Part I

Advanced Control of continuous-time LTI-SISO Systems

Chapter 1

Review of linear systems theory

This chapter contains a short review of linear systems theory. The contents of this review roughly correspond with the materials of "Systems and Control 1". Here and there however, new concepts are introduced that are relevant for the themes that will be treated further on in this course.

1.1 Descriptions and properties of linear systems

The first step in feedback controller design is to develop a mathematical model of the process or plant to be controlled. This model should describe the relation between a given input signal $u(t)$ and the system output signal $y(t)$. In this review, we will discuss the modelling of linear, time-invariant (LTI), single-input-single-output (SISO) systems, as depicted in Fig 1.1. These models can be described by differing dynamic models. In the time domain, the physics of an LTI system can be modelled by a linear, *ordinary differential equation* (ODE) or by a *state space model*. In the frequency domain, the dynamics of the system can be described by a *transfer function* or by the *frequency response* (in the form of a Bode or a Nyquist diagram).

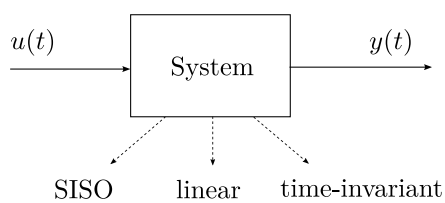


Figure 1.1: Basic setup of a LTI-SISO system.

1.1.1 Ordinary differential equations

An ordinary differential equation describes the dynamic relation between the input $u(t)$ and output $y(t)$ of a LTI-SISO system. It takes on the general form

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0y(t) = b_qu^{(q)}(t) + \cdots + b_1\dot{u}(t) + b_0u(t) \ , \qquad (1.1)$$

with the initial conditions

$$y^{(n-1)}(0) = y_{0n}, \ \cdots, \ y(0) = y_{00}, \tag{1.2}$$

$$u^{(q)}(0) = u_{0q}, \ \cdots, \ u(0) = u_{00}, \tag{1.3}$$

and with the convention

$$\frac{d^n y(t)}{dt^n} = y^{(n)}(t).$$

The value $y^{(n)}(0)$ is not required in (1.2), since it can be calculated by (1.1) when (1.2) and (1.3) are given.

Causality. In order to describe a technically realizable system, $n \geq q$ must hold. If this is not the case, and $n < q$, then $y(t)$ depends on future values of $u(t)$, which is physically impossible. This can be made more clear considering the example where $n = 0$ and $q = 1$. A certain choice of constants could then lead to the differential equation $y(t) = \dot{u}(t)$. The problem with this equation is that in order to determine the output $y(t_0)$ at a certain time t_0, a pure differentiation of the input signal at time t_0 is necessary (see Fig 1.2). The physical realization of a pure differentiation requires knowledge of the value of $u(t)$ for $t > t_0$, which is physically impossible.

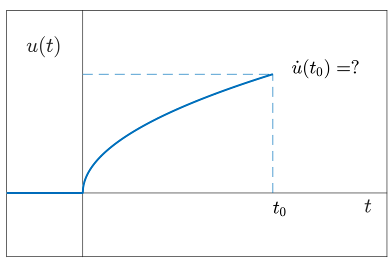


Figure 1.2: To evaluate the output $y(t_0)$ of the system $y(t) = \dot{u}(t)$, an instantaneous evaluation of $\dot{u}(t_0)$ is needed. This evaluation is not possible without any information about $u(t)$ for $t > t_0$.

Characteristic polynomial. The characteristic polynomial $p(\lambda)$ of an ODE of the form (1.1), is given by

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0 \quad . \tag{1.4}$$

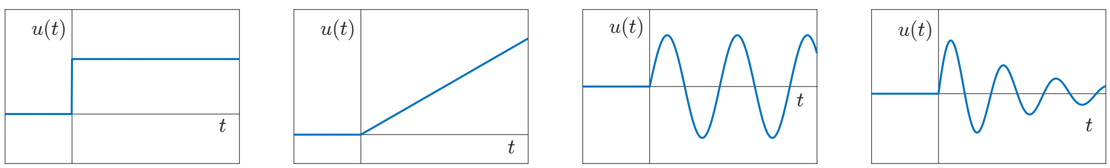
Notice that the characteristic polynomial only depends on the coefficients a_k of the derivatives $y^{(k)}(t)$ in (1.1) and not of the coefficients b_k. The characteristic polynomial can be derived by using the homogeneous solution approach $y(t) = ce^{\lambda t}$ (for $u^q(0) = 0, \cdots, u(0) = 0$ and c a constant) in the ODE. Thus, the solutions λ_i of 1.4 yield together with the approach $y(t) = ce^{\lambda t}$ a solution for the homogenous ODE. The λ_i are called also called the roots of $p(\lambda)$. These roots play an essential role in the analysis of the system dynamics, since they directly determine the *natural response* and *transient response* of the system. These partial solutions of the ODE are discussed next.

Basic ODE solution structure. The solution $y(t)$ of an ODE is the sum of three partial solutions: the natural response, the transient response and the steady state response. This can be seen by considering inputs of the type

$$u(t) = \sum_i \sum_m \bar{u}_{i,m} t^i e^{\mu_m t} \ , \tag{1.5}$$

where μ_m and $\bar{u}_{i,m}$ are real numbers or complex conjugate pairs (this ensures that u(t) is a real signal and has no imaginary part). Nearly all interesting input signals can be described by this class of inputs. In the following, some examples of inputs of this form are shown (see also Fig. 1.3). The convention is that $u(t) = 0$ for $t < 0$:

$$\begin{array}{rll} \text{Step:} & i=0,\ \mu_m = 0 & u(t) = \bar{u}_1 e^{0t} = \bar{u}_1 \sigma(t) \\ \text{Ramp:} & i=1,\ \mu_m = 0 & u(t) = \bar{u}_1 t e^{0t} = \bar{u}_1 t \\ \text{Sinusoids:} & i=0,\ \mu_m = \pm j\omega & u(t) = e^{j\omega t} + e^{-j\omega t} = 2\cos(\omega t) \\ \text{Damped Sinusoids:} & i=0,\ \mu_m = -a \pm j\omega & u(t) = e^{(-a+j\omega)t} + e^{(-a-j\omega)t} = 2e^{-at}\cos(\omega t) \end{array}$$



(a) Step input (b) Ramp input (c) Sinusoidal input (d) Damped sinusoid

Figure 1.3: Examples of inputs of the type $u(t) = \sum_i \sum_j \bar{u}_{i,m} t^i e^{\mu_m t}$.

For every input $u(t)$ of the type above, the output $y(t)$ can be computed analytically. It can be shown that the output consists of a natural response $y_{\text{nat}}(t)$ that appears due to the initial value of the system, and of a forced response $y_{\text{forced}}(t)$ that is caused by the input signal. The forced response itself consists of the transient response $y_{\text{trans}}(t)$ and the steady state response $y_{\text{ss}}(t)$.

