3

Dynamics

3.1 Introduction

From the perspective of control a dynamical system is such that the effects of actions do not occur immediately. Typical examples are: The velocity of a car does not change immediately when the gas pedal is pushed. The temperature in a room does not rise immediately when an air conditioner is switched on. Dynamical systems are also common in daily life. An headache does not vanish immediately when an aspirin is taken. Knowledge of school children do not improve immediately after an increase of a school budget. Training in sports does not immediately improve results. Increased funding for a development project does not increase revenues in the short term.

Dynamics is a key element of control because both processes and controllers are dynamical systems. Concepts, ideas and theories of dynamics are part of the foundation of control theory. Dynamics is also a topic of its own that is closely tied to the development of natural science and mathematics. There has been an amazing development due to contributions from intellectual giants like Newton, Euler, Lagrange and Poincare.

Dynamics is a very rich field that is partly highly technical. In this section we have collected a number of results that are relevant for understanding the basic ideas of control. The chapter is organized in separate sections which can be read independently. For a first time reader we recommend to read this section section-wise as they are needed for the other chapters of the book. To make this possible there is a bit of overlap between the different sections. in connection with the other chapters. There is a bit of overlap so that the different sections can be read independently.

Section 3.2 gives an overview of dynamics and how it is used in control which has inherited ideas both from mechanics and from electrical engineering. It also introduces the standard models that are discussed in the following sections. In Section 3.3 we introduce a model for dynamics in terms of linear, time-invariant differential equations. This material is sufficient for analysis and design of simple control systems of the type discussed in the beginning of Chapter 5. The concepts of transfer function, poles and zeros are also introduced in Section 3.3. Another view of linear, time-invariant systems is given in Section 3.4 introduces the Laplace transform which is a good formalism for linear systems. This also gives another view on transfer functions. Combining the block diagrams introduced in Chapter 2 with the transfer functions gives a simple way to model and analyze feedback systems. The material Section 3.4 gives a good theoretical base for Chapters 4 and 5. Frequency response is yet another useful way of describing dynamics that provides additional insight. The key idea is to investigate how sine waves are propagating through a dynamical system. This is one of the contributions from electrical engineering discussed in Section 3.5. This section together with Section 3.4 gives the basis for reading Chapters 4, 5, 6 and 7 of the book.

Section 3.6 presents the idea of state models which has its origin in Newtonian mechanics. The problems of control have added richness by the necessity to include the effect of external inputs and the information obtained from sensors. In Section 3.6 we also discuss how to obtain models from physics and how nonlinear systems can be approximated by linear systems, so called linearization. In

The main part of this chapter deals with linear time invariant systems. We will frequently only consider systems with one input and one output. This is true for Sections 3.3, 3.4 and 3.5. The state models in Section 3.5 can however be nonlinear and have many inputs and outputs.

3.2 Two Views on Dynamics

Dynamical systems can be viewed from two different ways: the internal view or the external views. 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. 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 state as a function of the state itself, i.e. a differential equation.

$$\frac{dx}{dt} = f(x) \tag{3.1}$$

This is illustrated in Figure 3.1 for a system with two state variables. The particular system represented in the figure is the van der Pol equation

$$\begin{aligned}\frac{dx_1}{dt} &= x_1 - x_1^3 - x_2\\ \frac{dx_2}{dt} &= x_1\end{aligned}$$

which is a model of an electronic oscillator. The model (3.1) 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

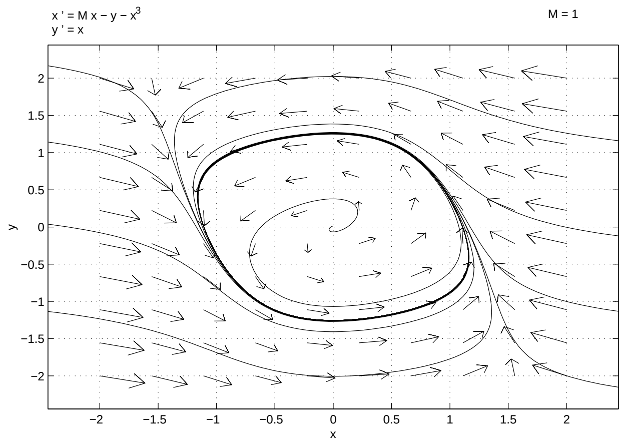

Figure 3.1 Illustration of a state model. A state model gives the rate of change of the state as a function of the state. The velocity of the state are denoted by arrows.

natural system is known ant 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 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 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, see Figure 3.2. 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 $(u_1,\ y_1)$ och $(u_2,\ y_2)$ denote two input-output pairs, and a and b be real numbers. A system is linear if $(au_1+bu_2,\ ay_1+ay_2)$ 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.

Time invariance is another concept. It means that the behavior of the

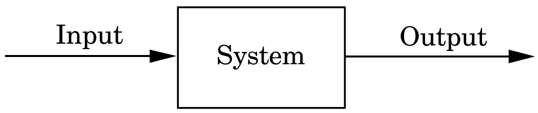

Figure 3.2 Illustration of the input-output view of a dynamical system.

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

$$y(t) = \int_0^t g(t-\tau)u(\tau)d\tau. \tag{3.2}$$

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

$$y(t) = h(t) = \int_0^t g(t-\tau)d\tau = \int_0^t g(\tau)u(\tau)d\tau \tag{3.3}$$

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 (3.4) is thus replaced by

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) \\ y &= g(x,u) \end{aligned} \tag{3.4}$$

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.

Standard Models

Standard models are very useful for structuring our knowledge. It also simplifies problem solving. Learn the standard models, transform the problem to a standard form and you are on familiar grounds. We will discuss four standard forms

- Ordinary differential equations
- Transfer functions
- Frequency responses
- State equations

The first two standard forms are primarily used for linear time-invariant systems. The state equations also apply to nonlinear systems.

3.3 Ordinary Differential Equations

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

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_n y = b_1 \frac{d^{n-1} u}{dt^{n-1}} + b_2 \frac{d^{n-2} u}{dt^{n-2}} + \ldots + b_n u, \tag{3.5}$$

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 (3.5). The form (3.5) adds richness and is much more relevant to control. The equation is sometimes called a controlled differential equation.

The Homogeneous Equation

If the input u to the system (3.5) is zero, we obtain the equation

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = 0, \tag{3.6}$$

which is called the homogeneous equation associated with equation (3.5). The characteristic polynomial of Equations (3.5) and (3.6) is

$$A(s) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n \tag{3.7}$$

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 (3.6).

If the characteristic equation has distinct roots α_k the solution is

$$y(t) = \sum_{k=1}^{n} C_k e^{\alpha_k t}, \tag{3.8}$$

where C_k are arbitrary constants. The Equation (3.6) 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 3.3. Notice that the linear approximations shown in dashed lines change by one unit for one unit of αt. Complex roots $s = \sigma \pm i\omega$ correspond to the time functions.

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

which have oscillatory behavior, see Figure 3.4. The distance between zero crossings is π/ω and corresponding amplitude change is $e^{\sigma\pi/\omega}$.

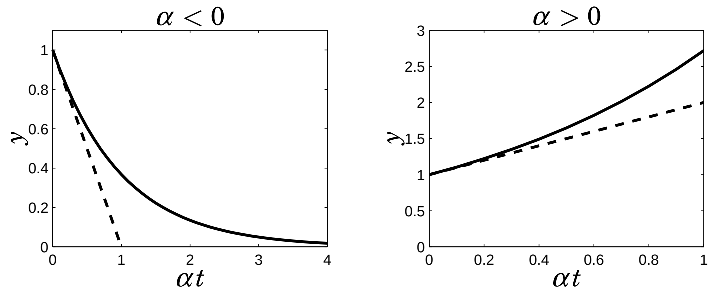

Figure 3.3 The exponential function $y(t) = e^{\alpha t}$. The linear approximations of of the functions for small αt are shown in dashed lines. The parameter $T = 1/\alpha$ is the time constant of the system.

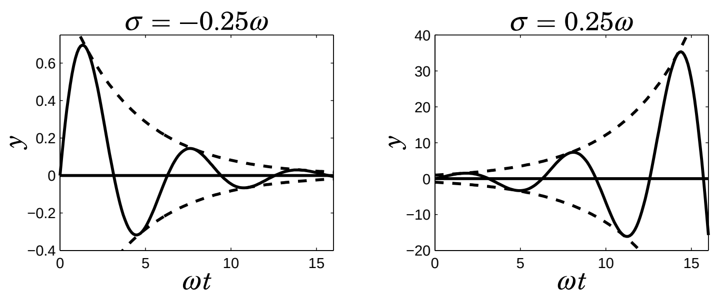

Figure 3.4 The exponential function $y(t) = e^{\sigma t} \sin \omega t$. The linear approximations of of the functions for small αt are shown in dashed lines. The dashed line corresponds to a first order system with time constant $T = 1/\sigma$. The distance between zero crossings is π/ω.

Multiple Roots

When there are multiple roots the solution to Equation (3.6) has the form

$$y(t) = \sum_{k=1}^{n} C_k(t) e^{\alpha_k t}, \tag{3.9}$$

Where $C_k(t)$ is a polynomial with degree less than the multiplicity of the root α_k. The solution (3.9) thus has n free parameters.

The Inhomogeneous Equation – A Special Case

The equation

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = u(t) \tag{3.10}$$

has the solution

$$y(t) = \sum_{k=1}^{n} C_{k-1}(t)e^{\alpha_k t} + \int_0^t h(t-\tau)u(\tau)d\tau, \tag{3.11}$$

where f is the solution to the homogeneous equation

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

with initial conditions

$$f(0) = 0, \quad f'(0) = 0, \ldots, \; f^{(n-2)}(0) = 0, \quad f^{(n-1)}(0) = 1. \tag{3.12}$$

The solution (3.11) 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.

The Inhomogeneous Equation - The General Case

The Equation (3.5) has the solution

$$y(t) = \sum_{k=1}^{n} C_{k-1}(t)e^{\alpha_k t} + \int_0^t g(t-\tau)u(\tau)d\tau, \tag{3.13}$$

where the function g, called the *impulse response*, is given by

$$g(t) = b_1 f^{(n-1)}(t) + b_2 f^{(n-2)}(t) + \; \ldots \; + b_n f(t). \tag{3.14}$$

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 depends on the input.

Notice that the impulse response has the form

$$g(t) = \sum_{k=1}^{n} c_k(t)e^{\alpha_k t}. \tag{3.15}$$

It thus has the same form as the general solution to the homogeneous equation (3.9). The coefficients c_k are given by the conditions (3.12).

The impulse response is also called the weighting function because the second term of (3.13) can be interpreted as a weighted sum of past inputs.

The Step Response

Consider (3.13) and assume that all initial conditions are zero. The output is then given by

$$y(t) = \int_0^t g(t-\tau)u(\tau)d\tau, \tag{3.16}$$

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

$$y(t) = \int_0^t g(t-\tau)d\tau = \int_0^t g(\tau)d\tau = H(t), \tag{3.17}$$

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

$$g(t) = \frac{dh(t)}{dt} \tag{3.18}$$

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.

Stability

The solution of system is described by the ordinary differential equation (3.5) is given by (3.9). The solution is stable if all solutions go to zero. A system is thus stable if the real parts of all α_i are negative, or equivalently that all the roots of the characteristic polynomial (3.7) have negative real parts.

Stability can be determined simply by finding the roots of the characteristic polynomial of a system. This is easily done in Matlab.

The Routh-Hurwitz Stability Criterion

When control started to be used for steam engines and electric generators computational tools were not available and it was it was a major effort to find roots of an algebraic equation. Much intellectual activity was devoted to the problem of investigating if an algebraic equation have all its roots in the left half plane without solving the equation resulting in the Routh-Hurwitz criterion. Some simple special cases of this criterion are given below.

- The polynomial $A(s) = s + a_1$ has its zero in the left half plane if $a_1 > 0$.
- The polynomial $A(s) = s^2 + a_1 s + a_2$ has all its zeros in the left half plane if all coefficients are positive.

- The polynomial $A(s) = s^3 + a_2 s^2 + a_3$ has all its zeros in the left half plane if all coefficients are positive and if $a_1 a_2 > a - 3$.

Transfer Functions, Poles and Zeros

The model (3.5) is characterized by two polynomials

$$\begin{aligned} A(s) &= s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_{n-1} s + a_n \\ B(s) &= b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_{n-1} s + b_n \end{aligned}$$

The rational function

$$G(s) = \frac{B(s)}{A(s)} \tag{3.19}$$

is called the transfer function of the system.

Consider a system described by (3.5) assume that the input and the output have constant values u_0 and y_0 respectively. It then follows from (3.5) that

$$a_n y_0 = b_n u_0$$

which implies that

$$\frac{y_0}{u_0} = \frac{b_n}{a_n} = G(0)$$

The number $G(0)$ is called the static gain of the system because it tells the ratio of the output and the input under steady state condition. If the input is constant $u = u_0$ and the system is stable then the output will reach the steady state value $y_0 = G(0)u_0$. The transfer function can thus be viewed as a generalization of the concept of gain.

Notice the symmetry between y and u. The *inverse system* is obtained by reversing the roles of input and output. The transfer function of the system is $\dfrac{B(s)}{A(s)}$ and the inverse system has the transfer function $\dfrac{A(s)}{B(s)}$.

The roots of $A(s)$ are called poles of the system. The roots of $B(s)$ are called zeros of the system. The poles of the system are the roots of the characteristic equation, they characterize the general solution to to the homogeneous equation and the impulse response. A pole $s = \lambda$ corresponds to the component $e^{\lambda t}$ of the solution, also called a mode. If $A(\alpha) = 0$, then $y(t) = e^{\alpha t}$ is a solution to the homogeneous equation (3.6). Differentiation gives

$$\frac{d^k y}{dt^k} = \alpha^k y(t)$$

and we find

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = A(\alpha) y(t) = 0$$

The modes thus correspond to the terms of the solution to the homogeneous equation (3.6) and the terms of the impulse response (3.15) and the step response.

If $s = \beta$ is a zero of $B(s)$ and $u(t) = Ce^{\beta t}$, then it follows that

$$b_1 \frac{d^{n-1}u}{dt^{n-1}} + b_2 \frac{d^{n-2}u}{dt^{n-2}} \ldots + b_n u = B(\beta)Ce^{\beta t} = 0.$$

A zero of $B(s)$ at $s = \beta$ blocks the transmission of the signal $u(t) = Ce^{\beta t}$.

3.4 Laplace Transforms

The Laplace transform is very convenient for dealing with linear time-invariant system. The reason is that it simplifies manipulations of linear systems to pure algebra. It also a natural way to introduce transfer functions and it also opens the road for using the powerful tools of the theory of complex variables. The Laplace transform is an essential element of the language of control.

