Chapter 7
Closed-Loop Control

In this chapter, an introduction to the basics of continuous-time feedback systems is given. For more detailed treatments, the reader is referred to textbooks such as [1–5]. A simple amplitude control loop serves as an example in the following sections. The concepts presented here, however, may also be applied to more advanced control loops (cf. [6, 7]). The RF control loops are often called **low-level RF (LLRF) systems** to distinguish them from the high-power parts.

7.1 Basics of Continuous-Time Feedback Systems

Since many discrete feedback systems may be treated as quasicontinuous if the sampling time is small enough, discrete-time systems are not covered in the following. The analysis of discrete-time systems is, however, possible in an analogous way to continuous-time systems with the $\mathcal{Z}$-transform instead of the Laplace transform [8]. Most feedback analysis and design methods may then be used for discrete systems in a very similar way.

7.1.1 Linear Time-Invariant Systems

The systems under consideration are assumed to be linear and time-invariant (they are so-called **LTI systems**). Assume a general dynamic system

$$y(t) = \varphi\{x(t)\}$$

that maps the input signal $x(t)$ to the output signal $y(t)$. If the system is time-invariant, a time shift at the input will lead to the shifted output

$$\varphi\{x(t - t_0)\} = y(t - t_0). \tag{7.1}$$

In case of a linear system, a linear combination of two input signals $x_1(t)$ and $x_2(t)$ will lead to the same linear combination of their corresponding outputs $y_1(t) = \varphi\{x_1(t)\}$ and $y_2(t) = \varphi\{x_2(t)\}$, i.e.,

$$\varphi\{a_1 x_1(t) + a_2 x_2(t)\} = a_1 y_1(t) + a_2 y_2(t) \tag{7.2}$$

holds for arbitrary constants a_1 and a_2.

A consequence of properties (7.1) and (7.2) is that the output of LTI systems can be calculated in the Laplace domain as

$$Y(s) = H(s)\, X(s), \tag{7.3}$$

where the transfer function $H(s)$ corresponds to the impulse response $h(t)$ of the system as defined in Sect. 2.3, and $X(s)$ is the Laplace transform of the system input $x(t)$. This fact is of particular importance, because it enables the analysis and design of feedback systems in the Laplace domain. For a demonstration of the fact that Eq. (7.3) holds for any LTI system, we follow [9] and approximate the input signal $x(t)$ by the step function

$$x_{\text{step}}(t) = \sum_{\nu=0}^{\infty} x(\tau_\nu)\left(\Theta(t - \tau_\nu) - \Theta(t - \tau_{\nu+1})\right) \approx x(t),$$

where $\tau_\nu = \nu\Delta\tau$ are discrete sampling times with distance $\Delta\tau$ and $\Theta(t)$ is the Heaviside step function. It is assumed that $x(t)$ is zero for $t < 0$, as introduced in Sect. 2.2, for all functions for which the one-sided Laplace transform is used. The step response of the system, i.e., the output for $x(t) = \Theta(t)$, will be denoted by $y_\Theta(t)$ in the following. For the input $x_{\text{step}}(t)$, the LTI properties then lead to the output response

$$y_{\text{step}}(t) = \sum_{\nu=0}^{\infty} x(\tau_\nu)\left(y_\Theta(t - \tau_\nu) - y_\Theta(t - \tau_{\nu+1})\right).$$

The continuous output response $y(t)$ is obtained for the limit $\Delta\tau \to 0$:[1]

$$y(t) = \lim_{\Delta\tau \to 0} \sum_{\nu=0}^{\infty} x(\tau_\nu)\, \frac{y_\Theta(t - \tau_\nu) - y_\Theta(t - \tau_{\nu+1})}{\Delta\tau}\, \Delta\tau$$

$$= \int_0^{\infty} x(\tau)\, \dot{y}_\Theta(t - \tau)\, d\tau.$$

---

[1]The assumption is made that the step response $y_\Theta(t)$ is continuous at $t = 0$, continuously differentiable for $t > 0$, and zero for $t < 0$. However, the proof is also possible if $y_\Theta(t)$ is piecewise analytic for $t > 0$ and zero for $t < 0$ [10].