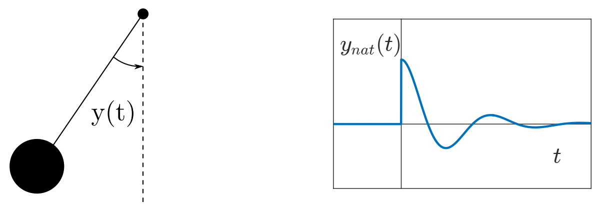
$$y(t) = \underbrace{y_{\text{nat}}(t)}_{\substack{\text{natural response} \\ \text{(depending on initial conditions} \\ \text{but not on system input)}}} + \underbrace{y_{\text{forced}}(t)}_{\substack{\text{forced response} \\ \text{(depending on input} \\ \text{but not on initial conditions)}}} \tag{1.6}$$

$$= y_{\text{nat}}(t) + y_{\text{trans}}(t) + y_{\text{ss}}(t) \tag{1.7}$$

$$= \sum_i \bar{y}_{\text{nat}_i} e^{\lambda_i t} + \sum_i \bar{y}_{\text{trans}_i} e^{\lambda_i t} + \sum_i \sum_m \bar{y}_{\text{ss}_{i,m}} t^i e^{\mu_m t}. \tag{1.8}$$

It can be shown that the coefficients $\bar{y}_{\text{nat}_i}$, $\bar{y}_{\text{trans}_i}$ and $\bar{y}_{\text{ss}_{i,m}}$ do not depend on time. Hence, the time behavior is only determined by the terms $e^{\lambda_i t}$ (with λ_i are the solutions of the characteristic polynomial) and the terms $t^i e^{\mu_m t}$ of the input. An important consequence is that if the real part of all λ_i is negative, then the natural response and the transient response will go to zero for $t \to \infty$. The system output for $t \to \infty$ is then solely represented by the steady state response. However, if at

least one λ_i has a positive real part, the natural and transient response will not die out and will grow unbounded.
In addition, it is interesting to note that the coefficients $\bar{y}_{\mathrm{nat}_i}$, $\bar{y}_{\mathrm{trans}_i}$ and $\bar{y}_{\mathrm{ss}_{i,m}}$ can be interpreted as the intensity or the "amplitude" of the time-dependent exponential terms. It turns out that $\bar{y}_{\mathrm{trans}_i}$ depends on the input parameters $\bar{u}_{i,m}$ and μ_m but not on the initial conditions (1.2) and (1.3). Hence, the transient behavior is triggered by the input, but evolves similarly to the natural response.
The example of a (linearised) pendulum system can be used to illustrate the different components of the output response $y(t)$. In this example the output is defined as the angle between the pendulum and the vertical axis, as shown in Fig. 1.4a. Suppose now that the pendulum has an initial non-zero displacement from its equilibrium position. Apart from gravity, no external forces are present. When released from its initial position, the pendulum will start an oscillating motion, that will fade out if any friction is present, until equilibrium is reached. This oscillating motion is described by the natural response $y_{\mathrm{nat}}(t)$ (see Fig 1.4b) and is solely a function of the initial displacement and of the roots λ_i of the characteristic polynomial.



(a) Non-zero initial state (b) Natural response

Figure 1.4: Natural response of a pendulum system.

Suppose now that the initial displacement of the pendulum is zero (hence the natural response is zero), but that it is placed in a homogeneous wind field, that exerts an external force $u(t)$ on the pendulum, as shown in Fig. 1.5a. The pendulum will be displaced from its equilibrium, following an oscillating motion, until the new equilibrium point is reached. This motion is described by the forced response $y_{\mathrm{forced}}(t)$, shown in Fig. 1.5b. The transient response $y_{\mathrm{trans}}(t)$ describes the oscillating motion around the new equilibrium point, and the steady state response $y(t)$ describes precisely this equilibrium point.

Linearization. The state equations of a physical systems are in general nonlinear. Many real systems are not described by a linear ODE of the form (1.1), but rather by a nonlinear ODE of the general form

$$f(y^{(n)}(t), ..., \dot{y}(t), y(t), u^{(q)}, ..., \dot{u}(t), u(t)) = 0 \quad . \tag{1.9}$$

An example of a nonlinear system is a water tank, with an input flow rate $u(t)$ (e.g. in m^3 per sec.), that has an outlet at the bottom (see Fig. 1.6). The outlet flow rate is a function of the pressure at the bottom of the tank, which itself is a function of the water level $h(t)$. The dynamics of the system, where the output $y(t)$ considered is the water level $h(t)$, are described by the nonlinear ODE

$$f(\dot{y}(t), y(t), u(t)) = \dot{y}(t) + \frac{k}{A}\sqrt{2gy(t)} - \frac{1}{A}u(t) = 0 \quad , \tag{1.10}$$

where k is an outlet characteristic and A is the cross section area of the tank. In order to design a

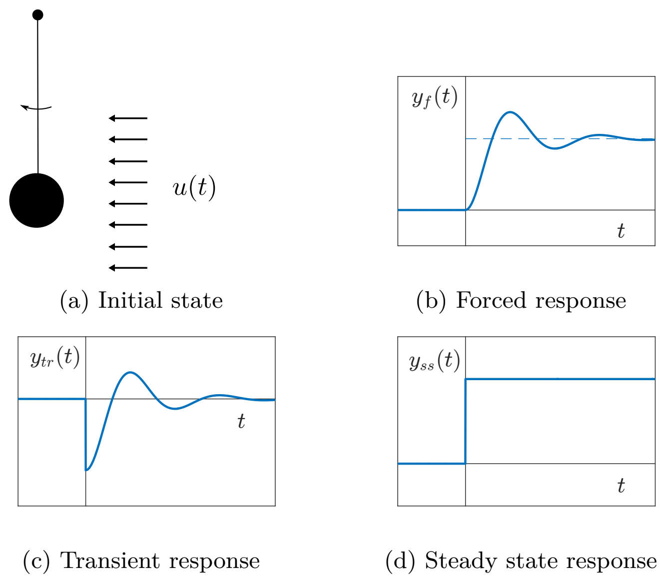


Figure 1.5: Forced response of a pendulum system in a homogeneous wind field. The forced response y_{forced} is the sum of the transient response y_{trans} and the steady state response y_{ss}.

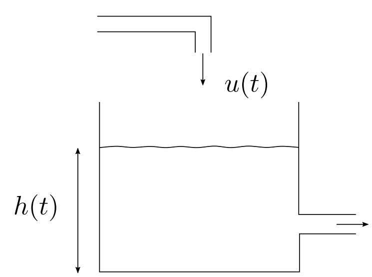


Figure 1.6: Setup of the water tank.

controller to regulate the water level in the tank, a more complex analysis of the system behavior (in terms of its solution structure, natural response,...) than the one discussed above is required.