The Laplace Transform

Consider a function f defined on $0 \le t < \infty$ and a real number $\sigma > 0$. Assume that f grows slower than $e^{\sigma t}$ for large t. The Laplace transform $F = \mathcal{L}f$ of f is defined as

$$\mathcal{L}f = F(s) = \int_0^\infty e^{-st} f(t) dt$$

We will illustrate computation of Laplace transforms with a few examples

Transforms of Simple Function The transform of the function $f_1(t) = e^{-at}$ is given by

$$F_1(s) = \int_0^\infty e^{-(s+a)t} dt = -\frac{1}{s+a} e^{-st} \Big|_0^\infty = \frac{1}{s+a}$$

Differentiating the above equation we find that the transform of the function $f_2(t) = te - at$ is

$$F_2(s) = \frac{1}{(s+a)^2}$$

Repeated differentiation shows that the transform of the function $f_3(t) = t^n e^{-at}$ is

$$F_3(s) = \frac{(n-1)!}{(s+a)^n}$$

Setting $a = 0$ in f_1 we find that the transform of the unit step function $f_4(t) = 1$ is

$$F_4(s) = \frac{1}{s}$$

Similarly we find by setting $a = 0$ in f_3 that the transform of $f_5 = t^n$ is

$$F_5(s) = \frac{n!}{s^{n+1}}$$

Setting $a = ib$ in f_1 we find that the transform of $f(t) = e^{-ibt} = \cos bt - i \sin bt$ is

$$F(s) = \frac{1}{s+ib} = \frac{s-ib}{s^2+b^2} = \frac{s}{s^2+b^2} - i\frac{b}{s^2+b^2}$$

Separating real and imaginary parts we find that the transform of $f_6(t) = \sin bt$ and $f_7(t) = \cos bt$ are

$$F_6(t) = \frac{b}{s^2+b^2}, \quad F_7(t) = \frac{s}{s^2+b^2}$$

Proceeding in this way it is possible to build up tables of transforms that are useful for hand calculations.

Properties of Laplace Transforms The Laplace transform also has many useful properties. First we observe that the transform is linear because

$$\begin{aligned}\mathcal{L}(af + bg) &= aF(s) + bF(s) = a\int_0^\infty e^{-st} f(t)dt + b\int_0^\infty e^{-st} g(t)dt\\ &= \int_0^\infty e^{-st}(af(t) + bg(t))dt = a\mathcal{L}f + b\mathcal{L}g\end{aligned}$$

Next we will calculate the transform of the derivative of a function, i.e. $f'(t) = \frac{df(t)}{dt}$. We have

$$\mathcal{L}\frac{df}{dt} = \int_0^\infty e^{-st} f'(t)dt = e^{-st}f(t)\Big|_0^\infty + s\int_0^\infty e^{-st} f(t)dt = -f(0) + s\mathcal{L}f$$

where the second equality is obtained by integration by parts. This formula is very useful because it implies that differentiation of a time function corresponds to multiplication of the transform by s provided that the initial value $f(0)$ is zero. We will consider the transform of an integral

$$\begin{aligned}
\mathcal{L}\int_0^t f(\tau)d\tau &= \int_0^\infty e^{-st}\int_0^t f(\tau)d\tau \\
&= -\frac{e^{-st}}{s}\int_0^t e^{-s\tau}f'(\tau)d\tau\Big|_0^\infty + \int_0^\infty \frac{e^{-s\tau}}{s}f(\tau)d\tau \\
&= \frac{1}{s}\int_0^\infty e^{-s\tau}f(\tau)d\tau = \mathcal{L}f
\end{aligned}$$

The relation between the input u and the output y of a linear time-invariant system is given by the convolution integral

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

see (3.18). We will now consider the Laplace transform of such an expression. We have

$$\begin{aligned}
Y(s) &= \int_0^\infty e^{-st}y(t)dt = \int_0^\infty e^{-st}\int_0^\infty g(t-\tau)u(\tau)d\tau dt \\
&= \int_0^\infty\int_0^t e^{-s(t-\tau)}e^{-s\tau}g(t-\tau)u(\tau)d\tau dt \\
&= \int_0^\infty e^{-s\tau}u(\tau)d\tau\int_0^\infty e^{-st}g(t)dt = G(s)U(s)
\end{aligned}$$

The description of a linear time-invariant systems thus becomes very simple when working with Laplace transforms.

Next we will consider the effect of a time shift. Let the number a be positive and let the function f_a be a time shift of the function f, i.e.

$$f_a(t) = \begin{cases} 0 & \text{for PI } t < 0 \\ f(t-a) & \text{for } t \geq 0 \end{cases}$$

The Laplace transform of f_a is given by

$$\begin{aligned}
F_a(s) &= \int_0^\infty e^{-st}f(t-a)dt = \int_a^\infty e^{-st}f(t-a)dt \\
&= \int_a^\infty e^{-as}e^{-s(t-a)}f(t-a)dt = e^{-as}\int_0^\infty e^{-st}f(t)dt = e^{-as}F(s)
\end{aligned} \tag{3.20}$$

Delaying a signal by a time units thus correspond to multiplication of its Laplace transform by e^{-as}.

The behavior of time functions for small arguments is governed by the behavior of the Laplace transform for large arguments. This is expressed by the so called initial value theorem.

$$\lim_{s\to\infty} sF(s) = \lim_{s\to\infty} \int_0^\infty se^{-st} f(t)dt = \lim_{s\to\infty} \int_0^\infty e^{-v} f(\frac{v}{s})dv = f(0)$$

This holds provided that the limit exists.

The converse is also true which means that the behavior of time functions for large arguments is governed by the behavior of the Laplace transform for small arguments. Final value theorem. Hence

$$\lim_{s\to 0} sF(s) = \lim_{s\to 0} \int_0^\infty se^{-st} f(t)dt = \lim_{s\to 0} \int_0^\infty e^{-v} f(\frac{v}{s})dv = f(\infty)$$

These properties are very useful for qualitative assessment of a time functions and Laplace transforms.

Linear Differential Equations

The differentiation property $\mathcal{L}\frac{df}{dt} = s\mathcal{L}f - f(0)$ makes the Laplace transform very convenient for dealing with linear differential equations. Consider for example the system

$$\frac{dy}{dt} = ay + bu$$

Taking Laplace transforms of both sides give

$$sY(s) - y(0) = aY(s) + bU(s)$$

Solving this linear equation for $Y(s)$ gives

$$Y(s) = \frac{y(0)}{s-a} + \frac{b}{s-a}U(s)$$

Transforming back to time function we find

$$y(t) = e^{at}y(0) + b\int_0^t e^{a(t-\tau)}u(\tau]d\tau$$

To convert the transforms to time functions we have used the fact that the transform

$$\frac{1}{s-a}$$

corresponds to the time function e^{at} and we have also used the rule for transforms of convolution.

Inverse Transforms

A simple way to find time functions corresponding to a rational Laplace transform. Write $F(s)$ in a partial fraction expansion

$$F(s) = \frac{B(s)}{A(s)} = \frac{B(s)}{(s-\alpha_1)(s-\alpha_2)\dots(s-\alpha_n)} = \frac{C_1}{s-\alpha_1} + \frac{C_2}{s-\alpha_2} + \dots + \frac{C_n}{s-\alpha_n}$$
$$C_k = \lim_{s\to\alpha_k}(s-\alpha_k)F(s) = \frac{B(\alpha_k)}{(\alpha_k-\alpha_1)\dots(\alpha_k-\alpha_{k-1})(s-\alpha_{k+1})\dots(\alpha_k-\alpha_n)}$$

The time function corresponding to the transform is

$$f(t) = C_1 e^{\alpha_1 t} + C_2 e^{\alpha_2 t} + \dots + C_n e^{\alpha_n t}$$

Parameters α_k give shape and numbers C_k give magnitudes.

Notice that α_k may be complex numbers. With multiple roots the constants C_k are instead polynomials.

The Transfer Function

The transfer function of an LTI system was introduced in Section 3.3 when dealing with differential equations. Using Laplace transforms it can also be defined as follows. Consider an LTI system with input u and output y. The transfer function is the ratio of the transform of the output and the input where *the Laplace transforms are calculated under the assumption that all initial values are zero.*

$$G(s) = \frac{Y(s)}{U(s)} = \frac{\mathcal{L}y}{\mathcal{L}u}$$

The fact that all initial values are assumed to be zero has some consequences that will be discussed later.

EXAMPLE 3.1—LINEAR TIME-INVARIANT SYSTEMS

Consider a system described by the ordinary differential equation (3.5), i.e

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

Taking Laplace transforms under the assumption that all initial values are zero we get.

$$\begin{aligned}&(s^n + a_1 s^{n-1} + a_2 s^{n-2} + \dots + a_{n-1}s + a_n)Y(s)\\ &\quad = (b_1 s^{n-1} + b_2 s^{n-2} + \dots + b_{n-1}s + b_n)U(s)\end{aligned}$$

The transfer function of the system is thus given by

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_{n-1}s + b_n}{s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_{n-1}s + a_n} = \frac{B(s)}{A(s)} \tag{3.21}$$

□

EXAMPLE 3.2—A TIME DELAY

Consider a system which simply delays the input T time units. It follows from 3.20 that the input output relation is

$$Y(s) = e^{-sT}U(s)$$

The transfer function of a time delay is thus

$$G(s) = \frac{Y(s)}{U(s)} = e^{-sT}$$

□

It is also possible to calculate the transfer functions for systems described by partial differential equations.

EXAMPLE 3.3—THE HEAT EQUATION

$$G(s) = e^{-\sqrt{sT}}$$
$$G(s) = \frac{1}{\cosh\sqrt{sT}}$$

□

Transfer functions and Laplace transforms are ideal to deal with block diagrams for linear time-invariant systems. We have already shown that a block is simply characterized by

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

The transform of the output of a block is simply the product of the transfer function of the block and the transform of the input system. Algebraically this is equivalent to multiplication with a constant. This makes it easy to find relations between the signals that appear in a block diagram. The combination of block diagrams and transfer functions is a very nice combination because they make it possible both to obtain an overview of a system and to guide the derivation of equations for the system. This is one of the reasons why block diagrams are so widely used in control.

Notice that it also follows from the above equation that signals and systems have the same representations. In the formula we can thus consider g as the input and u as the transfer function.

To illustrate the idea we will consider an example.

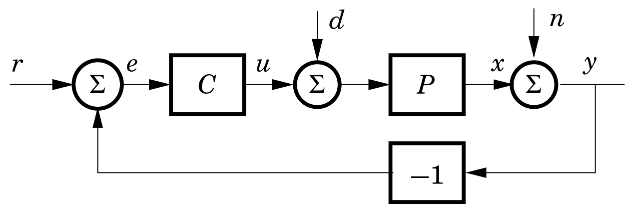

Figure 3.5 Block diagram of a feedback system.

EXAMPLE 3.4—RELATIONS BETWEEN SIGNALS IN A BLOCK DIAGRAM
Consider the system in Figure 3.5. The system has two blocks representing the process P and the controller C. There are three external signals, the reference r, the load disturbance d and the measurement noise n. A typical problem is to find out how the error e related to the signals r d and n? Introduce Laplace transforms and transfer functions. To obtain the desired relation we simply trace the signals around the loop. Starting with the signal e and tracing backwards in the loop we find that e is the difference between r and y, hence $E = R - Y$. The signal y in turn is the sum of n and the output of the block P, hence $Y = N + P(D + V)$. Finally the signal v is the output of the controller which is given by $V = PE$. Combining the results we get

$$E = R - \big(N + P(D + CE)\big)$$

With a little practice this equation can be written directly. Solving for E gives

$$E = \frac{1}{1 + PC}R - \frac{1}{1 + PC}N - \frac{P}{1 + PC}D$$

□

Notice *the form of the equations* and the use of *superposition*.

Simulating LTI Systems

Linear time-invariant systems can be conveniently simulated using Matlab. For example a system with the transfer function

$$G(s) = \frac{5s + 2}{s^2 + 3s + 2}$$

is introduced in matlab as

```
G=tf([5 2],[1 3 2])
```

The command `step(G)` gives the step response of the system.

Transfer Functions

The transfer function of a linear system is defined as

$$G(s) = \frac{Y(s)}{U(s)} = \frac{\mathcal{L}y}{\mathcal{L}u} \tag{3.22}$$

where $U(s) = \mathcal{L}u$ is the Laplace transform of the input u and $Y(s) = \mathcal{L}y$ is the Laplace transform of the output y. The Laplace transforms are computed under the assumption that all initial conditions are zero.

Circuit Analysis

Laplace transforms are very useful for circuit analysis. A resistor is described by the algebraic equation

$$V = RI$$

but inductors and capacitors are describe by the linear differential equations

$$\begin{aligned} CV &= \int_0^t I(\tau]d\tau \\ L\frac{dI}{dt} &= V \end{aligned}$$

Taking Laplace transforms we get

$$\begin{aligned} \mathcal{L}V &= RI \\ \mathcal{L}V &= \frac{1}{sC}\mathcal{L}I \\ \mathcal{L}V &= sL\mathcal{L}I \end{aligned}$$

The transformed equations for all components thus look identical, the transformed voltage $\mathcal{L}V$ is a generalized impedance Z multiplied by the transformed current $\mathcal{L}I$. The impedance is

$$\begin{aligned} Z(s) &= R \text{ for a resistor} \\ Z(s) &= \frac{1}{sC} \text{ for a capacitor} \\ Z(s) &= sL \text{ for an inductor} \end{aligned}$$

Operating with the transforms we can thus pretend that all elements of a circuit is a resistor which means that circuit analysis is reduced to pure algebra. This is just another illustration of the fact that differential equations are transformed to algebraic equations. We illustrate the procedure by an example.

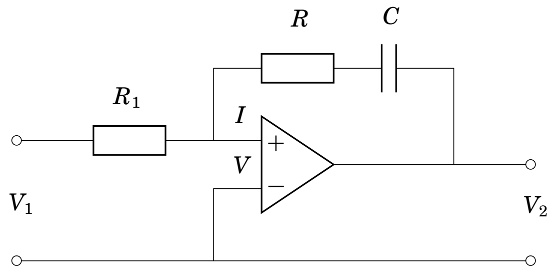

Figure 3.6 Schematic diagram of an electric circuit.

EXAMPLE 3.5—OPERATIONAL AMPLIFIERS
Consider the electric circuit shown in Figure 3.6. Assume that the problem is to find the relation between the input voltage V_1 and the output voltage V_2. Assuming that the gain of the amplifier is very high, say around 10^6, then the voltage V is negligible and the current I_0 is zero. The currents I_1 and I_2 then are the same which gives

$$\frac{\mathcal{L}V_1}{Z_1(s)} = -\frac{\mathcal{L}V_2}{Z_2(s)}$$

It now remains to determine the generalized impedances Z_1 and Z_2. The impedance Z_2 is a regular resistor. To determine Z_1 we use the simple rule for combining resistors which gives

$$Z_1(s) = R + \frac{1}{sC}$$

Hence

$$\frac{\mathcal{L}V_2}{\mathcal{L}V_2} = -\frac{Z_1(s)}{Z_2(s)} = -\frac{R}{R_2} - \frac{1}{R_2Cs}$$

Converting to the time domain we find

$$V_2(t) = -\frac{R}{R_2} - \frac{1}{R_2C}\int_0^t V_1(\tau)d\tau$$

The circuit is thus a PI controller. □

3.5 Frequency Response

The idea of frequency response is to characterize linear time-invariant systems by their response to sinusoidal signals. The idea goes back to

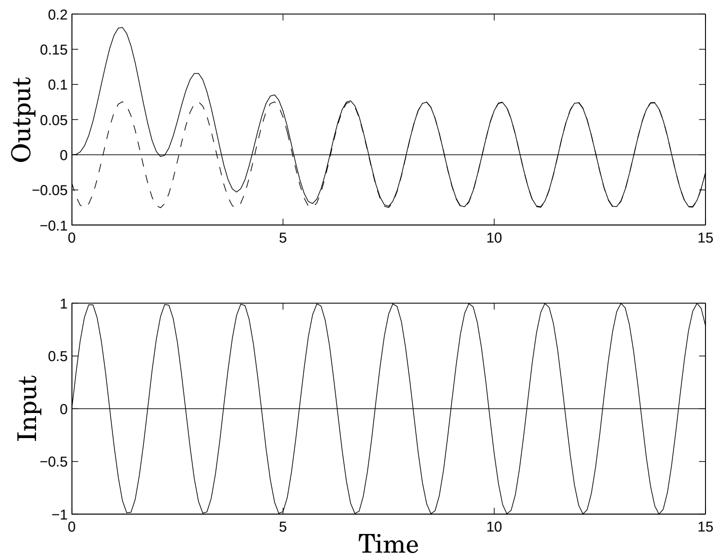

Figure 3.7 Response of a linear time-invariant system to a sinusoidal input (full lines). The dashed line shows the steady state output calculated from (3.23).