If the derivative $\dot{y}_\Theta(t)$ is denoted by the function $h(t)$, this is a convolution integral, and Eq. (2.27),

$$y(t) = h(t) * x(t) \quad \circ\!\!-\!\!\bullet \quad Y(s) = H(s)\,X(s),$$

holds. The choice of $h(t)$ is indeed not coincidental, because a comparison with Sect. 2.3 shows that due to $\dot{\Theta}(t) = \delta(t)$, this is the already defined impulse response, and the relation

$$h(t) = \dot{y}_\Theta(t) \quad \circ\!\!-\!\!\bullet \quad H(s) = s\,Y_\Theta(s)$$

holds for $t > 0$, i.e., the impulse response $h(t)$ is the derivative of the step response with respect to time. Conversely, it can easily be shown that systems defined by Eq. (7.3) are linear, because in the Laplace domain, the output $Y(s)$ results from a simple multiplication of the input $X(s)$ and the transfer function [10]. In addition, they are time-invariant, because the shifted input

$$x(t - t_0) \quad \circ\!\!-\!\!\bullet \quad X(s)\,e^{-t_0 s}$$

leads to the output

$$Y(s)\,e^{-t_0 s} \quad \bullet\!\!-\!\!\circ \quad y(t - t_0).$$

In summary, we can conclude that the definition of LTI systems by the properties (7.1) and (7.2) is equivalent to Definition (7.3).

In many cases, the transfer function $H(s)$ has the form

$$H(s) = \frac{b_0 + b_1 s + \ldots + b_m s^m}{a_0 + a_1 s + \ldots + a_n s^n} \tag{7.4}$$

with real coefficients b_ν and a_ν and nonzero coefficients $b_m \neq 0$ and $a_n \neq 0$. This is a rational transfer function, and the system (7.3) is then represented in the time domain by the linear ODE

$$a_0 y(t) + a_1 \dot{y}(t) + \ldots + a_n \frac{\mathrm{d}^n}{\mathrm{d}t^n} y(t) = b_0 x(t) + b_1 \dot{x}(t) + \ldots + b_m \frac{\mathrm{d}^m}{\mathrm{d}t^m} x(t) \tag{7.5}$$

with constant coefficients. A transfer function (7.4) is called **proper** if $m \leq n$ and **strictly proper** if $m < n$. In the latter case, $H(s)$ tends to zero as $|s| \to \infty$.

It is sometimes more convenient to use the **zero-pole-gain representation**

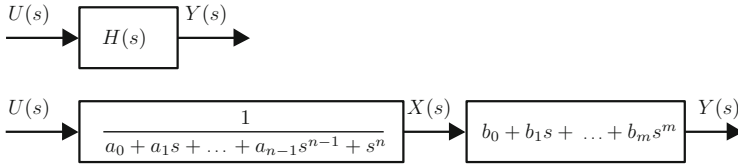$$H(s) = \frac{K(s - z_1)(s - z_2)\ldots(s - z_m)}{s^N(s - p_1)\ldots(s - p_{n-N})}. \tag{7.6}$$

Fig. 7.1 Derivation of the state-space representation

The **zeros** z_ν are those values for which $H(s)$ becomes zero, whereas the **poles** $p_\nu \neq 0$ are singularities of $H(s)$. In case a pole and a zero are exactly equal, they cancel and do not influence the input–output behavior of the system. The gain can also be expressed as $K = b_m/a_n$.

As will be discussed in the following, the system represented by $H(s)$ is called **stable** if all poles have a negative real part, i.e., $\text{Re}\{p_\nu\} < 0$ and $N = 0$. In this case, all poles lie in the open left half of the complex s-plane, which is referred to as OLHP. The abbreviations ORHP (open right half-plane), LHP (left half-plane), RHP (right half-plane) follow accordingly. If at least one pole has a positive real part, the system is **unstable**.