Therefore it is usual to simplify the nonlinear behavior of such a system, and use linear models in the design of the controller. This is especially the case for regulation control systems, where the goal of the controller is to keep the system at a given equilibrium point in the presence of disturbances. For this equilibrium point $(u_{\text{ss}}, y_{\text{ss}})$, it holds that

$$f(0, ..., 0, y_{\text{ss}}, 0, ..., u_{\text{ss}}) = 0 \quad . \tag{1.11}$$

If the controller does its work properly, the system will stay near its equilibrium point. Therefore, the system model is only required to describe the dynamic behavior around this point. Such a system model is then obtained by linearizing the nonlinear model via a first order Taylor approximation around the equilibrium point:

$$\bar{a}_n \Delta y^{(n)}(t) + ... + \bar{a}_1 \Delta \dot{y}(t) + \bar{a}_0 \Delta y(t) = \bar{b}_q \Delta u^{(q)}(t) + ... + \bar{b}_0 \Delta u(t) \quad , \tag{1.12}$$

where

$$\bar{a}_i = \quad \frac{\partial f}{\partial y^{(i)}} \left(y^{(n)}(t), ..., y(t), u^{(q)}(t), ..., u(t) \right) \Big|_{u_{\text{ss}}, y_{\text{ss}}} \tag{1.13}$$

$$\bar{b}_j = -\frac{\partial f}{\partial u^{(j)}} \left(y^{(n)}(t), ..., y(t), u^{(q)}(t), ..., u(t) \right) \Big|_{u_{\text{ss}}, y_{\text{ss}}} \quad , \tag{1.14}$$

and where

$$\Delta y(t) = y(t) - y_{\text{ss}} \quad \text{and} \quad \Delta u(t) = u(t) - u_{\text{ss}} \quad . \tag{1.15}$$

In order to get the linearized ODE in the canonical form (1.1), divide (1.12) by $\bar{a}_n$. In the case of the water tank, the linearization is done around a certain equilibrium point ($y_{\text{ss}} = h_{\text{ref}}$, $u_{\text{ss}} = k\sqrt{2gy_{\text{ss}}} = k\sqrt{2gh_{\text{ref}}}$), that needs to be maintained. Applying (1.13) and (1.14) then gives

$$\bar{a}_1 = 1 \quad , \tag{1.16}$$

$$\bar{a}_0 = \frac{\mathrm{d}}{\mathrm{d}y} \frac{k}{A} \sqrt{2gy(t)} \Big|_{y_{\text{ss}}, u_{\text{ss}}} = \frac{k}{A} \sqrt{\frac{g}{2h_{\text{ref}}}} \quad , \tag{1.17}$$

$$\bar{b}_0 = \frac{1}{A} \quad . \tag{1.18}$$

The linearized system model of the water tank around the reference height h_{ref} is then

$$\Delta \dot{y}(t) + \frac{k}{A} \sqrt{\frac{g}{2h_{\text{ref}}}} \Delta y(t) = \frac{1}{A} \Delta u(t) \quad , \tag{1.19}$$

which is a linear ODE of the form (1.1), so that all analyses discussed above as well as below can be applied to it.

1.1.2 State space models

The use of the general ODE in controller design is rather inconvenient. However, it forms the basis for other, mathematically equivalent, model types that are more frequently used. The state space model is preferred to the general ODE model because with the 'system states', it introduces variables that are easily interpretable from an engineering perspective. The system states are also an important element in controller design. Furthermore, the state space model has a convenient form for computational processing.

The state space model can be derived by transforming the general ODE, that is of degree n, into n ODE's of degree 1. It is expressed in the form

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}\mathbf{x}(t) + \mathbf{d}u(t) \ , \end{aligned} \tag{1.20}$$

with

$$\mathbf{x}(0) = \mathbf{x_0}. \tag{1.21}$$

The states of the system are contained in $\mathbf{x}(\mathrm{t})$, which is a vector with n elements, also for SISO systems. One of the advantages of this model is that it easily extends to multiple-inputs-multiple-outputs (MIMO) systems. This extension will be discussed in detail in Part 3. The matrix $\mathbf{A}$ and the vectors $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ of the state space model can easily be deduced from the ODE model (with $a_n = 1$) using the control canonical form given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{1.22}$$

and

$$\mathbf{c} = \begin{bmatrix} b_0 - b_n a_0 & b_1 - b_n a_1 & \cdots & b_{n-1} - b_n a_{n-1} \end{bmatrix} \quad \text{and} \quad \mathbf{d} = b_n \quad . \tag{1.23}$$

Characteristic polynomial. The characteristic polynomial of the system is obtained from the state space model by

$$p(\lambda) = \det(\lambda I - A) = 0. \tag{1.24}$$

Thus the solutions of $p(\lambda) = 0$ are the *eigenvalues* of A. Therefore, if $\Re\{\mathrm{eig}(A)\} < 0$, then $\lim_{t\to\infty} y_{\mathrm{nat}}(t) \;=\; 0$ and $\lim_{t\to\infty} y_{\mathrm{trans}}(t) \;=\; 0$.

Time solution. Solving the state space ODE (1.20) for $x(t)$ and $y(t)$ yields

$$\mathbf{x}(t) = \underbrace{e^{\mathbf{A}t}\mathbf{x}_0}_{\substack{\text{natural}\\ \text{response}}} + \underbrace{\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{b}u(\tau)d\tau}_{\text{forced response}} \tag{1.25}$$

$$y(t) = \underbrace{\mathbf{c}e^{\mathbf{A}t}\mathbf{x}_0}_{\substack{\text{natural}\\ \text{response}}} + \underbrace{\int_0^t \mathbf{c}e^{\mathbf{A}(t-\tau)}\mathbf{b}u(\tau)d\tau + \mathbf{d}u(t)}_{\text{forced response}} \quad . \tag{1.26}$$

The structure of the solution is the same as the one of the general ODE, since the models are equivalent. Hence, the output and the state trajectories can also be separated into a natural response and a forced response. The Matrix Exponential $e^{\mathbf{A}t}$ in (1.25) is defined as

$$e^{\mathbf{A}t} = I + \mathbf{A}t + \frac{\mathbf{A}^2}{2!}t^2 + \frac{\mathbf{A}^3}{3!}t^3 + \cdots .$$

It has the same property as the usual exponential ($\frac{\mathrm{d}e^{at}}{\mathrm{d}t} = ae^{at}$):

$$\frac{\mathrm{d}e^{\mathbf{A}t}}{\mathrm{d}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A} \quad .$$

The impulse response $g(t)$ for an input $u(t) = \delta(t)$ ($\mathbf{x}_0 = 0$), can be calculated by (1.25) and results in

$$g(t) = \mathbf{c}e^{\mathbf{A}t}\mathbf{b} + \mathbf{d}\delta(t) \quad . \tag{1.27}$$

Similarly, the step response $h(t)$ for an input $u(t) = \sigma(t)$ ($\mathbf{x}_0 = 0$), can be found by

$$h(t) = \int_0^t \mathbf{c}e^{\mathbf{A}t}\mathbf{b} + \mathbf{d}\sigma(t) \quad . \tag{1.28}$$

Remark. $\dot{h}(t) = g(t)$.