Fourier, who introduced the method to investigate propagation of heat in metals. Figure 3.7 shows the response of a linear time-invariant system to a sinusoidal input. The figure indicates that after a transient the output is a sinusoid with the same frequency as the input. The steady state response to a sinusoidal input of a stable linear system is in fact given by $G(i\omega)$. Hence if the input is

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

the output is

$$y(t) = a|G(i\omega)| \sin (\omega t + \arg G(i\omega)) = a \Im e^{i\omega t} G(i\omega) \tag{3.23}$$

The dashed line in Figure 3.7 shows the output calculated by this formula. It follows from this equation that the transfer function G has the interesting property that its value for $s = i\omega$ describes the steady state response to sinusoidal signals. The function $G(i\omega)$ is therefore called the frequency response. The argument of the function is frequency ω and the function takes complex values. The magnitude gives the magnitude of the steady state output for a unit amplitude sinusoidal input and the argument gives the phase shift between the input and the output. Notice that the system must be stable for the steady state output to exist.

The frequency response can be determined experimentally by analyzing how a system responds to sinusoidal signals. It is possible to make very accurate measurements by using correlation techniques.

To derive the formula we will first calculate the response of a system with the transfer function $G(s)$ to the signal e^{at} where all poles of the transfer function have the property $\Re p_k < \alpha < a$. The Laplace transform of the output is

$$Y(s) = G(s)\frac{1}{s-a}$$

Making a partial fraction expansion we get

$$Y(s) = \frac{G(a)}{s-a} + \sum \frac{R_k}{(s-p_k)(p_k-a)}$$

It follows the output has the property

$$\left|y(t) - G(a)e^{at}\right| < ce^{-\alpha t}$$

where c is a constant. Asymptotically we thus find that the output approaches $G(a)e^{at}$. Setting $a = ib$ we find that the response to the input

$$u(t) = e^{ibt} = \cos bt + i \sin bt$$

will approach

$$\begin{aligned} y(t) &= G(ib)e^{ibt} = |G(ib)|e^{i(b+\arg G(ib))} \\ &= |G(ib)|\cos\left(bt + \arg\arg G(ib)\right) + i|G(ib)|\sin\left(bt + \arg\arg G(ib)\right) \end{aligned}$$

Separation of real and imaginary parts give the result.

Nyquist Plots

The response of a system to sinusoids is given by the the frequency response $G(i\omega)$. This function can be represented graphically by plotting the magnitude and phase of $G(i\omega)$ for all frequencies, see Figure 3.8. The magnitude $a = |G(i\omega)|$ represents the amplitude of the output and the angle $\phi = \arg G(i\omega)$ represents the phase shift. The phase shift is typically negative which implies that the output will lag the input. The angle ψ in the figure is therefore called phase lag. One reason why the Nyquist curve is important is that it gives a totally new way of looking at stability of a feedback system. Consider the feedback system in Figure 3.9. 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

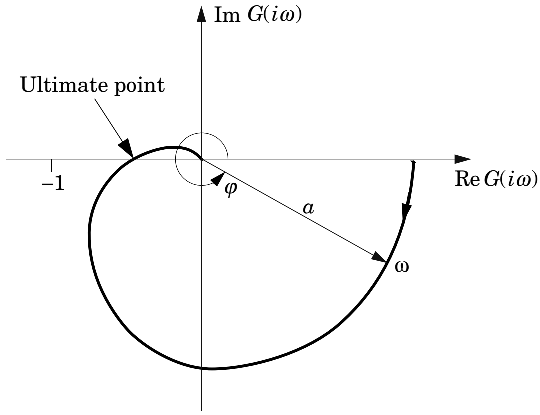

Figure 3.8 The Nyquist plot of a transfer function $G(i\omega)$.

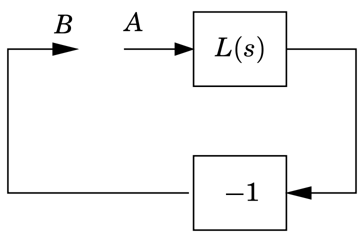

Figure 3.9 Block diagram of a simple feedback system.

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. We have also defined stability as a binary property, a system is either stable or unstable. In practice it is useful to be able to talk about degrees of stability. All of these issues are addressed by Nyquist's stability criterion. This result has a strong intuitive component which we will discuss first. There is also some beautiful mathematics associated with it that will be discussed in a separate section.

Consider the feedback system in Figure 3.9. Let the transfer functions of the process and the controller be $P(s)$ and $C(s)$ respectively. Introduce the loop transfer function

$$L(s) = P(s)C(s) \tag{3.24}$$

To get insight into the problem of stability we will start by investigating the conditions for oscillation. For that purpose we cut the feedback loop as indicated in the figure and we inject a sinusoid at point A. In steady state the signal at point B will also be a sinusoid with the same frequency. It seems reasonable that an oscillation can be maintained if the signal at B has the same amplitude and phase as the injected signal because we could then connect A to B. Tracing signals around the loop we find that the condition that the signal at B is identical to the signal at A is that

$$L(i\omega_0) = -1 \tag{3.25}$$

which we call the condition for oscillation. This condition means that the Nyquist curve of $L(i\omega)$ intersects the negative real axis at the point -1. Intuitively it seems reasonable that the system would be stable if the Nyquist curve intersects to the right of the point -1 as indicated in Figure 3.9. This is essentially true, but there are several subtleties that are revealed by the proper theory.

Stability Margins

In practice it is not enough to require that the system is stable. There must also be some margins of stability. There are many ways to express this. Many of the criteria are based on Nyquist's stability criterion. They are based on the fact that it is easy to see the effects of changes of the gain and the phase of the controller in the Nyquist diagram of the loop transfer function $L(s)$. An increase of controller gain simply expands the Nyquist curve radially. An increase of the phase of the controller twists the Nyquist curve clockwise, see Figure 3.10. The gain margin g_m tells how much the controller gain can be increased before reaching the stability limit. Let ω_{180} be the smallest frequency where the phase lag of the loop transfer function $L(s)$ is 180°. The gain margin is defined as

$$g_m = \frac{1}{|L(i\omega_{180})|} \tag{3.26}$$

The stability margin is a closely related concept which is defined as

$$s_m = 1 + |L(i\omega_{180})| = 1 - \frac{1}{g_m} \tag{3.27}$$

A nice feature of the stability margin is that it is a number between 0 and 1. Values close to zero imply a small margin.

The phase margin φ_m is the amount of phase lag required to reach the stability limit. Let ω_{gc} denote the lowest frequency where the loop transfer function $L(s)$ has unit magnitude. The phase margin is then given by

$$\varphi_m = \pi + \arg L(i\omega_{gc}) \tag{3.28}$$

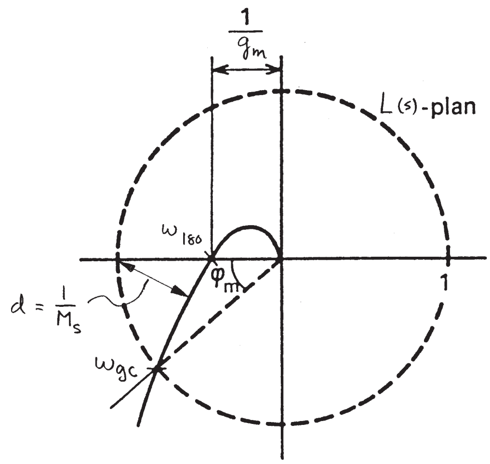

Figure 3.10 Nyquist curve of the loop transfer function L with indication of gain, phase and stability margins.

The margins have simple geometric interpretations in the Nyquist diagram of the loop transfer function as is shown in Figure 3.10. The stability margin s_m is the distance between the critical point and the intersection of the Nyquist curve with the negative real axis.

One possibility to characterize the stability margin with a single number is to choose the shortest distance d to the critical point. This is also shown in Figure 3.10.

Reasonable values of the margins are phase margin $\varphi_m = 30° - 60°$, gain margin $g_m = 2 - 5$, stability margin $s_m = 0.5 - 0.8$, and shortest distance to the critical point $d = 0.5 - 0.8$.

The gain and phase margins were originally conceived for the case when the Nyquist curve only intersects the unit circle and the negative real axis once. For more complicated systems there may be many intersections and it is then necessary to consider the intersections that are closest to the critical point. For more complicated systems there is also another number that is highly relevant namely the delay margin. The delay margin is defined as the smallest time delay required to make the system unstable. For loop transfer functions that decay quickly the delay margin is closely related to the phase margin but for systems where the amplitude ratio of the loop transfer function has several peaks at high frequencies the delay margin is a much more relevant measure.

Nyquist's Stability Theorem*

We will now prove the Nyquist stability theorem. This will require more results from the theory of complex variables than in many other parts of the book. Since precision is needed we will also use a more mathematical style of presentation. We will start by proving a key theorem about functions of complex variables.

THEOREM 3.1—PRINCIPLE OF VARIATION OF THE ARGUMENT
Let D be a closed region in the complex plane and let Γ be the boundary of the region. Assume the function f is analytic in D and on Γ except at a finite number of poles and zeros, then

$$w_n = \frac{1}{2\pi}\Delta_\Gamma \arg f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz = N - P$$

where N is the number of zeros and P the number of poles in D. Poles and zeros of multiplicity m are counted m times. The number w_n is called the winding number and $\Delta_\Gamma \arg f(z)$ is the variation of the argument of the function f as the curve Γ is traversed in the positive direction.

PROOF 3.1
Assume that $z = a$ is a zero of multiplicity m. In the neighborhood of $z = a$ we have

$$f(z) = (z-a)^m g(z)$$

where the function g is analytic and different form zero. We have

$$\frac{f'(z)}{f(z)} = \frac{m}{z-a} + \frac{g'(z)}{g(z)}$$

The second term is analytic at $z = a$. The function f'/f thus has a single pole at $z = a$ with the residue m. The sum of the residues at the zeros of the function is N. Similarly we find that the sum of the residues of the poles of is $-P$. Furthermore we have

$$\frac{d}{dz}\log f(z) = \frac{f'(z)}{f(z)}$$

which implies that

$$\int_\Gamma \frac{f'(z)}{f(z)}dz = \Delta_\Gamma \log f(z)$$

where Δ_Γ denotes the variation along the contour Γ. We have

$$\log f(z) = \log |f(z)| + i \arg f(z)$$

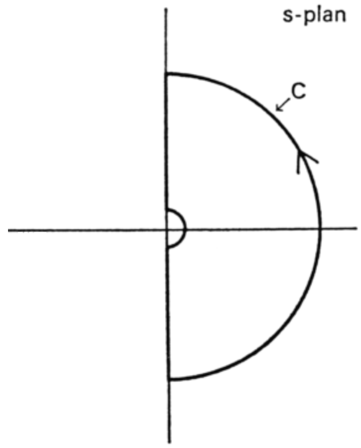

Figure 3.11 Contour Γ used to prove Nyquist's stability theorem.

Since the variation of $|f(z)|$ around a closed contour is zero we have

$$\Delta_\Gamma \log f(z) = i\Delta_\Gamma \arg f(z)$$

and the theorem is proven.

REMARK 3.1
The number w_n is called the winding number.

REMARK 3.2
The theorem is useful to determine the number of poles and zeros of an function of complex variables in a given region. To use the result we must determine the winding number. One way to do this is to investigate how the curve Γ is transformed under the map f. The variation of the argument is the number of times the map of Γ winds around the origin in the f-plane. This explains why the variation of the argument is also called the winding number. □

We will now use the Theorem 1 to prove Nyquist's stability theorem. For that purpose we introduce a contour that encloses the right half plane. For that purpose we choose the contour shown in Figure 3.11. The contour consists of a small half circle to the right of the origin, the imaginary axis and a large half circle to the right with with the imaginary axis as a diameter. To illustrate the contour we have shown it drawn with a small radius r and a large radius R. The Nyquist curve is normally the map of the positive imaginary axis. We call the contour Γ the full Nyquist contour.

Consider a closed loop system with the loop transfer function $L(s)$. The closed loop poles are the zeros of the function

$$f(s) = 1 + L(s)$$

To find the number of zeros in the right half plane we thus have to investigate the winding number of the function $f = 1 + L$ as s moves along the contour Γ. The winding number can conveniently be determined from the Nyquist plot. A direct application of the Theorem 1 gives.

THEOREM 3.2—NYQUIST'S STABILITY THEOREM
Consider a simple closed loop system with the loop transfer function $L(s)$. Assume that the loop transfer function does not have any poles in the region enclosed by Γ and that the winding number of the function $1+L(s)$ is zero. Then the closed loop characteristic equation has not zeros in the right half plane.

We illustrate Nyquist's theorem by an examples.

EXAMPLE 3.6—A SIMPLE CASE
Consider a closed loop system with the loop transfer function

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

Figure 3.12 shows the image of the contour Γ under the map L. The Nyquist curve intersects the imaginary axis for $\omega = 1$ the intersection is at $-k/2$. It follows from Figure 3.12 that the winding number is zero if $k < 2$ and 2 if $k > 2$. We can thus conclude that the closed loop system is stable if $k < 2$ and that the closed loop system has two roots in the right half plane if $k > 2$. □

By using Nyquist's theorem it was possible to resolve a problem that had puzzled the engineers working with feedback amplifiers. The following quote by Nyquist gives an interesting perspective.

> Mr. Black proposed a negative feedback repeater and proved by tests that it possessed the advantages which he had predicted for it. In particular, its gain was constant to a high degree, and it was linear enough so that spurious signals caused by the interaction of the various channels could be kept within permissible limits. For best results, the feedback factor, the quantity usually known as $\mu\beta$ (the loop transfer function), had to be numerically much larger than unity. The possibility of stability with a feedback factor greater than unity was

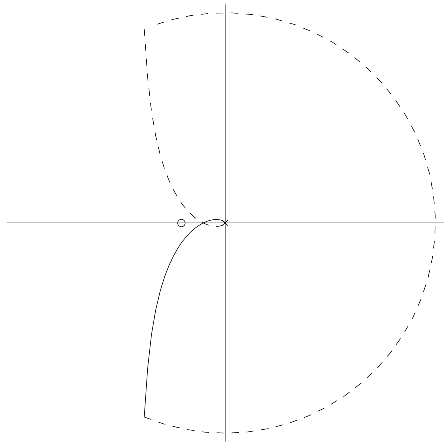

Figure 3.12 Map of the contour Γ under the map $L(s) = \frac{k}{s((s+1)^2}$. The curve is drawn for $k < 2$. The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

puzzling. Granted that the factor is negative it was not obvious how that would help. If the factor was -10, the effect of one round trip around the feedback loop is to change the magnitude of the current from, say 1 to -10. After a second trip around the loop the current becomes 100, and so forth. The totality looks much like a divergent series and it was not clear how such a succession of ever-increasing components could add to something finite and so stable as experience had shown. The missing part in this argument is that the numbers that describe the successive components 1, -10, 100, and so on, represent the steady state, whereas at any finite time many of the components have not yet reached steady state and some of them, which are destined to become very large, have barely reached perceptible magnitude. My calculations were principally concerned with replacing the indefinite diverging series referred to by a series which gives the actual value attained at a specific time t. The series thus obtained is convergent instead of divergent and, moreover, converges to values in agreement with the experimental findings.

