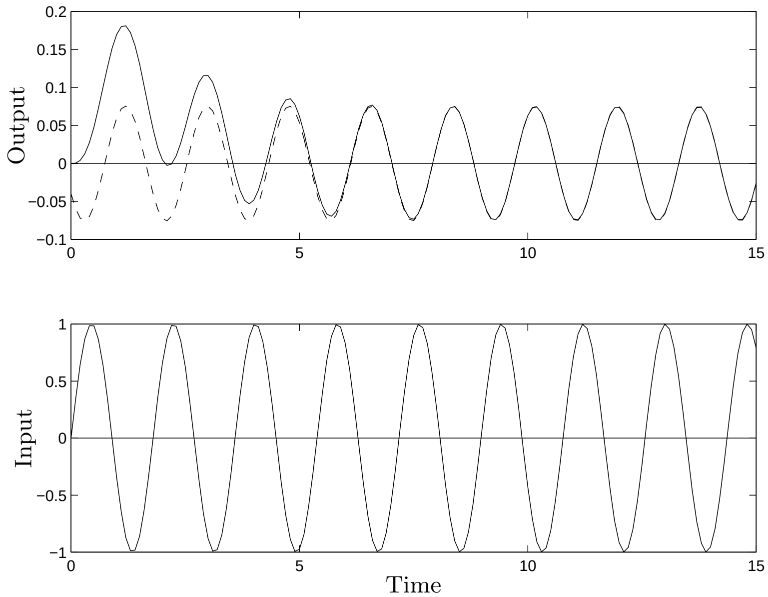

Figure 2.6: Response of a linear time-invariant system to a sinusoidal input (full lines). The system has the transfer function $G(s) = 1/(s+1)^2$. The dashed line shows the steady state output calculated from (2.23).

has poles λ_k. The output corresponding to the input $u(t) = \sin \omega t$ is

$$y(t) = \sum_k C_k(t) e^{\lambda_k t} + |G(i\omega)| \sin(\omega_t + \arg G(i\omega))$$

If the system is stable, i.e. $\mathrm{R}e\lambda_k < 0$ for all k, the first term will decay exponentially and the solution will converge to the steady state response given by

$$y(t) = |G(i\omega)| \sin(\omega_t + \arg G(i\omega)) \tag{2.23}$$

This is illustrated in Figure 2.6 which shows the response of a linear time-invariant system to a sinusoidal input. The figure shows the output of the system when it is initially at rest and the steady state output given by (2.23). The figure shows that after a transient the output is indeed a sinusoid with the same frequency as the input.

The steady state response to a sinusoid is completely characterized by the function $G(i\omega)$ which is called the frequency response of the system. The argument of the function is frequency ω and the function takes complex values. The magnitude $|G(i\omega)|$ is called the gain and the angle $\arg G(i\omega)$ is called the phase. The phase is often negative and the quantity -arg $G(i\omega)$, called the phase lag, is therefore also used. The gain $|G(i\omega)|$ is a generalization of the static gain $G(0)$ which tells steady state output when the input is a constant. It is thus possible to talk about the gain of the system for signals of different frequencies. The propagation of any signal can then be obtained by representing it as a sum of sinusoids, investigating each sinusoid individually and adding the outputs using superposition.

The frequency response can be determined experimentally by injecting a sinusoid and measuring the ratio of the amplitudes and the phase shift between input and output. Very accurate measurements are possible by using correlation techniques. This is very important in practice because it may be very time consuming or even impossible to obtain a mathematical model from first principles.

2.5 State Models

The state is a collection of variables that summarize the past of a system for the purpose of prediction the future. For an engineering system the state is composed of the variables required to account for storage of mass, momentum and energy. An key issue in modeling is to decide how accurate storage has to be represented. The state variables are gathered in a vector, the state vector $x \in R^n$. The control variables are represented by another vector $u \in R^p$ and the measured signal by the vector $y \in R^q$. A system can then be represented by the model

$$\begin{aligned} \frac{dx}{dt} &= f(x, u) \\ y &= g(x, u) \end{aligned} \tag{2.24}$$

The dimension of the state vector is called the order of the system. The system is called time-invariant because the functions f and g do not depend explicitly on time t. It is possible to have more general time-varying systems where the functions do depend on time. The model thus consists of two functions. The function f gives the velocity of the state vector as a function of state x, control u and time t and the function g gives the measured values as functions of state x, control u and time t. The function f is called the velocity function and the function g is called the sensor function or the measurement function. A system is called linear if the functions f and g are linear in x and u. A linear system can thus be

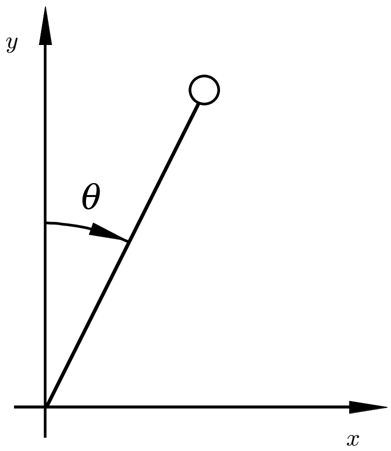

Figure 2.7: An inverted pendulum. *The picture should be mirrored.*

represented by

$$\begin{aligned}\frac{dx}{dt} &= Ax + Bu\\ y &= Cx + Du\end{aligned}$$

where A, B, C and D are constant varying matrices. Such a system is said to be linear and time-invariant, or LTI for short. The matrix A is called the dynamics matrix, the matrix B is called the control matrix, the matrix C is called the sensor matrix and the matrix D is called the direct term. Frequently systems will not have a direct term indicating that the control signal does not influence the output directly. We will illustrate by a few examples.