7.1.2 State-Space Representation

The higher-order ODE (7.5) can be rewritten as a system of ODEs of first order. Consider the transfer function $H(s)$ with input U and output Y, as shown in Fig. 7.1. The input variable $U(s)$ corresponds to $X(s)$ in the previous section. The notation is changed here to be consistent with the standard notation in the control system literature. Without loss of generality, it is assumed that $H(s)$ has the form (7.4) but with $a_n = 1$, i.e., the coefficients of $H(s)$ are normalized by $a_n \neq 0$. By splitting $H(s)$ in two blocks with its denominator and numerator, a new variable $X(s)$ may be defined as shown in Fig. 7.1.

In the time domain, the following ODEs can be derived from this block diagram:

$$u(t) = a_0 x(t) + a_1 \dot{x}(t) + \ldots + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} x(t) + \frac{d^n}{dt^n} x(t),$$

$$y(t) = b_0 x(t) + b_1 \dot{x}(t) + \ldots + b_m \frac{d^m}{dt^m} x(t).$$

Defining the **states** (see also Sect. 2.8.1)

$$x_1 := x, \quad x_2 := \dot{x}, \quad \ldots, \quad x_n := \frac{d^{n-1}}{dt^{n-1}} x(t), \tag{7.7}$$

leads to the system of equations

$$\dot{x}_1(t) = x_2(t),$$
$$\dot{x}_2(t) = x_3(t),$$
$$\vdots$$
$$\dot{x}_{n-1}(t) = x_n(t),$$
$$\dot{x}_n(t) = -a_0 x_1(t) - \ldots - a_{n-1} x_n(t) + u(t),$$
$$y(t) = b_0 x_1(t) + \ldots + b_m x_{m+1}(t).$$

With the definition of the **state vector**

$$\vec{x}(t) := \begin{bmatrix} x_1(t) & x_2(t) & \ldots & x_n(t) \end{bmatrix}^{\mathrm{T}},$$

the matrix representation

$$\frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t} = \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & \ldots \\ & & & \ddots & \\ & & & & 1 \\ -a_0 & -a_1 & \ldots & & -a_{n-1} \end{bmatrix} \cdot \vec{x}(t) + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \cdot u(t),$$

$$y(t) = \begin{bmatrix} b_0 & \ldots & b_m & 0 & \ldots & 0 \end{bmatrix} \cdot \vec{x}(t),$$

is obtained, which is called the **controllable canonical form** and is a special case of a **state space representation**. Different choices of the states (7.7) lead to different representations, but these have the general form

$$\frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t} = A \cdot \vec{x}(t) + B \cdot \vec{u}(t),$$
$$\vec{y}(t) = C \cdot \vec{x}(t),$$

with the state vector $\vec{x}$ of dimension n, the input vector $\vec{u}$ of dimension p, the output vector $\vec{y}$ of dimension q, the $n \times n$ system matrix A, the $n \times p$ input matrix B, and the $q \times n$ output matrix C. All matrices are assumed to have constant and real elements. A feedthrough matrix for a direct influence of $\vec{u}$ on $\vec{y}$ can be avoided in most practical cases. The Laplace transform of these equations yields[2]

$$s\,\vec{X}(s) - \vec{x}(0) = A\,\vec{X}(s) + B\,\vec{U}(s),$$

---

[2]In the following, we write $\vec{x}(0)$ instead of $\vec{x}(0+)$ because we assume that the value at $t = 0$ is defined by the limit $t \to 0$ for positive values of t.

or

$$\vec{X}(s) = (s\,I - A)^{-1}\left(B\,\vec{U}(s) + \vec{x}(0)\right),\qquad\qquad(7.8)$$

where I denotes the $n \times n$ identity matrix. The Laplace transform for the output $\vec{y}$ leads to

$$\vec{Y}(s) = C\,(s\,I - A)^{-1}\left(B\,\vec{U}(s) + \vec{x}(0)\right).$$

In case of a system with a single input and a single output (**SISO system**), the transfer function $H(s)$ is obtained as

$$H(s) = C(s\,I - A)^{-1}B,$$

where C is a row vector and B a column vector ($p = 1, q = 1$).