This explains how I came to undertake the work. It should perhaps be explained also how it come to be so detailed. In the course of the calculations, the facts with which the term conditional stability have come to be associated, became apparent. One aspect of this was that it is possible to have a feedback loop which is stable and can be made unstable by by increasing the loop loss. this seemed a very surprising

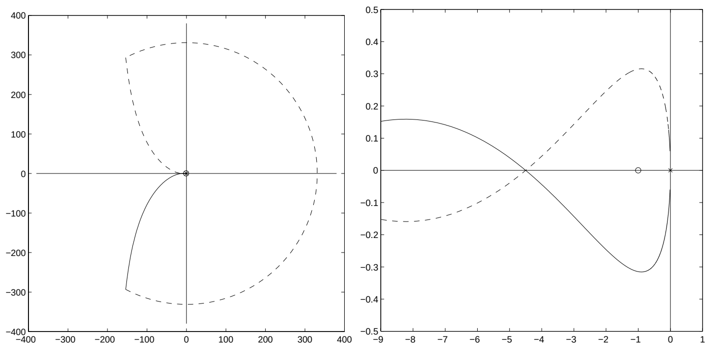

Figure 3.13 Map of the contour Γ under the map $L(s) = \frac{3(s+1)^2}{s(s+6)^2}$, see (3.29), which is a conditionally stable system. The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semicircle at the origin in dashed lines. The plot on the right is an enlargement of the area around the origin of the plot on the left.

> result and appeared to require that all the steps be examined and set forth in full detail.

This quote clearly illustrate the difficulty in understanding feedback by simple qualitative reasoning. We will illustrate the issue of conditional stability by an example.

EXAMPLE 3.7—CONDITIONAL STABILITY
Consider a feedback system with the loop transfer function

$$L(s) = \frac{3(s+1)^2}{s(s+6)^2} \tag{3.29}$$

The Nyquist plot of the loop transfer function is shown in Figure 3.13 The figure shows that the Nyquist curve intersects the negative real axis at a point close to -5. The naive argument would then indicate that the system would be unstable. The winding number is however zero and stability follows from Nyquist's theorem. □

Notice that Nyquist's theorem does not hold if the loop transfer function has a pole in the right half plane. There are extensions of the Nyquist theorem to cover this case but it is simpler to invoke Theorem 1 directly. We illustrate this by two examples.

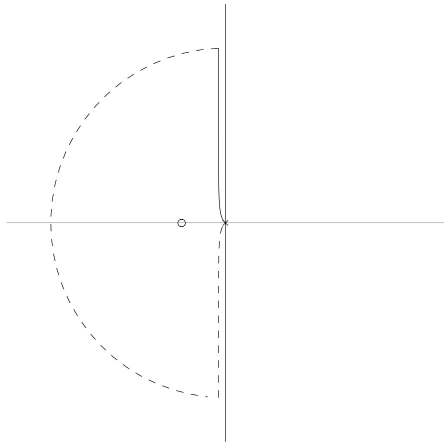

Figure 3.14 Map of the contour Γ under the map $L(s) = \frac{k}{s(s-1)(s+5)}$. The curve on the right shows the region around the origin in larger scale. The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

EXAMPLE 3.8—LOOP TRANSFER FUNCTION WITH RHP POLE

Consider a feedback system with the loop transfer function

$$L(s) = \frac{k}{s(s-1)(s+5)}$$

This transfer function has a pole at $s = 1$ in the right half plane. This violates one of the assumptions for Nyquist's theorem to be valid. The Nyquist curve of the loop transfer function is shown in Figure 3.14. Traversing the contour Γ in clockwise we find that the winding number is 1. Applying Theorem 1 we find that

$$N - P = 1$$

Since the loop transfer function has a pole in the right half plane we have $P = 1$ and we get $N = 2$. The characteristic equation thus has two roots in the right half plane. □

EXAMPLE 3.9—THE INVERTED PENDULUM

Consider a closed loop system for stabilization of an inverted pendulum with a PD controller. The loop transfer function is

$$L(s) = \frac{s+2}{s^2-1} \tag{3.30}$$

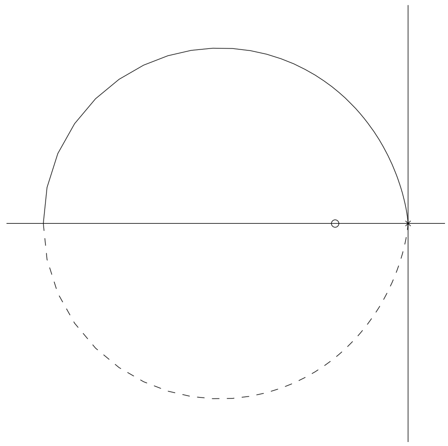

Figure 3.15 Map of the contour Γ under the map $L(s) = \frac{s+2}{s^2-1}$ given by (3.30). The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

This transfer function has one pole at $s = 1$ in the right half plane. The Nyquist curve of the loop transfer function is shown in Figure 3.15. Traversing the contour Γ in clockwise we find that the winding number is -1. Applying Theorem 1 we find that

$$N - P = -1$$

Since the loop transfer function has a pole in the right half plane we have $P = 1$ and we get $N = 0$. The characteristic equation thus has no roots in the right half plane and the closed loop system is stable. $\square$

Bode Plots

The Nyquist curve is one way to represent the frequency response $G(i\omega)$. Another useful representation was proposed by Bode who represented it by two curves, the gain curve and the phase curve. The gain curve gives the value of $G(i\omega)$ as a function of ω and the phase curve gives $\arg G(i\omega)$ as a function of ω. The curves are plotted as shown below with logarithmic scales for frequency and magnitude and linear scale for phase, see Figure 3.16 An useful feature of the Bode plot is that both the gain curve and the phase curve can be approximated by straight lines, see Figure 3.16 where the approximation is shown in dashed lines. This fact was particularly useful when computing tools were not easily accessible.

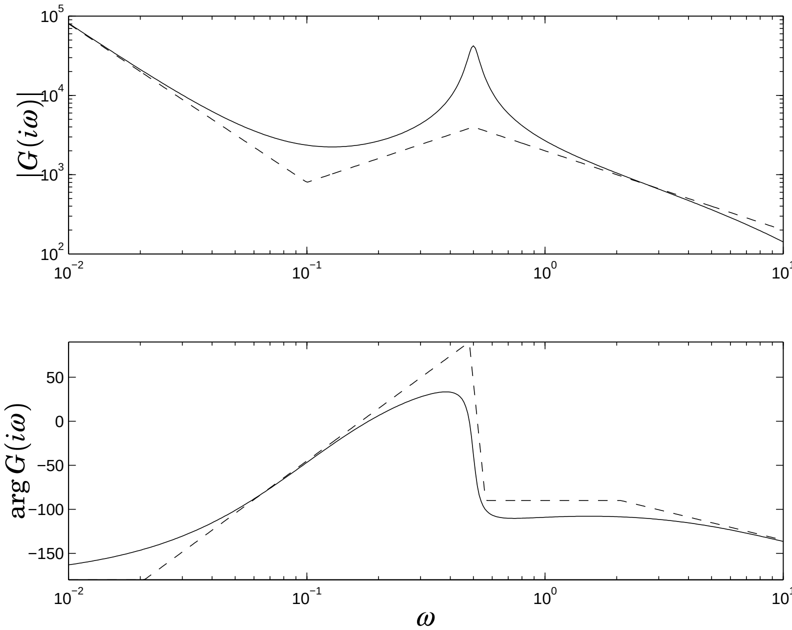

Figure 3.16 Bode diagram of a frequency response. The top plot is the gain curve and bottom plot is the phase curve. The dashed lines show straight line approximations of the curves.

The fact that logarithmic scales were used also simplified the plotting. We illustrate Bode plots with a few examples.

It is easy to sketch Bode plots because with the right scales they have linear asymptotes. This is useful in order to get a quick estimate of the behavior of a system. It is also a good way to check numerical calculations.

Consider first a transfer function which is a polynomial $G(s) = B(s)/A(s)$. We have

$$\log G(s) = \log B(s) - \log A(s)$$

Since a polynomial is a product of terms of the type :

$$s, \quad s+a, \quad s^2 + 2\zeta a s + a^2$$

it suffices to be able to sketch Bode diagrams for these terms. The Bode plot of a complex system is then obtained by composition.

EXAMPLE 3.10—BODE PLOT OF A DIFFERENTIATOR

Consider the transfer function

$$G(s) = s$$

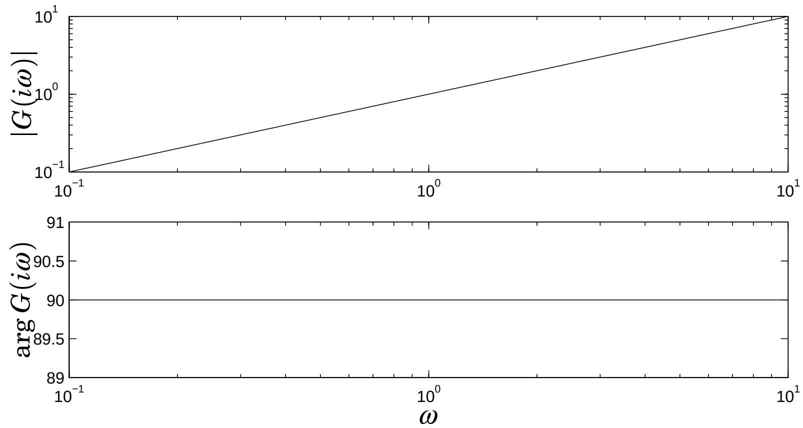

Figure 3.17 Bode plot of a differentiator.

We have $G(i\omega) = i\omega$ which implies

$$\begin{aligned}\log |G(i\omega)| &= \log \omega \\ \arg G(i\omega) &= \pi/2\end{aligned}$$

The gain curve is thus a straight line with slope 1 and the phase curve is a constant at 90°.. The Bode plot is shown in Figure 3.17 □

EXAMPLE 3.11—BODE PLOT OF AN INTEGRATOR

Consider the transfer function

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

We have $G(i\omega) = 1/i\omega$ which implies

$$\begin{aligned}\log |G(i\omega)| &= -\log \omega \\ \arg G(i\omega) &= -\pi/2\end{aligned}$$

The gain curve is thus a straight line with slope -1 and the phase curve is a constant at $-90°$. The Bode plot is shown in Figure 3.18 □

Compare the Bode plots for the differentiator in Figure 3.17 and the integrator in Figure 3.18. The sign of the phase is reversed and the gain curve is mirror imaged in the horizontal axis. This is a consequence of the property of the logarithm.

$$log \frac{1}{G} = -log G = -log|G| - i \arg G$$

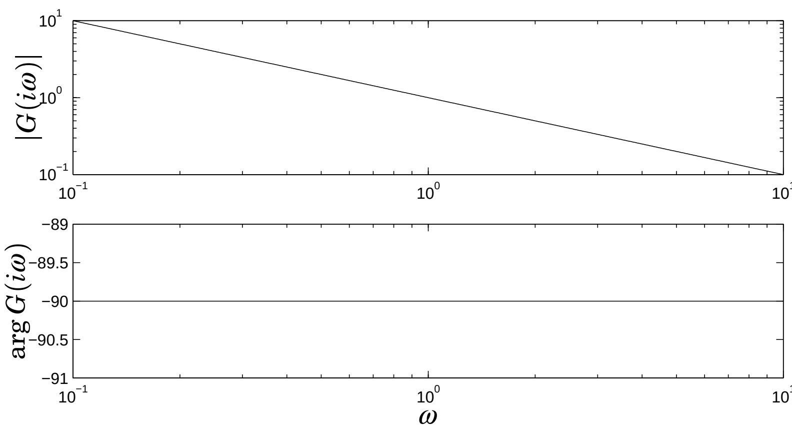

Figure 3.18 Bode plot of an integrator.

EXAMPLE 3.12—BODE PLOT OF A FIRST ORDER FACTOR
Consider the transfer function

$$G(s) = s + a$$

We have

$$G(i\omega) = a + i\omega$$

and it follows that

$$|G(i\omega)| = \sqrt{\omega^2 + a^2}, \quad \arg G(i\omega) = \arctan \omega/a$$

Hence

$$\log |G(i\omega)| = \frac{1}{2} \log (\omega^2 + a^2), \quad \arg G(i\omega) = \arctan \omega/a$$

The Bode Plot is shown in Figure 3.19. Both the gain curve and the phase curve can be approximated by straight lines if proper scales are chosen

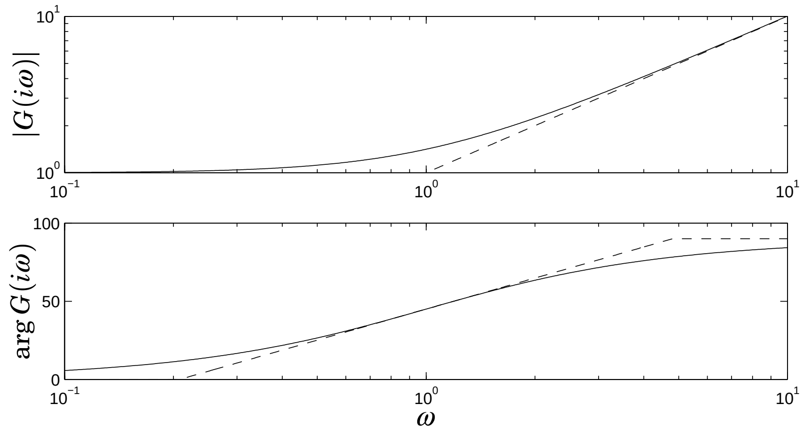

Figure 3.19 Bode plot of a first order factor. The dashed lines show the piece-wise linear approximations of the curves.

and we obtain the following approximations.

$$\log |G(i\omega)| \approx \begin{cases} \log a & \text{if } \omega << a, \\ \log a + \log \sqrt{2} & \text{if } \omega = a, \\ \log \omega & \text{if } \omega >> a \end{cases},$$

$$\arg G(i\omega) \approx \begin{cases} 0 & \text{if } \omega << a, \\ \dfrac{\pi}{4} + \dfrac{1}{2}\log\dfrac{\omega}{a} & \text{if } \omega \approx a, \\ \dfrac{\pi}{2} & \text{if } \omega >> a \end{cases},$$

Notice that a first order system behaves like an integrator for high frequencies. Compare with the Bode plot in Figure 3.18. □

EXAMPLE 3.13—BODE PLOT OF A SECOND ORDER SYSTEM
Consider the transfer function

$$G(s) = s^2 + 2a\zeta s + a^2$$

We have

$$G(i\omega) = a^2 - \omega^2 + 2i\zeta a\omega$$

Hence

$$\log|G(i\omega)| = \frac{1}{2}\log\left(\omega^4 + 2a^2\omega^2(2\zeta^2 - 1) + a^4\right)$$
$$\arg G(i\omega) = \arctan 2\zeta a\omega/(a^2 - \omega^2)$$

Notice that the smallest value of the magnitude $\min_\omega |G(i\omega)| = 1/2\zeta$ is obtained for $\omega = a$ The gain is thus constant for small ω. It has an asymptote with zero slope for low frequencies. For large values of ω the gain is proportional to ω^2, which means that the gain curve has an asymptote with slope 2. The phase is zero for low frequencies and approaches 180° for large frequencies. The curves can be approximated with the following piece-wise linear expressions

$$\log|G(i\omega)| \approx \begin{cases} 2\log a & \textit{if } \omega << a, \\ 2\log a + \log 2\zeta & \textit{if } \omega = a, \\ 2\log\omega & \textit{if } \omega >> a \end{cases},$$
$$\arg G(i\omega) \approx \begin{cases} 0 & \textit{if } \omega << a, \\ \dfrac{\pi}{2} + \dfrac{\omega - a}{a\zeta} & \textit{if } \omega = a, \\ \pi & \textit{if } \omega >> a \end{cases},$$

The Bode Plot is shown in Figure 3.20, the piece-wise linear approximations are shown in dashed lines. □

Sketching a Bode Plot

It is easy to sketch the asymptotes of the gain curves of a Bode plot. This is often done in order to get a quick overview of the frequency response. The following procedure can be used

- Factor the numerator and denominator of the transfer functions.
- The poles and zeros are called break points because they correspond to the points where the asymptotes change direction.
- Determine break points sort them in increasing frequency
- Start with low frequencies
- Draw the low frequency asymptote
- Go over all break points and note the slope changes

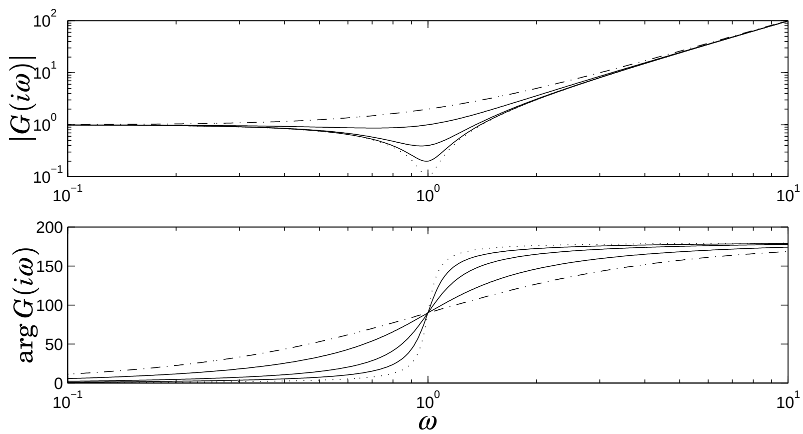

Figure 3.20 Bode plot of a second order factor with $\zeta = 0.05$ (dotted), 0.1, 0.2, 0.5 and 1.0 (dash-dotted). The dashed lines show the piece-wise linear approximations of the curves.