Linearization. When the system contains nonlinear dynamics, the state space model takes on a more general, nonlinear form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), u(t)) \tag{1.29}$$

$$y(t) = g(\mathbf{x}(t), u(t)) \quad , \tag{1.30}$$

where

$$\mathbf{x}(0) = \mathbf{x}_0 \quad . \tag{1.31}$$

The function $\mathbf{f}$ is vector-valued in general. Therefore, (1.29) is a simplified expression for

$$\begin{bmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_i(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} f_1(x_1, ..., x_n, u) \\ \vdots \\ f_i(x_1, ..., x_n, u) \\ \vdots \\ f_n(x_1, ..., x_n, u) \end{bmatrix} \quad . \tag{1.32}$$

The linearization of the general ODE model, as discussed in section 1.1.1, can also be directly performed on the nonlinear state space model. Likewise, the linearization is applied around an equilibrium point $(\mathbf{x}_{ss}, u_{ss})$, and results in a linearized state space model with respect to the deviations

$$\Delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}_{ss} \tag{1.33}$$

$$\Delta u(t) = u(t) - u_{ss} \tag{1.34}$$

from the equilibrium point. The linearized model is given by

$$\Delta\dot{\mathbf{x}}(t) = \underbrace{\left.\frac{\partial \mathbf{f}(\mathbf{x}, u)}{\partial \mathbf{x}}\right|_{x_{ss}, u_{ss}}}_{\mathbf{A}} \Delta\mathbf{x}(t) + \underbrace{\left.\frac{\partial \mathbf{f}(\mathbf{x}, u)}{\partial u}\right|_{x_{ss}, u_{ss}}}_{\mathbf{b}} \Delta u(t) \;=\; \mathbf{A}\Delta\mathbf{x}(t) + \mathbf{b}\Delta u(t) \tag{1.35}$$

$$\Delta y(t) = \underbrace{\left.\frac{\partial g(\mathbf{x}, u)}{\partial \mathbf{x}}\right|_{x_{ss}, u_{ss}}}_{\mathbf{c}} \Delta\mathbf{x}(t) + \underbrace{\left.\frac{\partial g(\mathbf{x}, u)}{\partial u}\right|_{x_{ss}, u_{ss}}}_{\mathbf{d}} \Delta u(t) \;=\; \mathbf{c}\Delta\mathbf{x}(t) + \mathbf{d}\Delta u(t) \quad , \tag{1.36}$$

where the partial derivatives of the state function $\mathbf{f}(t)$ and output function $g(t)$ are Jacobi matrices, e.g.

$$\left.\frac{\partial \mathbf{f}(\mathbf{x}, u)}{\partial \mathbf{x}}\right|_{x_{ss}, u_{ss}} = \left.\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}\right|_{x_{ss}, u_{ss}} , \tag{1.37}$$

For the equilibrium point it holds that $\mathbf{f}(\mathbf{x}_{ss}, u_{ss}) = 0$. The linearized state space model of the form (1.35) and (1.36) is mathematically equivalent to the linearized general ODE of the form (1.12).

1.1.3 Transfer function

To analyze in a straightforward way the dynamic behavior of complex LTI-SISO systems (e.g. systems connected in series or systems in a feedback loop), it is also useful to work with models in the *frequency domain*. Such a model can be obtained based on the time-domain model by the *Laplace transform*. The Laplace transform of a function $f(t)$ to a function $F(s)$ (with $s = \alpha + j\omega$) is defined as

Definition 1. $\mathcal{L}\{f(t)\} = \int_0^\infty f(t)e^{-st}dt = F(s)$.

The relation between a time-signal and its Laplace transform is also often symbolized by ○—●, with the full circle on the side of the Laplace transformed signal. We thus write

$$f(t) \quad ○\!\!-\!\!\!-\!\!● \quad F(s)$$

in order to express that $f(t)$ and $F(s)$ are a pair. Some important Laplace transforms are:

Step:	$\sigma(t)$	○—●	$\frac{1}{s}$
Ramp:	t	○—●	$\frac{1}{s^2}$
Dirac impulse:	$\delta(t)$	○—●	1

The inverse Laplace transform is defined as

$$\mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi j} \int_{\alpha-\infty}^{\alpha+\infty} F(s)e^{st}ds = f(t) \quad . \tag{1.38}$$

Relevance. The transfer function of a system can be defined as the Laplace transform of the impulse response $g(t)$, under the presupposition that all initial conditions equal zero:

Definition 2. $G(s) = \mathcal{L}\{g(t)\}$ *with* $\mathbf{x}_0 = 0$.

Why is this transfer function model of the plant so useful? There are at least two important reasons:

1. Combining (1.25) and (1.27), for $\mathbf{x}_0 = 0$ it can be seen that

$$y(t) = y_{\text{forced}}(t) = \int_0^t g(t-\tau)u(\tau)d\tau \quad . \tag{1.39}$$

   This expression is the *convolution* of the impulse response $g(t)$ with the input signal $u(t)$. The Laplace transform of (1.39) results in

$$Y_{\text{forced}}(s) = G(s) \cdot U(s) \quad . \tag{1.40}$$

   It so turns out that the transfer function model is very useful to calculate the forced system output, which is the sum of the transient and the steady state response: $y_{\text{forced}}(t) = y_{\text{trans}}(t) + y_{\text{ss}}(t)$. This is done in three steps:

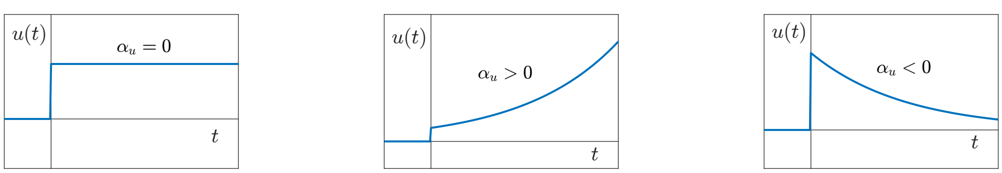
   (a) Transform the input signal $u(t)$ to its representation $U(s)$ in the frequency domain.

   (b) Use $G(s)$ to determine the output signal in the frequency domain via $Y(s) = G(s)U(s)$.

   (c) Transform $Y(s)$ into its time-domain representation via the inverse Laplace transform.

   In certain circumstances, this procedure is easier than solving the state equation or to calculate the convolution of the impulse response with the input signal.

2. For *exponential inputs*, the transfer function $G(s)$ can also be used to calculate the steady state output $y_{\text{ss}}(t)$ in a direct manner. An exponential input takes the form $u(t) = \bar{u}e^{s_{\text{u}}t} = \bar{u}e^{(\alpha_{\text{u}}+j\omega_{\text{u}})t}$. Fig. 1.7 and Fig. 1.8 show some examples of exponential inputs. It can be shown that the steady



(a) (b) (c)

Figure 1.7: Examples of exponential inputs of the type $u(t) = e^{\alpha_{\text{u}}t}$.