Example 1 *(The Double Integrator).* Consider a system described by

$$\begin{aligned}\frac{dx}{dt} &= \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} x + \begin{bmatrix}0 & 1\end{bmatrix} u\\ y &= \begin{bmatrix}1 & 0\end{bmatrix} x\end{aligned} \tag{2.25}$$

This is a linear time-invariant system of second order with no direct term.

Example 2 *(The Inverted Pendulum).* Consider the inverted pendulum in Figure 2.7. The state variables are the angle $\theta = x_1$ and the angular velocity $d\theta/dt = x_2$, the control variable is the acceleration ug of the pivot, and the output is the angle θ.

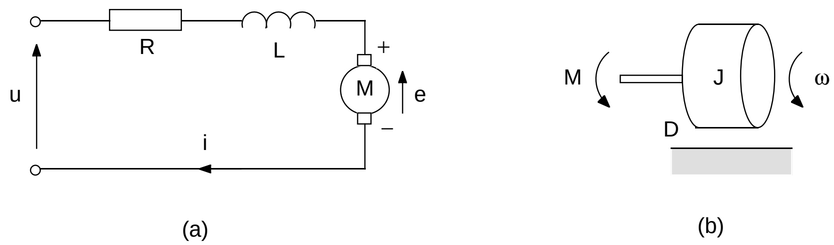

Figure 2.8: Schematic diagram of an electric motor.

Newton's law of conservation of angular momentum becomes

$$J\frac{d^2\theta}{dt^2} = mgl\sin\theta + mul\cos\theta$$

Introducing $x_1 = \theta$ and $x_2 = d\theta/dt$ the state equations become

$$\begin{aligned} \frac{dx}{dt} &= \begin{bmatrix} x_2 \\ \frac{mgl}{J}\sin x_1 + \frac{mlu}{J}\cos x_1 \end{bmatrix} \\ y &= x_1 \end{aligned}$$

It is convenient to normalize the equation by choosing $\sqrt{J/mgl}$ as the unit of time. The equation then becomes

$$\begin{aligned} \frac{dx}{dt} &= \begin{bmatrix} x_2 \\ \sin x_1 + u\cos x_1 \end{bmatrix} \\ y &= x_1 \end{aligned} \tag{2.26}$$

This is a nonlinear time-invariant system of second order.

Example 3 *(An Electric Motor).* A schematic picture of an electric motor is shown in Figure 2.8 Energy stored is stored in the capacitor, and the inductor and momentum is stored in the rotor. Three state variables are needed if we are only interested in motor speed. Storage can be represented by the current I through the rotor, the voltage V across the capacitor and the angular velocity ω of the rotor. The control signal is the voltage E applied to the motor. A momentum balance for the rotor gives

$$J\frac{d\omega}{dt} + D\omega = kI$$

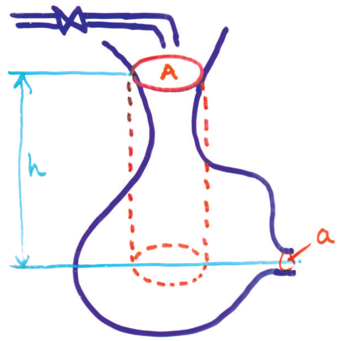

Figure 2.9: A schematic picture of a water tank.

and Kirchoffs laws for the electric circuit gives

$$E = RI + L\frac{dI}{dt} + V - k\frac{d\omega}{dt}$$
$$I = C\frac{dV}{dt}$$

Introducing the state variables $x_1 = \omega$, $x_2 = V$, $x_3 = I$ and the control variable $u = E$ the equations for the motor can be written as

$$\frac{dx}{dt} = \begin{bmatrix} -\frac{D}{J} & 0 & \frac{k}{J} \\ 0 & 0 & \frac{1}{C} \\ -\frac{kD}{JL} & -\frac{1}{L} & \frac{k^2}{JL} - \frac{R}{L} \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} u y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x \qquad (2.27)$$

This is a linear time-invariant system with three state variables and one input.

Example 4 *(The Water Tank).* Consider a tank with water where the input is the inflow and there is free outflow, see Figure 2.9 Assuming that the density is constant a mass balance for the tank gives

$$\frac{dV}{dt} = q_{in} - q_{out}$$

The outflow is given by

$$q_{out} = a\sqrt{2gh}$$

There are several possible choices of state variables. One possibility is to characterize the storage of water by the height of the tank. We have the following relation between height h and volume

$$V = \int_0^h A(x)dx$$

Simplifying the equations we find that the tank can be described by

$$\frac{dh}{dt} = \frac{1}{A(h)}(q_{in} - a\sqrt{2gh})$$
$$q_{out} = a\sqrt{2gh}$$

The tank is thus a nonlinear system of first order.

2.6 Difference Equations

2.7 References