- A crude sketch of the phase curve is obtained by using the relation that, for systems with no RHP poles or zeros, one unit slope corresponds to a phase of 90°

We illustrate the procedure with the transfer function

$$G(s) = \frac{200(s+1)}{s(s+10)(s+200)} = \frac{1+s}{10s(1+0.1s)(1+0.01s)}$$

The break points are 0.01, 0.1, 1. For low frequencies the transfer function can be approximated by

$$G(s) \approx \frac{1}{10s}$$

Following the procedure we get

- The low frequencies the system behaves like an integrator with gain 0.1. The low frequency asymptote thus has slope -1 and it crosses the axis of unit gain at $\omega = 0.1$.
- The first break point occurs at $\omega = 0.01$. This break point corresponds to a pole which means that the slope decreases by one unit to -2 at that frequency.
- The next break point is at $\omega = 0.1$ this is also a pole which means that the slope decreases to -3.

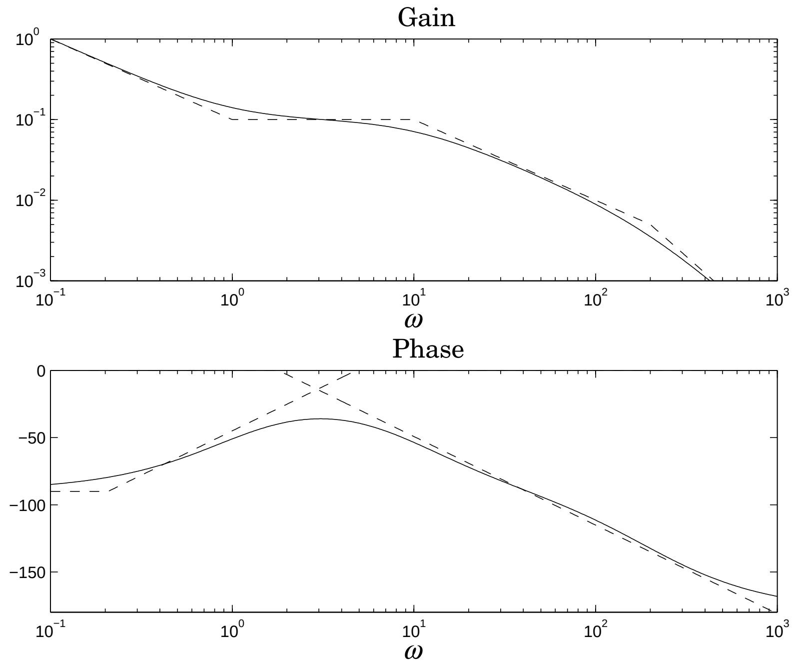

Figure 3.21 Illustrates how the asymptotes of the gain curve of the Bode plot can be sketched. The dashed curves show the asymptotes and the full lines the complete plot.

- The next break point is at $\omega = 1$, since this is a zero the slope increases by one unit to -2.

Figure 3.21 shows the asymptotes of the gain curve and the complete Bode plot.

Gain and Phase Margins

The gain and phase margins can easily be found from the Bode plot of the loop transfer function. Recall that the gain margin tells how much the gain has to be increased for the system to reach instability. To determine the gain margin we first find the frequency ω_{pc} where the phase is $-180°$. This frequency is called the phase crossover frequency. The gain margin is the inverse of the gain at that frequency. The phase margin tells how the phase lag required for the system to reach instability. To determine the phase margin we first determine the frequency ω_{gc} where the gain of the loop transfer function is one. This frequency is called the gain crossover frequency. The phase margin is the phase of the loop transfer function

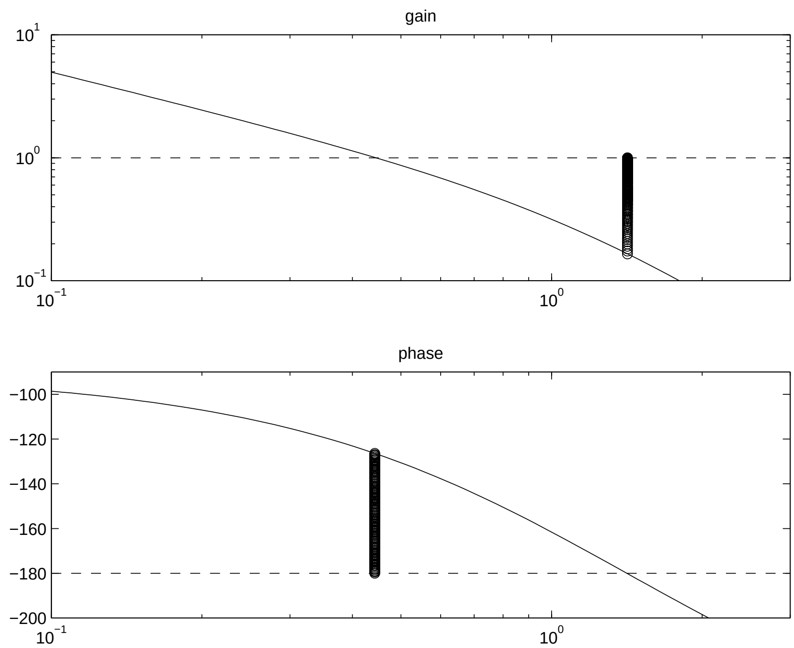

Figure 3.22 Finding gain and phase margins from the Bode plot of the loop transfer function.

at that frequency plus 180°. Figure 3.22 illustrates how the margins are found in the Bode plot of the loop transfer function.

Bode's Relations

Analyzing the Bode plots in the examples we find that there appears to be a relation between the gain curve and the phase curve. Consider e.g. the curves for the differentiator in Figure 3.17 and the integrator in Figure 3.18. For the differentiator the slope is $+1$ and the phase is constant $pi/2$ radians. For the integrator the slope is -1 and the phase is $-pi/2$. Bode investigated the relations between the curves and found that there was a unique relation between amplitude and phase for many systems. In particular he found the following relations for system with no poles and zeros in the right half plane.

$$
\begin{aligned}
\arg G(i\omega_0) &= \frac{2\omega_0}{\pi}\int_0^\infty \frac{\log|G(i\omega)| - \log|G(i\omega_0)|}{\omega^2-\omega_0^2} d\omega\\
&= \frac{1}{\pi}\int_0^\infty \frac{d\log|G(i\omega)|}{d\log\omega}\log\Big|\frac{\omega+\omega_0}{\omega-\omega_0}\Big| d\\
\approx \frac{\pi}{2}\frac{d\log|G(i\omega)|}{d\log\omega}\\
\frac{\log|G(i\omega)|}{\log|G(i\omega_0)|} &= -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{\omega^{-1}\arg G(i\omega) - \omega_0^{-1}\arg G(i\omega_0)}{\omega^2-\omega_0^2} d\omega\\
&= -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{d\big(\omega^{-1}\arg G(i\omega)\big)}{d\omega}\log\Big|\frac{\omega+\omega_0}{\omega-\omega_0}\Big| d\omega
\end{aligned}
\tag{3.31}
$$

The formula for the phase tells that the phase is a weighted average of the logarithmic derivative of the gain, approximatively

$$
\arg G(i\omega) \approx \frac{\pi}{2}\frac{d\log|G(i\omega)|}{d\log\omega} \tag{3.32}
$$

This formula implies that a slope of $+1$ corresponds to a phase of $\pi/2$, which holds exactly for the differentiator, see Figure 3.17. The exact formula (3.31) says that the differentiated slope should be weighted by the kernel

$$
\int_0^\infty \log\Big|\frac{\omega+\omega_0}{\omega-\omega_0}\Big| d\omega = \frac{\pi^2}{2}
$$

Figure 3.23 is a plot of the kernel.

Minimum Phase and Non-minimum Phase

Bode's relations hold for systems that do not have poles and zeros in the left half plane. Such systems are called minimum phase systems. One nice property of these systems is that the phase curve is uniquely given by the gain curve. These systems are also relatively easy to control. Other systems have larger phase lag, i.e. more negative phase. These systems are said to be non-minimum phase, because they have more phase lag than the equivalent minimum phase systems. Systems which do not have minimum phase are more difficult to control. Before proceeding we will give some examples.

EXAMPLE 3.14—A TIME DELAY

The transfer function of a time delay of T units is

$$
G(s) = e^{-sT}
$$

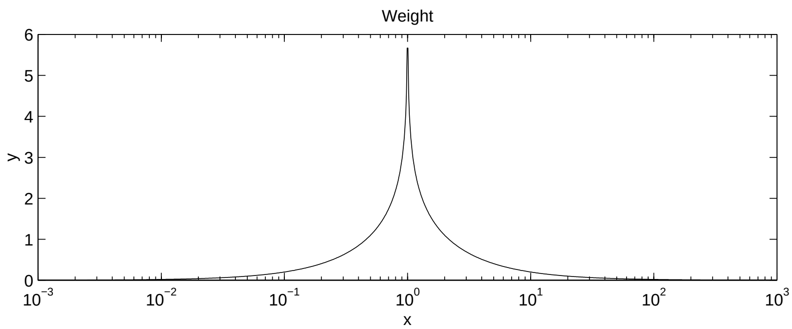

Figure 3.23 The weighting kernel in Bodes formula for computing the phase from the gain.

This transfer function has the property

$$|G(i\omega)| = 1, \quad \arg G(i\omega) = -\omega T$$

Notice that the gain is one. The minimum phase system which has unit gain has the transfer function $G(s) = 1$. The time delay thus has an additional phase lag of ωT. Notice that the phase lag increases with increasing frequency. Figure 3.24 □

It seems intuitively reasonable that it is not possible to obtain a fast response of a system with time delay. We will later show that this is indeed the case. □

Next we will consider a system with a zero in the right half plane

EXAMPLE 3.15—SYSTEM WITH A RHP ZERO
Consider a system with the transfer function

$$G(s) = \frac{a-s}{a+s}$$

This transfer function has the property

$$|G(i\omega)| = 1, \quad \arg G(i\omega) = -2 \arctan \frac{\omega}{a}$$

Notice that the gain is one. The minimum phase system which has unit gain has the transfer function $G(s) = 1$. In Figure 3.25 we show the Bode plot of the transfer function. The Bode plot resembles the Bode plot for a time delay which is not surprising because the exponential function e^{-sT}

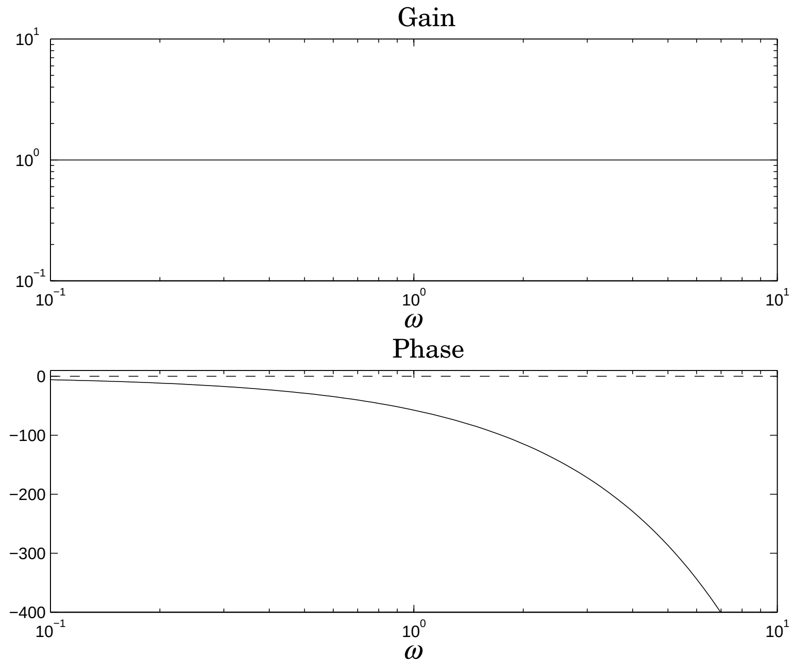

Figure 3.24 Bode plot of a time delay which has the transfer function $G(s) = e^{-s}$.

can be approximated by

$$e^{-sT} = \frac{1 - sT/2}{1 + sT/2}$$

The largest phase lag of a system with a zero in the RHP is however *pi*. □

We will later show that the presence of a zero in the right half plane severely limits the performance that can be achieved. We can get an intuitive feel for this by considering the step response of a system with a right half plane zero. Consider a system with the transfer function $G(s)$ that has a zero at $s = -\alpha$ in the right half plane. Let h be the step response of the system. The Laplace transform of the step response is given by

$$H(s) = \frac{G(s)}{s} = \int_0^t e^{-st} h(t) dt$$

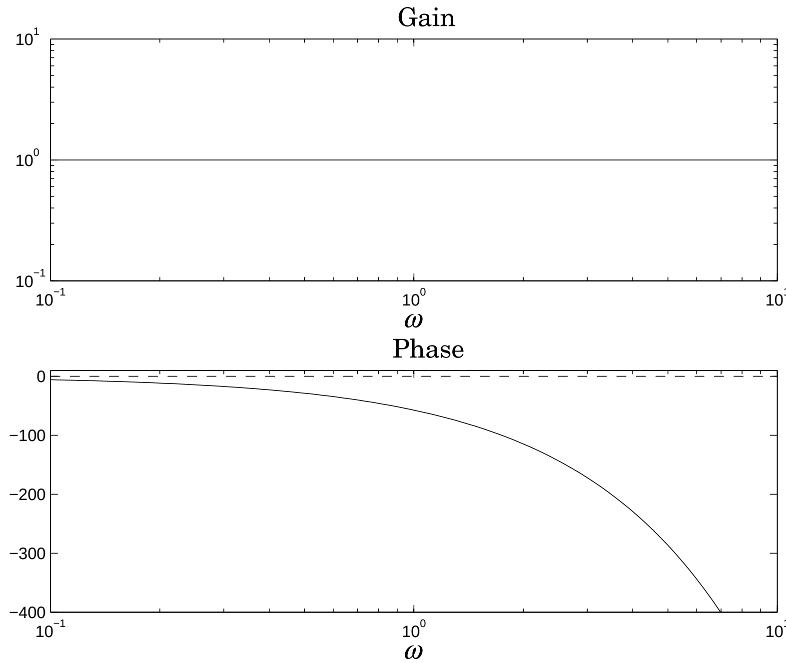

Figure 3.25 Bode plot of a the transfer function $G(s) = \dfrac{a-s}{a+s}$.