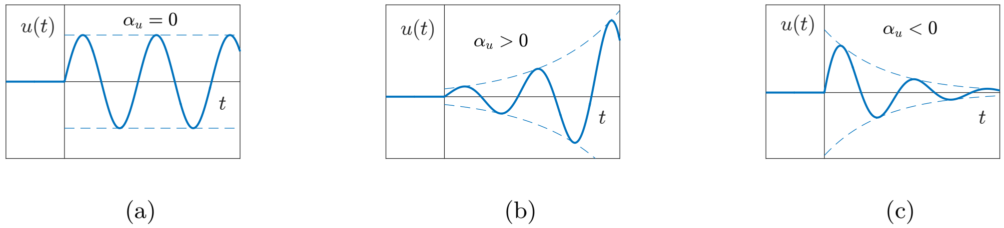


Figure 1.8: Examples of exponential inputs of the type $u(t) = e^{\alpha_u \pm j\omega}$.

state response $y_{ss}(t)$ for an exponential input $u(t)$ is given by

$$\begin{aligned} y_{ss}(t) &= G(s_u)u(t) && (1.41)\\ &= G(\alpha_u + j\omega_u) \cdot \bar{u}e^{(\alpha_u + j\omega_u)t} && (1.42)\\ &= \underbrace{|G(\alpha_u + j\omega_u)|}_{\text{gain}} \cdot \underbrace{e^{j\phi(\alpha_u + j\omega_u)}}_{\text{phase shift}} \cdot \bar{u}e^{(\alpha_u + j\omega_u)t} \quad . && (1.43) \end{aligned}$$

Interestingly, (1.41) - (1.43) describe the relation of the *time signals* $y_{ss}(t)$ and $u(t)$ in terms of a complex-valued function that is defined in the *frequency domain*. In particular, the relation is given by the transfer function $G(s)$, evaluated at the complex frequency s_u of the input, i.e., by $G(s_u)$. The value $G(s_u)$ is a complex number. The magnitude of this number represents the steady state 'gain' with which an exponential input with the complex frequency $s = \alpha_u + j\omega_u$ appears in the steady state output. The argument $\arg G(\alpha_u + j\omega_u) = \phi(\alpha_u + j\omega_u)$ (i.e. the angle the complex number makes with the x-axis in the complex plane) describes the respective phase shift that such an input undergoes.

Calculation. The transfer function can be calculated as the Laplace transform of the impulse response of the system. However, it can also be determined by departing from the state space model:

$$G(s) = \mathbf{c}(sI - \mathbf{A})^{-1}\mathbf{b} + \mathbf{d} \quad . \tag{1.44}$$

Alternatively the transfer function can also be worked out from the parameters $a_n, \ldots, a_0$ and $b_q, \ldots, b_0$ of the general ODE-description of the plant:

$$\begin{aligned} G(s) &= \frac{b_q s^q + b_{q-1}s^{q-1} + \cdots + b_1 s + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} && (1.45)\\ &= k\frac{\prod_i^q (s - s_{0i})}{\prod_j^n (s - s_i)} && (1.46) \end{aligned}$$

The roots s_{0i} of the numerator of $G(s)$ are called the *zeros* of the transfer function. The roots s_i of the denominator are called the *poles* of the transfer function.

Remark. *In case there is no cancellation of a pole with a zero, the denominator of the transfer function equals the characteristic polynomial. Hence, the poles of the transfer function are equal to the eigenvalues of any state space model of the same system. However, if there is a pole-zero cancellation,* *then the denominator of the transfer function is not equal to the characteristic equation anymore and there are less poles than eigenvalues. This phenomenon occurs when the state space model is not observable or controllable. These terms will be defined and discussed in Part 3.*

Remark. *Due to the causality criterion ($n \geq q$), the transfer function must at least have as many poles as zeros, in order to be realizable in practice.*

Definition 3. *The relative degree Δr of the transfer function is $n - q$.*

Definition 4. *$G(s)$ is proper if $n \geq q$.*

Definition 5. *$G(s)$ is strictly proper if $n > q$.*

Definition 6. *$G(s)$ is biproper if $n = q$.*

Minimum-phase systems and All-passes. A rational transfer function $G(s)$ is said to be *minimum-phase* if all zeros lie in the left half of the complex plane (LHP). Consequently, $G(s)$ is non-minimum-phase if at least one zero lies in the right half plane (RHP). The concept is relevant since non-minimum-phase systems are more difficult to control than minimum-phase systems and therefore demand special considerations.

Remark. *The term 'minimum-phase' is grounded in the the fact that it is possible to directly construct the Bode phase plot of a minimum-phase system from its Bode magnitude plot: a gradient of $z \cdot 20dB/decade$ is accompanied by a phase of $z \cdot 90°$, where z is a random integer.*

Further on, $G(s)$ is said to be an *all-pass* system if $|G(s)| = 1$ for all frequencies and $G(s)$ is non-minimum-phase, such as e.g.

$$G(s) = \frac{-Ts + 1}{Ts + 1} \quad . \tag{1.47}$$

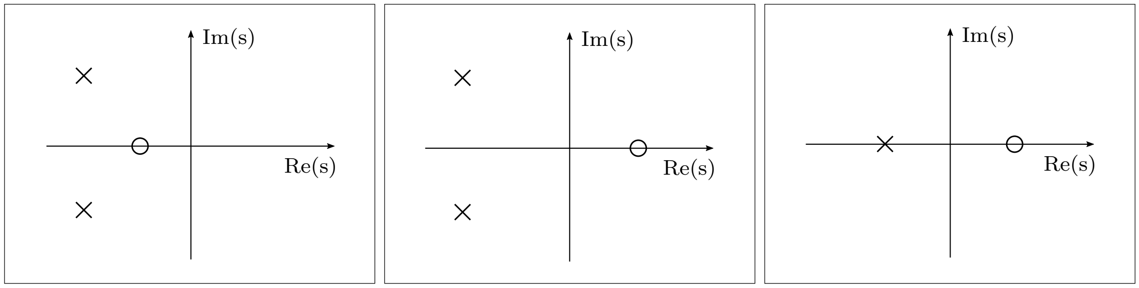
All-pass systems are therefore systems that amplify all frequencies equally and thus only affect the phase of sinusoidal signals. All-passes are characterized by the fact that for every all-pass zero $s_{0i} = \alpha_i + j\omega_i$, it also has a pole with the opposite real part $s_i = -\alpha_i + j\omega_i$. It is easy to see that every non-minimum-phase system can be separated into a minimum-phase system and an all-pass system. Fig. 1.9 gives examples of the Pole-Zero diagram for a minimum-phase, a non-minimum phase and an all-pass system. A circle represents a zero and a cross signifies a pole.

Finite value theorem. A property of the Laplace transform that is very useful in control is known as the 'Finite Value Theorem'.

Theorem 1. *The finite value theorem holds that*

$$\lim_{t \to \infty} f(t) = \lim_{s \to 0} s \cdot F(s) \tag{1.48}$$

under the condition that both limits exist. This is the case if the Laplace transform $F(s)$ is stable, i.e., only has poles with negative real part.



(a) Minimum-phase (b) Non-minimum-phase (c) All-pass

Figure 1.9: Examples of a Pole-Zero diagram of a (non)-minimum-phase and an all-pass system.