Since $G(\alpha)$ is zero we have

$$0 = \int_0^t e^{-\alpha t} h(t) dt$$

Since $e^{-\alpha t}$ is positive it follows that the step response $h(t)$ must be negative for some t. This is illustrated in Figure 3.26 which shows the step response of a system having a zero in the right half plane. Notice that the output goes in the wrong direction initially. This is sometimes referred to as inverse response. It seems intuitively clear that such systems are difficult to control fast. This is indeed the case as will be shown in Chapter 5. We have thus found that systems with time delays and zeros in the right half plane have similar properties. Next we will consider a system with a right half plane pole.

EXAMPLE 3.16—SYSTEM WITH A RHP POLE
Consider a system with the transfer function

$$G(s) = \frac{s+a}{s-a}$$

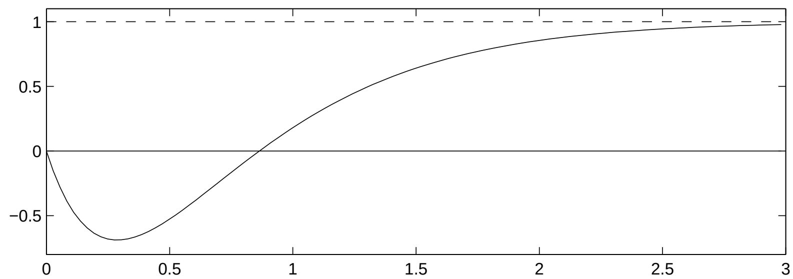

Figure 3.26 Step response of a system with a zero in the right half plane. The system has the transfer function $G(s) = \dfrac{6(-s+1)}{s^2+5s+6}$.

This transfer function has the property

$$|G(i\omega)| = 1, \quad \arg G(i\omega) = -2\arctan\frac{a}{\omega}$$

Notice that the gain is one. The minimum phase system which has unit gain has the transfer function $G(s) = 1$. In Figure 3.27 we show the Bode plot of the transfer function. □

Comparing the Bode plots for systems with a right half plane pole and a right half plane zero we find that the additional phase lag appears at high frequencies for a system with a right half plane zero and at low frequencies for a system with a right half plane pole. This means that there are significant differences between the systems. When there is a right half plane pole high frequencies must be avoided by making the system slow. When there is a right half plane zero low frequencies must be avoided and it is necessary to control these systems rapidly. This will be discussed more in Chapter 5.

It is a severe limitation to have poles and zeros in the right half plane. Dynamics of this type should be avoided by redesign of the system. The zeros of a system can also be changed by moving sensors or by introducing additional sensors. Unfortunately systems which are non-minimum phase are not uncommon i real life. We end this section by giving a few examples.

EXAMPLE 3.17—HYDRO ELECTRIC POWER GENERATION

The transfer function from tube opening to electric power in a hydroelectric power station has the form

$$\frac{P(s)}{A(s)} = \frac{P_0}{A_0}\frac{1-2sT}{1+sT}$$

where T is the time it takes sound to pass along the tube. □

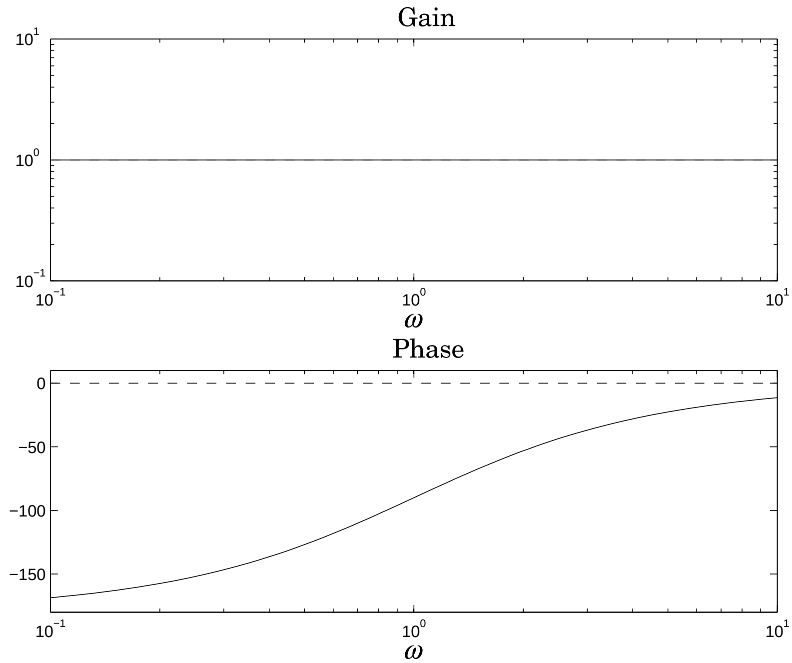

Figure 3.27 Bode plot of a the transfer function $G(s) = \dfrac{s+a}{s-a}$ which has a pole in the right half plane.

EXAMPLE 3.18—LEVEL CONTROL IN STEAM GENERATORS
Consider the problem of controlling the water level in a steam generator. The major disturbance is the variation of steam taken from the unit. When more steam is fed to the turbine the pressure drops. There is typically a mixture of steam and water under the water level. When pressure drops the steam bubbles expand and the level increases momentarily. After some time the level will decrease because of the mass removed from the system. □

EXAMPLE 3.19—FLIGHT CONTROL
The transfer function from elevon to height in an airplane is non-minimum phase. When the elevon is raised there will be a force that pushes the rear of the airplane down. This causes a rotation which gives an increase of the angle of attack and an increase of the lift. Initially the aircraft will however loose height. The Wright brothers understood this and used control surfaces in the front of the aircraft to avoid the effect. □

EXAMPLE 3.20—BACKING A CAR
Consider backing a car close to a curb. The transfer function from steering angle to distance from the curve is non-minimum phase. This is a mechanism that is similar to the aircraft. □

EXAMPLE 3.21—REVENUE FROM DEVELOPMENT
The relation between revenue development effort in a new product development is a non-minimum phase system. This means that such a system is very difficult to control tightly. □

3.6 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. The control variables are represented by another vector u and the measured signal by the vector y. A system can then be represented by the model

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) \\ y &= g(x,u) \end{aligned} \tag{3.33}$$

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 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

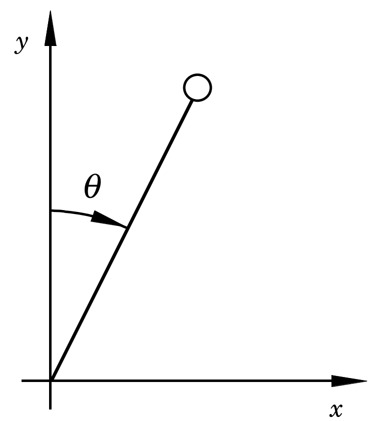

Figure 3.28 An inverted pendulum. *The picture should be mirrored.*

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 3.22—THE DOUBLE INTEGRATOR
Consider a system described by

$$\begin{aligned}\frac{dx}{dt} &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 1 \end{pmatrix} u \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} x\end{aligned} \tag{3.34}$$

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

EXAMPLE 3.23—THE INVERTED PENDULUM
Consider the inverted pendulum in Figure 3.28. 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.

Newtons law of conservation of angular momentum becomes

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

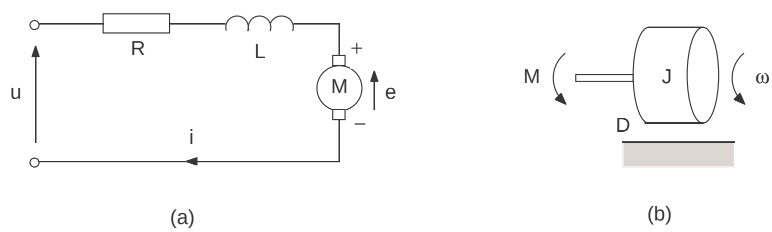

Figure 3.29 Schematic diagram of an electric motor.

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

$$\frac{dx}{dt} = \begin{pmatrix} x_2 \\ \dfrac{mgl}{J}\sin x_1 + \dfrac{mlu}{J}\cos x_1 \end{pmatrix}$$
$$y = x_1$$

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

$$\frac{dx}{dt} = \begin{pmatrix} x_2 \\ \sin x_1 + u\cos x_1 \end{pmatrix} \tag{3.35}$$
$$y = x_1$$

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

EXAMPLE 3.24—AN ELECTRIC MOTOR

A schematic picture of an electric motor is shown in Figure 3.29 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

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

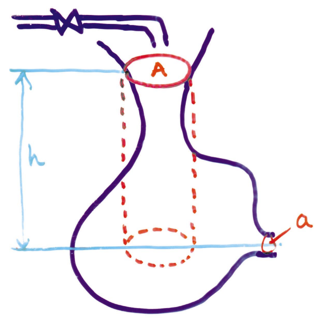

Figure 3.30 A schematic picture of a water tank.

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

$$\frac{dx}{dt} = \begin{pmatrix} -\frac{D}{J} & 0 & \frac{k}{J} \\ 0 & 0 & \frac{1}{C} \\ -\frac{kD}{JL} & -\frac{1}{L} & \frac{k^2}{JL} - \frac{R}{L} \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{L} \end{pmatrix} u y = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} x \tag{3.36}$$

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

EXAMPLE 3.25—THE WATER TANK
Consider a tank with water where the input is the inflow and there is free outflow, see Figure 3.30 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

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

The tank is thus a nonlinear system of first order. □

Equilibria

To investigate a system we will first determine the equilibria. Consider the system given by (3.33) which is assumed to be time-invariant. Let the control signal be constant $u = u_0$. The equilibria are states x_0 such that the $dx/dt = 0$. Hence

$$f(x_0, u_0) = 0$$

Notice that there may be several equilibria.

For second order systems the state equations can be visualized by plotting the velocities for all points in the state space. This graph is called the phase plane shows the behavior qualitative. The equilibria corresponds to points where the velocity is zero. We illustrate this with an example.

EXAMPLE 3.26—THE PHASE PLANE
Consider the

$$\frac{dx}{dt} = \begin{pmatrix} x_2 - x_2^3 \\ -x_1 - x_2^2 \end{pmatrix}$$

The equilibria are given by

$$\begin{aligned} x_2 - x_2^3 &= 0 \\ x_1 - x_2^2 &= 0 \end{aligned}$$

There are three equilibria:

$$x_1 = -1 \quad x_2 = -1$$

$$x_1 = -1 \quad x_2 = 1$$

$$x_1 = 0 \quad x_2 = 0$$

The phase plane is shown in Figure 3.31. The phase plane is a good visualization of solutions for second order systems. It also illustrates that nonlinear systems can be interpreted as a vector field or a flow. □

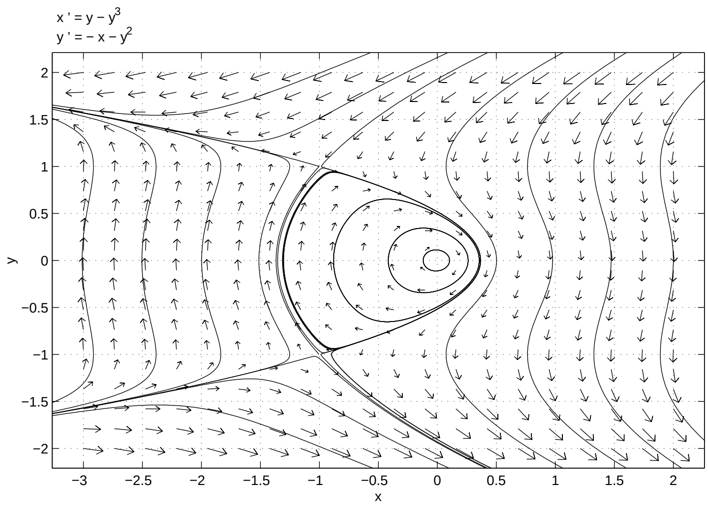

Figure 3.31 Phase plane for the second order system $dx_1/dt = x_2 - x_2^3, dx_2/dt = -x_1 - x_2^2$.

Linearization

Nonlinear systems are unfortunately difficult. It is fortunate that many aspects of control can be understood from linear models. This is particularly true for regulation problems where it is intended to keep variables close to specified values. When deviations are small the nonlinearities can be approximated by linear functions. With efficient control the deviations are small and the approximation works even better. In this section we will show how nonlinear dynamics systems are approximated. We will start with an example that shows how static systems are approximated.

EXAMPLE 3.27—LINEARIZATION OF STATIC SYSTEM
Consider the system

$$y = g(u)$$

A Taylor series expansion around $u = u_0$ gives

$$y = g(u_0) + g'(u_0)(u - u_0) + \dots$$

The linearized model is

$$y - y_0 = g'(u_0)(u - u_0)$$

The linearized model thus replaces the nonlinear curve by its tangent at the operating point. □

Linearization of dynamic systems is done in the same way. We start by determining the appropriate equilibria. The nonlinear systems are then approximated using Taylor series expansions. Consider the system

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

Consider small deviations from the equilibrium!

$$x = x_0 + \delta x, \quad u = u_0 + \delta u, \quad y = y_0 + \delta y$$

Make a series expansion of the differential equation and neglect terms of second and higher order. This gives

$$\begin{aligned}\frac{dx}{dt} &= f(x_0 + \delta x, u_0 + \delta u) \approx f(x_0, u_0) + \frac{\partial f}{\partial x}(x_0, u_0)\delta x + \frac{\partial f}{\partial u}(x_0, u_0)\delta u\\ y &= g(x_0 + \delta x, u_0 + \delta u) \approx y_0 + \frac{\partial g}{\partial x}(x_0, u_0)\delta x + \frac{\partial g}{\partial u}(x_0, u_0)\delta u\end{aligned}$$

We have $f(x_0, u_0) = 0$ because x_0 is an equilibrium and we find the following approximation for small deviations around the equilibrium.

$$\begin{aligned}\frac{d(x - x_0)}{dt} &= A(x - x_0) + B(u - u_0)\\ y - y_0 &= C(x - x_0) + D(u - u_0)\end{aligned}$$

where

$$\begin{aligned}A &= \frac{\partial f}{\partial x}(x_0, u_0) & B &= \frac{\partial f}{\partial u}(x_0, u_0)\\ C &= \frac{\partial g}{\partial x}(x_0, u_0) & D &= \frac{\partial g}{\partial u}(x_0, u_0)\end{aligned}$$

The linearized equation is thus a linear time-invariant system, compare with (3.37). It is common practice to relabel variables and simply let x, y and u denote deviations from the equilibrium.