The finite value theorem allows us to compute the steady state impulse response $g_{ss}(t) = \lim_{t\to\infty} g(t)$ of a stable system, based on its transfer function $G(s)$ in the Laplace domain:

$$\lim_{t\to\infty} g(t) = \lim_{s\to 0} s \cdot G(s) \quad . \tag{1.49}$$

For example, the steady state impulse response of the system $G(s) = \frac{1}{Ts+1}$ can be calculated as

$$\lim_{t\to\infty} g(t) = \lim_{s\to 0} \frac{s}{Ts+1} = 0 \quad . \tag{1.50}$$

The theorem can also be used to calculate the steady state step response of a stable system:

$$\lim_{t\to\infty} h(t) = \lim_{s\to 0} s \cdot \underbrace{G(s) \cdot \frac{1}{s}}_{H(s)} = \lim_{s\to 0} G(s) \quad . \tag{1.51}$$

For example, the steady state step response of the system $G(s) = \frac{V}{Ts+1}$ can then be calculated as:

$$h_{ss}(t) = \lim_{t\to\infty} h(t) = \lim_{s\to 0} \frac{V}{Ts+1} = V \quad . \tag{1.52}$$

1.1.4 Frequency response

The frequency response describes the steady state transfer function of the system for a sinusoidal input.

Definition 7. *The frequency response $G(j\omega)$ of a system is the transfer function $G(s)$ of that system with $s = j\omega$, i.e. for s only on the imaginary axis.*

Consider exponential inputs with arguments that strictly lie on the imaginary axis: $u(t) = e^{j\omega t} + e^{-j\omega t}$. This is the general form of pure sinusoids or of a DC-signal (if $\omega = 0$). Using the relationship (1.41), the frequency response is extremely useful to study the steady state output for sinusoidal inputs:

$$y_{ss}(t) = G(j\omega)u(t) \tag{1.53}$$

$$= |G(jw)|e^{j\cdot\arg(G(j\omega))}u(t) \ . \tag{1.54}$$

The effect of $G(j\omega)$ on sinusoidal inputs is twofold. The system causes an amplification of the sinusoidal input, as well as a phase shift. Both effects can be visualized in a *Bode plot*, which contains an magnitude plot and a phase plot, as shown in Fig 1.10. Because the frequency span of interest stretches across multiple decimal powers, and because the gain of a system can vary many orders of magnitude, the Bode plot makes use of a log-log-scale. The frequency is logarithmically scaled with $\log \omega$, whereas the gain is scaled to dB by

$$|G|_{\mathrm{dB}} = 20 \log |G| \quad .$$

The phase plot is the representation of the phase ϕ as a function of the logarithm of the frequency (ϕ-$\log \omega$-plot). Another way to visualize the frequency response of a system is by using a *Nyquist plot*,

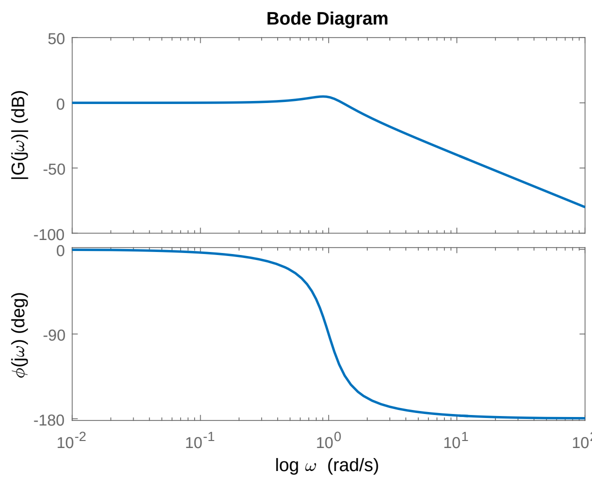


Figure 1.10: Bode diagram for a second order system.

as shown in Fig. 1.11. This plot shows both magnitude and phase shift in one plot. The diagram is constructed by drawing the complex value of the function $G(j\omega)$ as a curve in the complex plane for $\omega \in [0, \infty)$. The Nyquist curve for the frequency span $\omega \in (-\infty...0)$ can be obtained through a reflection of the Nyquist curve for positive frequencies about the real axis. The Nyquist curve starts off on the real axis, at the static gain $G(0) = \frac{b_0}{a_0} = k_s$. For a strictly proper system, it approximates zero for high frequencies. The Nyquist curve of a biproper system, however, approximates the point d on the real axis for $\omega \to \infty$. This point is determined by the control ratio $\frac{b_n}{a_n}$, since $\lim_{\omega \to \infty} G(j\omega) = \frac{b_n}{a_n} = d$, when $q = n$.

1.2 Stability

The general idea of stability entails that a system is stable if a bounded excitation leads to a bounded movement of the system. Starting from two different perspectives, this very general concept of stability can be pinned down to the following two exact definitions.

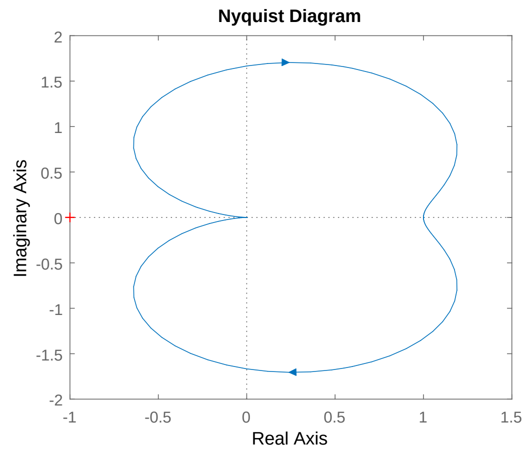


Figure 1.11: Nyquist diagram for a second order system

1. **BIBO-stability** (Bounded-Input-Bounded-Output-stability). Starting from the equilibrium state $\mathbf{x}_0 = 0$, the system is excited by a bounded *external signal* $u(t)$. The system is stable when the output of the system is also bounded. Since the system start at $x_0 = 0$, the output of the system only consists of the the forced response. Hence, BIBO-stability describes stability in terms of the forced response.

2. **Lyapunov stability**. The excitation is an initial *displacement* $\mathbf{x}_0$ from an equilibrium state, in the absence of an external input ($u(t) = 0$). The system is said to be stable when the system state stays close to, or even returns to the equilibrium state. Since the system has no input but starts with an $\mathbf{x}_0 \neq 0$, the system response only consists of the natural response. Therefore, Lyapunov stability describes stability in terms of the natural response.

1.2.1 BIBO-stability

BIBO-stability characterizes the behavior of the system due to a bounded input, starting from an zero initial state ($\mathbf{x}_0 = 0$) (see Fig. 1.12). A system is BIBO-stable if it reacts to every arbitrary bounded input $u(t)$ with a bounded output signal $y(t)$.

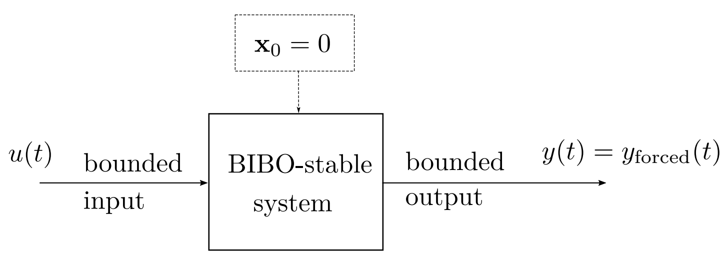


Figure 1.12: Setup of a BIBO-stable system.

Definition 8. *A system is BIBO-stable if (with $\boldsymbol{x}_0 = 0$) it holds that*

$$|u(t)| < u_{max} \implies |y(t)| < y_{max} \quad \forall \quad t > 0. \tag{1.55}$$

It can be shown that this condition is equivalent to the condition

$$\int_0^\infty |g(t)|dt < \infty \quad . \tag{1.56}$$

A system is thus BIBO-stability when the integral of its impulse response $g(t)$ exists for all t. This condition implies that $g(t) \to 0$ for $t \to \infty$. Therefore, the system is BIBO-stable if and only if all poles of the system are located in the LHP (i.e. $\Re\{s_i\} < 0 \; \forall i = 1, ..., n$).

1.2.2 Lyapunov stability

Lyapunov stability is a very versatile concept, that easily extends to nonlinear systems. It characterizes the behavior of the system state due to a displacement of the initial state $\mathbf{x}_0$ from an equilibrium point $\mathbf{x}_{\text{ss}}$, as shown in Fig. 1.13. After defining the notion of an equilibrium point, the concept of Lyapunov stability is further defined below. The different definitions are illustrated in Fig. 1.14.

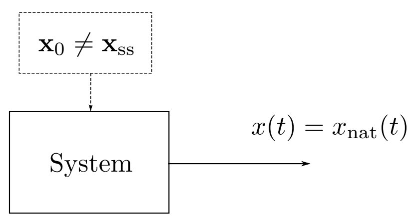


Figure 1.13: Setup of a Lyapunov stable system.

Definition 9. $\mathbf{x}_{\text{ss}}$ *is an* ***equilibrium point*** *if the system stays in* $\mathbf{x}_{\text{ss}}$ *for ever once* $\mathbf{x}_{\text{ss}}$ *is reached. In the general (nonlinear) case, for a state equation* $\dot{\mathbf{x}}(t) = f(\mathbf{x}(t))$, *the condition for an equilibrium point boils down to* $f(\mathbf{x}_{\text{ss}}) = 0$. *In the linear case, when the system is described by the state equation* $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$, *the condition for equilibrium translates to* $\mathbf{A}\mathbf{x}_{\text{ss}} = 0$. *Hence, a linear system thus has exactly one equilibrium point* ($\mathbf{x}_{\text{ss}} = 0$), *or an infinite number of equilibrium points (e.g. a single integrator system:* $\mathbf{A} = 0$, $\mathbf{x}_{\text{ss}} \in \mathbb{R}$*).*

Definition 10. *An equilibrium point* $\mathbf{x}_{\text{ss}}$ *is* ***Lyapunov stable*** *if for any* $R > 0$, *there exists a* $r > 0$ *such that*

$$\|\mathbf{x}_0 - \mathbf{x}_{\text{ss}}\| < r \implies \|\mathbf{x}(t) - \mathbf{x}_{\text{ss}}\| < R \quad \forall \quad t \geq 0 \quad . \tag{1.57}$$

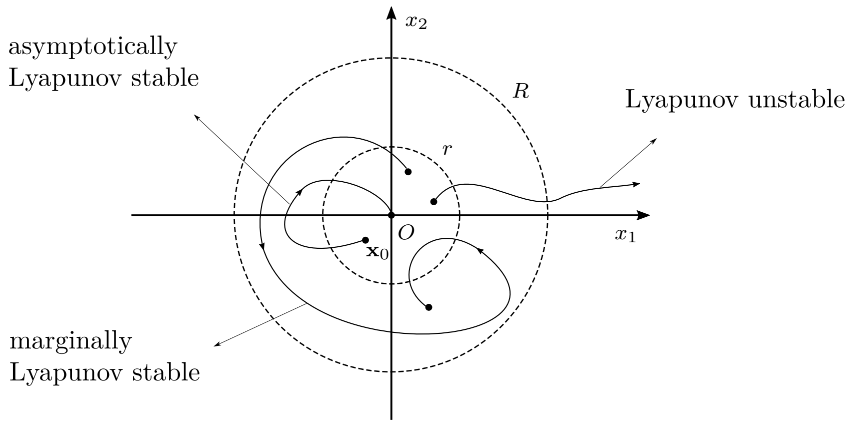


Figure 1.14: Trajectories of Lyapunov stable and unstable systems.

Definition 11. *An equilibrium point* $\mathbf{x}_{ss}$ *is* ***asymptotically Lyapunov stable*** *if* $\mathbf{x}_{ss}$ *is Lyapunov stable and there exists an* $r > 0$ *such that*

$$\|\mathbf{x}_0 - \mathbf{x}_{ss}\| < r \implies \lim_{t\to\infty} |\mathbf{x}(t)| = \mathbf{x}_{ss} \quad . \tag{1.58}$$

Definition 12. *An equilibrium point* $\mathbf{x}_{ss}$ *is* ***marginally Lyapunov stable*** *if* $\mathbf{x}_{ss}$ *is Lyapunov stable but not asymptotically Lyapunov stable.*

Definition 13. *A* ***system*** *is called (asymptotically) Lyapunov stable if all of its equilibrium points are (asymptotically) Lyapunov stable. A linear system is called marginally stable if all of its equilibrium points are Lyapunov stable and at least one equilibrium point is marginally stable.*

Conditions for Lyapunov stability. Given these definitions, we can investigate when a linear system is asymptotically Lyapunov stable, marginally Lyapunov stable or Lypapunov unstable:

- A linear system is asymptotically Lyapunov stable if and only if all eigenvalues of **A** lie strictly in the LHP, i.e., if

$$\Re\{\lambda_i\} < 0 \text{ for all } i = 1, \ldots, n \; . \tag{1.59}$$

- A linear system is Lyapunov unstable if at least one eigenvalue of the system matrix lies in the RHP, i.e., if

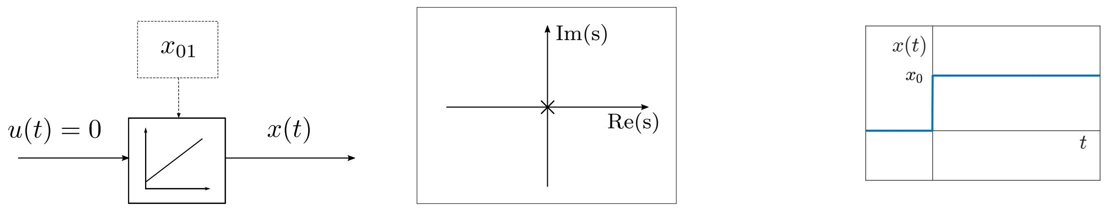
$$\Re\{\lambda_i\} > 0 \quad \text{for one or more} \quad i = 1, \ldots, n \quad . \tag{1.60}$$

- A linear system is marginally Lyapunov stable if it has no eigenvalues in the RHP and only non-repeated eigenvalues on the imaginary axis, i.e., if

$$\Re\{\lambda_i\} \leq 0 \quad \text{for all} \quad i = 1, \ldots, n \tag{1.61}$$

$$\text{and} \quad \lambda_i \neq \lambda_j \quad \text{if} \quad \Re\{\lambda_i\} = \Re\{\lambda_j\} = 0 \quad \forall \quad i, j = 1, \ldots, n \text{ with } i \neq j. \tag{1.62}$$

  - An example of a marginally stable system is an integrator (state equations: $\dot{\mathbf{x}}(t) = \mathbf{0}\mathbf{x}(t)$, transfer function: $G(s) = \frac{1}{s}$, initial state: $\mathbf{x}(0) = \mathbf{x}_0$). As the system matrix $\mathbf{A}$ of a single integrator system is zero, every initial state will immediately also be the new equilibrium point The system state will not grow infinite, but it will also not return to the original equilibrium. Thus, the system is Lyapunov stable but not asymptotically (see Fig. 1.15).