We illustrate with a few examples

EXAMPLE 3.28—LINEARIZATION OF THE WATER TANK

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

To determine the equilibrium we assume that the inflow is constant $qin = q_0$. It follows that

$$q_{out} = q_{in} = q_0 = a\sqrt{2gh_0}$$
$$h_0 = \frac{q_0^2}{2ga^2}$$

Let A_0 be the cross section A at level h_0, introduce the deviations. The linearized equations are

$$\frac{d\delta h}{dt} = -\frac{a\sqrt{2gh_0}}{2A_0h_0}\delta h + \frac{1}{A_0}\delta q_{in}$$
$$\delta q_{out} = \frac{a\sqrt{2gh_0}}{h_0}\delta h = \frac{q_0}{h_0}\delta h$$

The parameter

$$T = \frac{2A_0h_0}{q_0} = 2 \times \frac{\text{Total water volume } [m^3]}{\text{Flow rate } [\mathrm{m}^3/\mathrm{s}]}$$

is called the time constant of the system. Notice that $T/2$ is the time it takes to fill the volume A_0h_0 with the steady state flow rate q_0 □

EXAMPLE 3.29—LINEARIZATION OF THE INVERTED PENDULUM
Consider the inverted pendulum in Example 3.23 which is described by (3.35). If the control signal is zero the equilibria are given by

$$x_2 = 0$$
$$\sin x_1 = 0$$

i.e. $x_2 = \theta/dt$ and $x_1 = \theta = 0$ and $x_1 = \theta = \pi$. The first equilibrium corresponds to the pendulum standing upright and the second to the pendulum hanging straight down. We have

$$\frac{\partial f(x,0)}{\partial x} = \begin{pmatrix} 0 & 1 \\ \cos x_1 - u \sin x_1 & 0 \end{pmatrix}, \quad \frac{\partial f}{\partial u} = \begin{pmatrix} 0 \\ \cos x_1 \end{pmatrix},$$

Evaluating the derivatives at the upper equilibrium $u = 0$, $x_1 = 0$ and $x_2 = 0$ we get

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

For the equilibrium when then pendulum is hanging down, $u = 0$, $x_1 = \pi$ and $x_2 = 0$ we have instead

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \end{pmatrix}.$$

□

3.7 Linear Time-Invariant Systems

The model

$$\begin{aligned}\frac{dx}{dt} &= Ax + Bu \\ y &= Cx + Du\end{aligned} \tag{3.37}$$

is one of the standard models in control. In this section we will present an in depth treatment. Let us first recall that x is the state vector, u the control, y the measurement. The model is nice because it can represent systems with many inputs and many outputs in a very compact form. Because of the advances in numeric linear algebra there are also much powerful software for making computations. Before going into details we will present some useful results about matrix functions. It is assumed that the reader is familiar with the basic properties of matrices.

Matrix Functions

Some basic facts about matrix functions are summarized in this section. Let A be a square matrix, since it is possible to compute powers of matrices we can define a matrix polynomial as follows

$$f(A) = a_0 I + a_1 A + \ldots + a_n A^n$$

Similarly if the function $f(x)$ has a converging series expansion we can also define the following matrix function

$$f(A) = a_0 I + a_1 A + \ldots + a_n A^n + \ldots$$

The matrix exponential is a nice useful example which can be defined as

$$e^{At} = I + At + \frac{1}{2}(At)^2 + \ldots + \frac{1}{n!}A^n t^n + \ldots$$

Differentiating this expression we find

$$\begin{aligned}\frac{de^{At}}{dt} &= A + A^2 t + \frac{1}{2}A^3 t^2 + \ldots + \frac{1}{(n-1)!}A^n t^{n-1} + \ldots \\ &= A(= I + At + \frac{1}{2}(At)^2 + \ldots + \frac{1}{n!}A^n t^n + \ldots) = Ae^{At}\end{aligned}$$

The matrix exponential thus has the property

$$\frac{de^{At}}{dt} = Ae^{At} = e^{At}A \tag{3.38}$$

Matrix functions do however have other interesting properties. One result is the following.

THEOREM 3.3—CAYLEY-HAMILTON
Let the $n \times n$ matrix A have the characteristic equation

$$\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} \ldots + a_n = 0$$

then it follows that

$$\det(\lambda I - A) = A^n + a_1 A^{n-1} + a_2 A^{n-2} \ldots + a_n I = 0$$

A matrix satisfies its characteristic equation.

PROOF 3.2
If a matrix has distinct eigenvalues it can be diagonalized and we have $A = T^{-1}\Lambda T$. This implies that

$$\begin{aligned}
A^2 &= T^{-1}\Lambda T T^{-1}\Lambda T = T^{-1}\Lambda^2 T \\
A^3 &= T^{-1}\Lambda T A^2 = T^{-1}\Lambda T T^{-1}\Lambda^2 T = T^{-1}\Lambda^3 T
\end{aligned}$$

and that $A^n = T^{-1}\Lambda^n T$. Since λ_i is an eigenvalue it follows that

$$\lambda_i^n + a_1\lambda_i^{n-1} + a_2\lambda_i^{n-2} \ldots + a_n = 0$$

Hence

$$\Lambda_i^n + a_1\Lambda_i^{n-1} + a_2\Lambda_i^{n-2} \ldots + a_n I = 0$$

Multiplying by T^{-1} from the left and T from the right and using the relation $A^k = T^{-1}\Lambda^k T$ now gives

$$A^n + a_1 A^{n-1} + a_2 A^{n-2} \ldots + a_n I = 0$$

$\square$

The result can actually be sharpened. The minimal polynomial of a matrix is the polynomial of lowest degree such that $g(A) = 0$. The characteristic polynomial is generically the minimal polynomial. For matrices with common eigenvalues the minimal polynomial may, however, be different from the characteristic polynomial. The matrices

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

have the minimal polynomials

$$g_1(\lambda) = \lambda - 1, \qquad g_2(\lambda) = (\lambda - 1)^2$$

A matrix function can thus be written as

$$f(A) = c_0 I + c_1 A + \ldots + c_{k-1} A^{k-1}$$

where k is the degree of the minimal polynomial.

Solving the Equations

Using the matrix exponential the solution to (3.37) can be written as

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau \tag{3.39}$$

To prove this we differentiate both sides and use the property 3.38) of the matrix exponential. This gives

$$\frac{dx}{dt} = Ae^{At}x(0) + \int_0^t Ae^{A(t-\tau)}Bu(\tau)d\tau + Bu(t) = Ax + Bu$$

which prove the result. Notice that the calculation is essentially the same as for proving the result for a first order equation.

Input-Output Relations

It follows from Equations (3.37) and (3.39) that the input output relation is given by

$$y(t) = Ce^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t)$$

Taking the Laplace transform of (3.37) under the assumption that $x(0) = 0$ gives

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

Solving the first equation for $X(s)$ and inserting in the second gives

$$\begin{aligned} X(s) &= [sI - A]^{-1}BU(s) \\ Y(s) &= \big(C[sI - A]^{-1}B + D\big)U(s) \end{aligned}$$

The transfer function is thus

$$G(s) = C[sI - A]^{-1}B + D \tag{3.40}$$

we illustrate this with an example.

EXAMPLE 3.30—TRANSFER FUNCTION OF INVERTED PENDULUM
The linearized model of the pendulum in the upright position is characterized by the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad D = 0.$$

The characteristic polynomial of the dynamics matrix A is

$$\det\,(sI - A) = \det \begin{pmatrix} s & -1 \\ -1 & s \end{pmatrix} = s^2 - 1$$

Hence

$$(sI - A)^{-1} = \frac{1}{s^2 - 1} \det \begin{pmatrix} s & 1 \\ 1 & s \end{pmatrix}$$

The transfer function is thus

$$G(s) = C[sI - A]^{-1}B = \frac{1}{s^2 - 1} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} s & 1 \\ 1 & s \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{s^2 - 1}$$

□

Transfer function and impulse response remain invariant with coordinate transformations.

$$\tilde{g}(t) = \tilde{C}e^{\tilde{A}t}\tilde{B} = CT^{-1}e^{TAT^{-1}t}TB = Ce^{At}B = g(t)$$

and

$$\begin{aligned} \tilde{G}(s) &= \tilde{C}(sI - \tilde{A})^{-1}\tilde{B} = CT^{-1}(sI - TAT^{-1})^{-1}TB \\ XS &= C(sI - A)^{-1}B = G(s) \end{aligned}$$

Consider the system

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

To find the input output relation we can differentiate the output and we obtain

$$
\begin{aligned}
y &= Cx \\
\frac{dy}{dt} &= C\frac{dx}{dt} = CAx + CBu \\
\frac{d^2y}{dt^2} &= CA\frac{dx}{dt} + CB\frac{du}{dt} = CA^2x + CABu + CB\frac{du}{dt} \\
&\vdots \\
\frac{d^ny}{dt^n} &= CA^nx + CA^{n-1}Bu + CA^{n-2}B\frac{du}{dt} + \ldots + CB\frac{d^{n-1}u}{dt^{n-1}}
\end{aligned}
$$

Let a_k be the coefficients of the characteristic equation. Multiplying the first equation by a_n, the second by a_{n-1} etc we find that the input-output relation can be written as.

$$
\frac{d^ny}{dt^n} + a_1\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_ny = B_1\frac{d^{n-1}u}{dt^{n-1}} + B_2\frac{d^{n-2}u}{dt^{n-2}} + \ldots + B_nu,
$$

where the matrices B_k are given by.

$$
\begin{aligned}
B_1 &= CB \\
B_2 &= CAB + a_1CB \\
B_3 &= CA^2B + a_1CAB + a_2CB \\
&\vdots \\
B_n &= CA^{n-1}B + a_1CA^{n-1}B + \ldots + a_{n-1}CB
\end{aligned}
$$

Coordinate Changes

The components of the input vector u and the output vector y are unique physical signals, but the state variables depend on the coordinate system chosen to represent the state. The elements of the matrices A, B and C also depend on the coordinate system. The consequences of changing coordinate system will now be investigated. Introduce new coordinates z by the transformation $z = Tx$, where T is a regular matrix. It follows from (3.37) that

$$
\begin{aligned}
\frac{dz}{dt} &= T(Ax + Bu) = TAT^{-1}z + TBu = \tilde{A}z + \tilde{B}u \\
y &= Cx + DU = CT^{-1}z + Du = \tilde{C}z + Du
\end{aligned}
$$

The transformed system has the same form as (3.37) but the matrices A, B and C are different

$$
\tilde{A} = TAT^{-1}, \qquad \tilde{B} = TB, \qquad \tilde{C} = CT^{-1}, \qquad \tilde{D} = D \tag{3.41}
$$

It is interesting to investigate if there are special coordinate systems that gives systems of special structure.

The Diagonal Form Some matrices can be transformed to diagonal form, one broad class is matrices with distinct eigenvalues. For such matrices it is possible to find a matrix T such that the matrix TAT^{-1} is a diagonal i.e.

$$TAT^{-1} = \Lambda == \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

The transformed system then becomes

$$\begin{aligned} \frac{dz}{dt} &= \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} z + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} u \\ y &= \begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_n \end{pmatrix} z + Du \end{aligned} \tag{3.42}$$

The transfer function of the system is

$$G(s) = \sum_{i=1}^{n} \frac{\beta_i \gamma_i}{s - \lambda_i} + D$$

Notice appearance of eigenvalues of matrix A in the denominator.

Reachable Canonical Form Consider a system described by the n-th order differential equation

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

To find a representation in terms of state model we first take Laplace transforms

$$Y(s) = \frac{b_1 s^{n-1} + \ldots + b_1 s + b_n}{s^n + a_1 s^{n-1} + \ldots + a_{n-1} s + a_n} U(s) = \frac{b_1 s^{n-1} + \ldots + b_1 s + b_n}{A(s)} U(s)$$

Introduce the state variables

$$
\begin{aligned}
X_1(s) &= \frac{s^{n-1}}{A(s)}U(s)\\
X_2(s) &= \frac{s^{n-2}}{A(s)}U(s) = \frac{1}{s}X_1(s)\\
X_3(s) &= \frac{s^{n-2}}{A(s)}U(s) = \frac{1}{s^2}X_1(s) = \frac{1}{s}X_2(s)\\
&\vdots\\
X_n(s) &= \frac{1}{A(s)}U(s) = \frac{1}{s^{n-1}}X_1(s) = \frac{1}{s}X_{n-1}(s)
\end{aligned}
\tag{3.43}
$$

Hence

$$
\begin{aligned}
(s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n)X_1(s) = s^{n-1}U(s)\\
sX_1(s) + a_1X_1(s) + a_2\frac{1}{s}X_1(s) + \ldots + a_n\frac{1}{s^{n-1}X_1(s)} = U(s)\\
sX_1(s) + a_1X_2(s) + a_2X_2(s) + \ldots + a_nX_n(s) = U(s)
\end{aligned}
$$

Consider the equation for $X_1(s)$, dividing by s^{n-1} we get

$$
sX_1(s) + a_1X_2(s) + a_2X_2(s) + \ldots + a_nX_n(s) = U(s)
$$

Conversion to time domain gives

$$
\frac{dx_1}{dt} = -a_1x_1 - a_2x_2 - \ldots - a_nx_n + u
$$

(3.43) also implies that

$$
\begin{aligned}
X_2(s) &= \frac{1}{s}X_1(s)\\
X_3(s) &= \frac{1}{s}X_2(s)\\
&\vdots\\
X_n(s) &= \frac{1}{s}X_{n-1}
\end{aligned}
$$

Transforming back to the time domain gives

$$
\begin{aligned}
\frac{dx_2}{dt} &= x_1\\
\frac{dx_3}{dt} &= x_2\\
&\vdots\\
\frac{dx_n}{dt} &= x_{n-1}
\end{aligned}
$$

With the chosen state variables the output is given by

$$Y(s) = b_1X_1(s) + b_2X_2(s) + \ldots + b_nX_n(s)$$

Collecting the parts we find that the equation can be written as

$$
\begin{aligned}
\frac{dz}{dt} &= \begin{pmatrix} -a_1 & -a_2 & \ldots & a_{n-1} & -a_n\\ 1 & 0 & & 0 & 0\\ 0 & 1 & & 0 & 0\\ \vdots & & & & \\ 0 & 0 & & 1 & 0 \end{pmatrix} z + \begin{pmatrix} 1\\ 0\\ 0\\ \vdots\\ 0 \end{pmatrix} u\\
y &= \begin{pmatrix} b_1 & b_2 & \ldots & b_{n-1} & b_n \end{pmatrix} z + Du
\end{aligned}
\tag{3.44}
$$

The system has the characteristic polynomial

$$D_n(s) = \det \begin{pmatrix} s+a_1 & a_2 & \ldots & a_{n-1} & a_n\\ -1 & s & & 0 & 0\\ 0 & -1 & & 0 & 0\\ \vdots & & & & \\ 0 & 0 & & -1 & s \end{pmatrix}$$

Expanding the determinant by the last row we find that the following recursive equation for the polynomial $D_n(s)$.

$$D_n(s) = sD_{n-1}(s) + a_n$$

It follows from this equation that

$$D_n(s) = s^n + a_1s^{n-1} + \ldots + a_{n-1}s + a_n$$

Transfer function

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n} + D$$

The numerator of the transfer function $G(s)$ is the characteristic polynomial of the matrix A. This form is called the reachable canonical for for reasons that will be explained later in this Section.

Observable Canonical Form The reachable canonical form is not the only way to represent the transfer function

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n}$$

another representation is obtained by the following recursive procedure. Introduce the Laplace transform X_1 of first state variable as

$$X_1 = Y = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n} U$$

then

$$(s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n) X_1 = (b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n) U$$

Dividing by s^{n-1} and rearranging the terms we get

$$sX_1 = -a_1 X_1 + b_1 U + X_2$$

where

$$\begin{aligned} s^{n-1} X_2 = &-(a_2 s^{n-2} + a_3 s^{n-3} + \ldots + a_n) X_1 \\ &+ \big(b_2 s^{n-2} + b_3 s^{n-3} + \ldots + b_n\big) U \end{aligned}$$

Dividing by s^{n-2} we get

$$sX_2 = -a_2 X_2 + b_2 U + X_3$$

where

$$s^{n-2} X_3 = -(a_3 s^{n-3} + a_4^{n-4} \ldots + a_n\big) X_1 + \big(b_3 s^{n-3} + \ldots + b_n\big) U$$

Dividing by s^{n-3} gives

$$sX_3 = -a_3 X_1 ? b_3 U + X_4$$

Proceeding in this we we finally obtain

$$X_n = -a_n X_1 + b_1 U$$

Collecting the different parts and converting to the time domain we find that the system can be written as

$$\frac{dz}{dt} = \begin{pmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & & 0 \\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & & 1 \\ -a_n & 0 & 0 & & 0 \end{pmatrix} z + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} u \tag{3.45}$$

$$y = \begin{pmatrix} 1 & 0 & 0 \dots & 0 \end{pmatrix} z + Du$$

Transfer function

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n} + D$$

The numerator of the transfer function $G(s)$ is the characteristic polynomial of the matrix A.

Consider a system described by the n-th order differential equation

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

Reachability

We will now disregard the measurements and focus on the evolution of the state which is given by

$$\frac{sx}{dt} = Ax + Bu$$

where the system is assumed to be or order n. A fundamental question is if it is possible to find control signals so that any point in the state space can be reached. For simplicity we assume that the initial state of the system is zero, the state of the system is then given by

$$x(t) = \int_0^t e^{A(t-\tau)} Bu(\tau) d\tau = \int_0^t e^{A(\tau)} Bu(t-\tau) d\tau$$

It follows from the theory of matrix functions that

$$e^{A\tau} = I\alpha_0(s) + A\alpha_1(s) + \ldots + A^{n-1}\alpha_{n-1}(s)$$

and we find that

$$x(t) = B\int_0^t \alpha_0(\tau)u(t-\tau)d\tau + AB\int_0^t \alpha_1(\tau)u(t-\tau)d\tau + \\ \ldots + A^{n-1}B\int_0^t \alpha_{n-1}(\tau)u(t-\tau)d\tau$$

The right hand is thus composed of a linear combination of the columns of the matrix.

$$W_r = \begin{pmatrix} B & AB & \ldots & A^{n-1}B \end{pmatrix}$$

To reach all points in the state space it must thus be required that there are n linear independent columns of the matrix W_c. The matrix is therefor called the reachability matrix. We illustrate by an example.

EXAMPLE 3.31—REACHABILITY OF THE INVERTED PENDULUM
The linearized model of the inverted pendulum is derived in Example 3.29. The dynamics matrix and the control matrix are

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The reachability matrix is

$$W_r = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.46}$$

This matrix has full rank and we can conclude that the system is reachable. □

Next we will consider a the system in (3.44), i.e

$$\frac{dz}{dt} = \begin{pmatrix} -a_1 & -a_2 & \ldots & a_{n-1} & -a_n \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & & 1 & 0 \end{pmatrix} z + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} u = \tilde{A}z + \tilde{B}u$$

The inverse of the reachability matrix is

$$\tilde{W}_r^{-1} = \begin{pmatrix} 1 & a_1 & a_2 & \ldots & a_n \\ 0 & 1 & a_1 1 & \ldots & a_{n-1} \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \tag{3.47}$$

To show this we consider the product

$$\begin{pmatrix} \tilde{B} & \tilde{A}\tilde{B} & \cdots & \tilde{A}^{n-1}B \end{pmatrix} W_r^{-1} = \begin{pmatrix} w_0 & w_1 & \cdots & w_{n-1} \end{pmatrix}$$

where

$$\begin{aligned} w_0 &= \tilde{B} \\ w_1 &= a_1\tilde{B} + \tilde{A}\tilde{B} \\ &\vdots \\ w_{n-1} &= a_{n-1}B + a_{n-2}\tilde{A}B + \cdots + \tilde{A}^{n-1}B \end{aligned}$$

The vectors w_k satisfy the relation

$$w_k = a_k + \tilde{w}_{k-1}$$

Iterating this relation we find that

$$\begin{pmatrix} w_0 & w_1 & \cdots & w_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

which shows that the matrix (3.47) is indeed the inverse of $\tilde{W}_r$.

Systems That are Not Reachable It is useful of have an intuitive understanding of the mechanisms that make a system unreachable. An example of such a system is given in Figure 3.32. The system consists of two identical systems with the same input. The intuition can also be demonstrated analytically. We demonstrate this by a simple example.

EXAMPLE 3.32—NON-REACHABLE SYSTEM
Assume that the systems in Figure 3.32 are of first order. The complete system is then described by

$$\begin{aligned} \frac{dx_1}{dt} &= -x_1 + u \\ \frac{dx_2}{dt} &= -x_2 + u \end{aligned}$$

The reachability matrix is

$$W_r = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

This matrix is singular and the system is not reachable. □

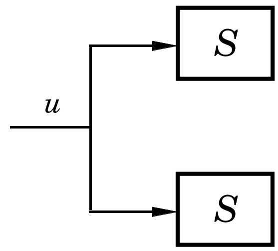

Figure 3.32 A non-reachable system.

Coordinate Changes

It is interesting to investigate how the reachability matrix transforms when the coordinates are changed. Consider the system in (3.37). Assume that the coordinates are changed to $z = Tx$. It follows from (3.41) that the dynamics matrix and the control matrix for the transformed system are

$$\begin{aligned}\tilde{A} &= TAT^{-1}\\ \tilde{B} &= TB\end{aligned}$$

The reachability matrix for the transformed system then becomes

$$\tilde{W}_r = \begin{pmatrix} \tilde{B} & \tilde{A}\tilde{B} & \dots & \tilde{A}^{n-1}\tilde{B} \end{pmatrix} =$$

We have

$$\begin{aligned}\tilde{A}\tilde{B} &= TAT^{-1}TB = TAB\\ \tilde{A}^2\tilde{B} &= (TAT^{-1})^2TB = TAT^{-1}TAT^{-1}TB = TA^2B\\ &\vdots\\ \tilde{A}^n\tilde{B} &= TA^nB\end{aligned}$$

and we find that the reachability matrix for the transformed system has the property

$$\tilde{W}_r = \begin{pmatrix} \tilde{B} & \tilde{A}\tilde{B} & \dots & \tilde{A}^{n-1}\tilde{B} \end{pmatrix} = T\begin{pmatrix} B & AB & \dots & A^{n-1}B \end{pmatrix} = TW_r$$

This formula is very useful for finding the transformation matrix T.

Observability

When discussing reachability we neglected the output and focused on the state. We will now discuss a related problem where we will neglect the input and instead focus on the output. Consider the system

$$\begin{aligned} \frac{dx}{dt} &= Ax \\ y &= Cx \end{aligned} \tag{3.48}$$

We will now investigate if it is possible to determine the state from observations of the output. This is clearly a problem of significant practical interest, because it will tell if the sensors are sufficient.

The output itself gives the projection of the state on vectors that are rows of the matrix C. The problem can clearly be solved if the matrix C is invertible. If the matrix is not invertible we can take derivatives of the output to obtain.

$$\frac{dy}{dt} = C\frac{sc}{dt} = CAx$$

From then derivative of the output we thus get the projections of the state on vectors which are rows of the matrix CA. Proceeding in this way we get

$$\begin{pmatrix} y \\ \frac{dy}{dt} \\ \frac{d^2y}{dt^2} \\ \vdots \\ \frac{d^{n-1}y}{dt^{n-1}} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} x$$

We thus find that the state can be determined if the matrix

$$W_o = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} \tag{3.49}$$

has n independent rows. Notice that because of the Cayley-Hamilton equation it is not worth while to continue and take derivatives higher than d^{n-1}/dt^{n-1}. The matrix W_o is called the observability matrix. A system is called observable if the observability matrix has full rank. We illustrate with an example.

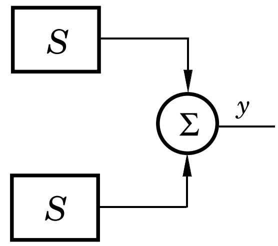

Figure 3.33 A non-observable system.

EXAMPLE 3.33—OBSERVABILITY OF THE INVERTED PENDULUM
The linearized model of inverted pendulum around the upright position is described by (3.41). The matrices A and C are

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

The observability matrix is

$$W_o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which has full rank. It is thus possible to compute the state from a measurement of the angle.

A Non-observable System

It is useful to have an understanding of the mechanisms that make a system unobservable. Such a system is shown in Figure 3.33. Next we will consider the system in (3.45) on observable canonical form, i.e.

$$\begin{aligned} \frac{dz}{dt} &= \begin{pmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & & 0 \\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & & 1 \\ -a_n & 0 & 0 & & 0 \end{pmatrix} z + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} u \\ y &= \begin{pmatrix} 1 & 0 & 0 \dots & 0 \end{pmatrix} z + Du \end{aligned}$$

A straight forward but tedious calculation shows that the inverse of the observability matrix has a simple form. It is given by

$$W_o^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_1 & 1 & 0 & \dots & 0 \\ a_2 & a_1 & 1 & \dots & 0 \\ \vdots & & & & \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & 1 \end{pmatrix}$$

This matrix is always invertible. The system is composed of two identical systems whose outputs are added. It seems intuitively clear that it is not possible to deduce the states from the output. This can also be seen formally.

Coordinate Changes

It is interesting to investigate how the observability matrix transforms when the coordinates are changed. Consider the system in (3.37). Assume that the coordinates are changed to $z = Tx$. It follows from (3.41) that the dynamics matrix and the output matrix are given by

$$\begin{aligned} \tilde{A} &= TAT^{-1} \\ \tilde{C} &= CT^{-1} \end{aligned}$$

The observability matrix for the transformed system then becomes

$$\tilde{W}_o = \begin{pmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \tilde{C}\tilde{A}^2 \\ \vdots \\ \tilde{C}\tilde{A}^{n-1} \end{pmatrix}$$

We have

$$\begin{aligned} \tilde{C}\tilde{A} &= CT^{-1}TAT^{-1} = CAT^{-1} \\ \tilde{C}\tilde{A}^2 &= CT^{-1}(TAT^{-1})^2 = CT^{-1}TAT^{-1}TAT^{-1} = CA^2T^{-1} \\ &\vdots \\ \tilde{C}\tilde{A}^n &= CA^nT^{-1} \end{aligned}$$

and we find that the observability matrix for the transformed system has the property

$$\tilde{W}_o = \begin{pmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \tilde{C}\tilde{A}^2 \\ \vdots \\ \tilde{C}\tilde{A}^{n-1} \end{pmatrix} T^{-1} = W_o T^{-1}$$

This formula is very useful for finding the transformation matrix T.

Kalman's Decomposition

The concepts of reachability and observability make it possible understand the structure of a linear system. We first observe that the reachable states form a linear subspace spanned by the columns of the reachability matrix. By introducing coordinates that span that space the equations for a linear system can be written as

$$\frac{d}{dt}\begin{pmatrix} x_c \\ x_{\bar{c}} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}\begin{pmatrix} x_c \\ x_{\bar{c}} \end{pmatrix} + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} u$$

where the states x_c are reachable and $x_{\bar{c}}$ are non-reachable. Similarly we find that the non-observable or quiet states are the null space of the observability matrix. We can thus introduce coordinates so that the system can be written as

$$\frac{d}{dt}\begin{pmatrix} x_o \\ x_{\bar{o}} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} x_o \\ x_{\bar{0}} \end{pmatrix}$$
$$y = \begin{pmatrix} C_1 & 0 \end{pmatrix}\begin{pmatrix} x_o \\ x_{\bar{o}} \end{pmatrix}$$

where the states x_o are observable and $x_{\bar{o}}$ not observable (quiet) Combining the representations we find that a linear system can be transformed to the form

$$\frac{dx}{dt} = \begin{pmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{pmatrix} x + \begin{pmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{pmatrix} u$$
$$y = \begin{pmatrix} C_1 & 0 & C_2 & 0 \end{pmatrix} x$$

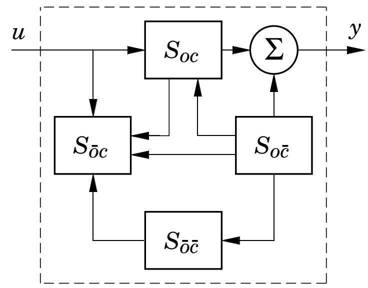

Figure 3.34 Kalman's decomposition of a system.

where the state vector has been partitioned as

$$x = \begin{pmatrix} x_{ro} \\ x_{r\bar{o}} \\ x_{\bar{r}o} \\ x_{\bar{r}\bar{o}} \end{pmatrix}^T$$

A linear system can thus be decomposed into four subsystems.

- S_{ro} reachable and observable
- $S_{r\bar{o}}$ reachable not observable
- $S_{\bar{r}o}$ not reachable observable
- $S_{\bar{r}\bar{o}}$ not reachable not observable

This decomposition is illustrated in Figure 3.34. By tracing the arrows in the diagram we find that the input influences the systems S_{oc} and $S_{\bar{o}c}$ and that the output is influenced by S_{oc} and $S_{o\bar{c}}$. The system $S_{\bar{o}\bar{c}}$ is neither connected to the input nor the output.

The transfer function of the system is

$$G(s) = C_1(sI - A_{11})^{-1}B_1 \tag{3.50}$$

It is thus uniquely given by the subsystem S_{ro}.

The Cancellation Problem Kalman's decomposition resolves one of the longstanding problems in control namely the problem of cancellation of poles and zeros. To illustrate the problem we will consider a system described by the equation.

EXAMPLE 3.34—CANCELLATION OF POLES AND ZEROS

$$\frac{dy}{dt} - y = \frac{du}{dt} - u \tag{3.51}$$

Integrating this system we find that

$$y(t) = u(t) + ce^t$$

where c is a constant. The transfer function of the system is

$$\frac{Y(s)}{U(s)} = \frac{s-1}{s-1} = 1$$

Since s is a complex variable the cancellation is clearly permissible and we find that the transfer function is $G(s) = 1$ and we have seemingly obtained a contradiction because the system is not equivalent to the system

$$y(t) = u(t)$$

The problem is easily resolved by using the Kalman representation. In this particular case the system has two subsystems S_{ro} and $S_{\bar{r}o}$. The system S_{ro} is a static system with transfer function $G(s) = 1$ and the subsystem $S_{\bar{r}o}$ which is observable but non reachable has the dynamics.

$$\frac{dx}{dt} = x$$

□

Notice that cancellations typically appear when using Laplace transforms because of the assumption that all initial values are zero. The consequences are particularly serious when factors like $s - 1$ are cancelled because they correspond to exponentially growing signals. In the early development of control cancellations were avoided by ad hoc rules forbidding cancellation of factors with zeros in the right half plane. Kalman's decomposition gives a very clear picture of what happens when poles and zeros are cancelled.

3.8 Summary

This chapter has summarized some properties of dynamical systems that are useful for control. Both input-output descriptions and state descriptions are given. Much of the terminology that is useful for control has also been introduced.