(a) Block diagram. (b) Pole-Zero diagram. (c) Natural response.

Figure 1.15: An integrator plant is marginally stable.

  - Another example is a second order system with zero damping, described by the state equations

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix} \quad , \tag{1.63}$$

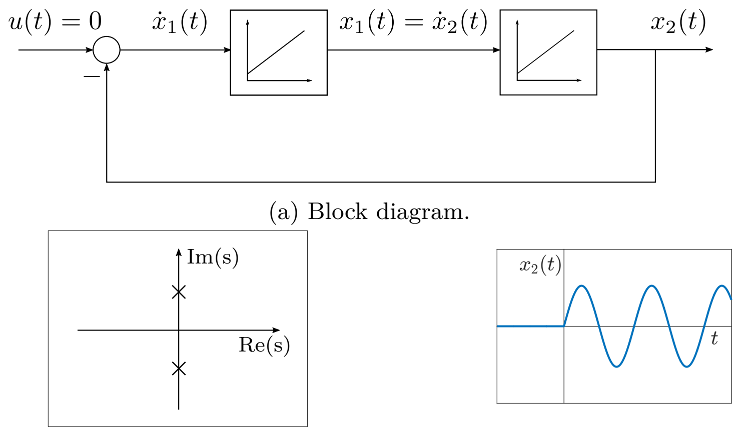
  that is equivalently described (for $\mathbf{x}(0) = 0$) by the transfer function

$$G(s) = \frac{1}{s^2 + 1} \quad . \tag{1.64}$$

  Fig. 1.16 shows the practical realization of such a system, that consists of two integrators in a feedback loop. The poles of this system s_i are $\pm j$. Therefore, an initial displacement will result in a non-damped sinusoidal motion at the eigenfrequency $\omega = 1\frac{\text{rad}}{\text{s}}$. The natural response does not fade out due to the lack of damping, this system is marginally stable.

- If a linear system has no eigenvalues in the RHP, and repeated poles on the imaginary axis, then

  - The system is marginally stable if $\mathbf{A}$ is diagonalizable, i.e. if all eigenvectors are independent. An example are two parallel integrators as shown in Fig. 1.17. This system can be described by the state equations



(a) Block diagram.

(b) Pole-Zero diagram.

(c) Natural response.

Figure 1.16: A second order system without damping is marginally stable.

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix} \quad . \tag{1.65}$$

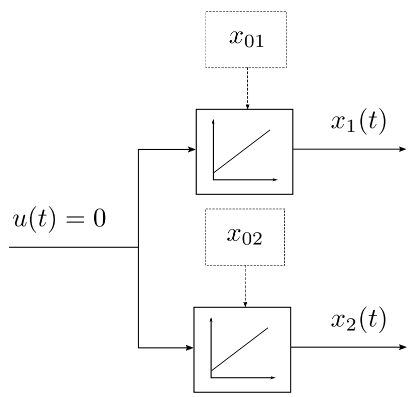
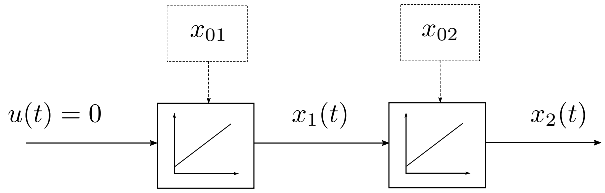


Figure 1.17: Block diagram of a parallel integrator system.

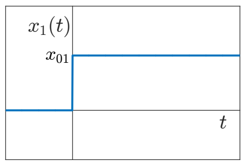
Since the states do not exert influence on each other, the system behaves as two separate (marginally stable) integrator plants. A parallel integrator system is therefore also marginally stable.

– The system is Lyapunov unstable if $\mathbf{A}$ is not diagonalizable, i.e. if some eigenvectors are linearly dependent. An example of such a system is a double integrator plant, as shown in Fig. 1.18.

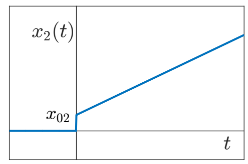
  The transfer function of the double integrator is $G(s) = \frac{1}{s^2}$. The system can also be described by the state equations



(a) Block diagram.



(b) Natural response of state 1.



(c) Natural response of state 2.

Figure 1.18: A double integrator system is Lyapunov unstable

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad . \tag{1.66}$$

The system has a double pole $s_i = 0$ and the system matrix is not diagonalizable. It is easy to see that the system output $x_2(t)$ will grow linearly in time for a given initial displacement x_{01}. The system is thus Lyapunov unstable.

Relation BIBO- and Lyapunov stability. If a linear system is Lyapunov stable, then it follows necessarily that the system is BIBO-stable as well. However, if a system is BIBO-stable it is not necessarily Lyapunov stable, since it could happen that there is an unstable eigenvalue of $\mathbf{A}$ that does not show up in the transfer function due to a pole-zero-cancellation. Then the system is BIBO-stable but Lyapunov unstable.

Remark. However, if the system is *controllable* and *observable* (notions that will be defined in part 3), then BIBO-stability also implies asymptotic Lyapunov stability. In this case, these stability concepts are equivalent.

1.2.3 The Hurwitz trick

The criteria for asymptotic Lyapunov stability and BIBO-stability of linear systems are very similar. The condition for asymptotic Lyapunov stability is that all eigenvalues λ_i of the system matrix $\mathbf{A}$ have a negative real part, and BIBO-stability of a system implies that all poles s_i of the transfer function have a negative real part. Using the Hurwitz-criterion discussed in SC1, we can easily check whether it is possible that that all solutions λ_i (or s_i) of a polynomial can have a negative real part, without having to calculate all λ_i (or s_i).

Theorem 2. *It follows from the Hurwitz criterion that the solutions λ_i of a polynomial*

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = 0 \quad , \tag{1.67}$$

can not all have a negative real part if at least one of following conditions is satisfied:

1. *it exists* $i = 1, \ldots, n$ *such that* $a_i = 0$.
2. *it exist* $i,j = 1, \ldots, n$ *such that* $\text{sign}(a_i) \neq \text{sign}(a_j)$.

In order to test for asymptotical Lyapunov stability, we can now apply this trick to the characteristic polynomial $p(\lambda)$ of the general ODE model or the state space model. If condition (1) or (2) holds for $p(\lambda)$, then the system is not asymptotically Lyapunov stable. Likewise, in order to test for BIBO-stability, we can inspect the denominator of the transfer function. If condition (1) or (2) holds, then the system is not BIBO-stable. Logically, if the system is not BIBO-stable, it is also not asymptotically Lyapunov stable.