7.1.3   Linearization of Nonlinear Systems

Every practical system contains nonlinearities. Examples are nonlinear friction and constraints on the input that lead to saturation. Fortunately, in many cases, the considered nonlinear system behaves similarly to a linear system in the vicinity of its operating point. Consider a nonlinear system described by

$$\frac{d\vec{x}(t)}{dt} = \vec{v}(\vec{x}(t), \vec{u}(t))$$

with the analytic vector function $\vec{v}$. Suppose that $\vec{x} = \vec{x}_F$ and $\vec{u} = \vec{u}_F$ constitute a constant equilibrium point, i.e.,

$$\vec{v}(\vec{x}_F, \vec{u}_F) = 0.$$

With the use of the Jacobian matrix

$$\frac{\partial\vec{v}}{\partial\vec{x}} = \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \cdots & \frac{\partial v_1}{\partial x_n} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \cdots & \frac{\partial v_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial v_n}{\partial x_1} & \frac{\partial v_n}{\partial x_2} & \cdots & \frac{\partial v_n}{\partial x_n} \end{pmatrix},$$

the Taylor series expansion around the equilibrium can be written as

$$\frac{d\vec{x}(t)}{dt} = \vec{v}(\vec{x}_F, \vec{u}_F) + \left.\frac{\partial\vec{v}}{\partial\vec{x}}\right|_F \cdot (\vec{x} - \vec{x}_F) + \left.\frac{\partial\vec{v}}{\partial\vec{u}}\right|_F \cdot (\vec{u} - \vec{u}_F) + \vec{v}_{ho}(\vec{x} - \vec{x}_F, \vec{u} - \vec{u}_F),$$

where $\big|_F$ denotes the value at the equilibrium and $\vec{v}_{ho}$ are higher-order terms. For small deviations

$$\Delta\vec{x}(t) = \vec{x}(t) - \vec{x}_F, \quad \Delta\vec{u}(t) = \vec{u}(t) - \vec{u}_F$$

from equilibrium, the higher-order terms may be neglected, and the linear system

$$\frac{d\Delta\vec{x}(t)}{dt} = A \cdot \Delta\vec{x}(t) + B \cdot \Delta\vec{u}(t)$$

with

$$A = \frac{\partial\vec{v}}{\partial\vec{x}}\bigg|_F, \quad B = \frac{\partial\vec{v}}{\partial\vec{u}}\bigg|_F,$$

can be used as a linearization of the nonlinear system.

7.1.4 Dynamic Response of LTI Systems

The output of an LTI system depends on its transfer function $H(s)$ and on the input signal $u(t)$. In the following, the response of a general LTI system with respect to important test signals is discussed. This prepares the definition of stability. It is assumed that the poles and zeros of $H(s)$ are all distinct, apart from N poles at $s = 0$. In most cases, this is a valid assumption. The calculations for the case with poles or zeros of higher multiplicity are similar but more intricate. Because the coefficients in Eq. (7.4) are real, nonreal poles p or zeros z are always accompanied by their complex conjugate counterparts p^* and z^*. The complex conjugate operator commutes with every holomorphic function $f(x)$ on its domain of definition if $f(x)$ is real for real x. Thus in this case, $f^*(x)$ equals $f(x^*)$. In particular, this applies to every polynomial and rational function with real coefficients.

According to Eq. (7.6), the considered transfer function can be written as

$$H(s) = \frac{K \prod_{v=1}^{m_1}(s - z_{r,v}) \prod_{v=1}^{m_2}(s - z_{c,v})(s - z_{c,v}^*)}{s^N \prod_{v=1}^{n_1}(s - p_{r,v}) \prod_{v=1}^{n_2}(s - p_{c,v})(s - p_{c,v}^*)}, \tag{7.9}$$

where the $z_{r,v}$ and $p_{r,v}$ are the nonzero real zeros and poles, $z_{c,v}$ and $p_{c,v}$ are the nonzero complex zeros and poles, and K is the real gain. The total polynomial degree equals $m = m_1 + 2m_2$ for the numerator and $n = N + n_1 + 2n_2$ for the denominator. For a proper transfer function, $n \geq m$ holds.

7.1.4.1 Impulse Response

The impulse response is of practical interest for the study of pulse-shaped disturbances that may act on the feedback loop. In addition, this case is equivalent to the response of the state-space representation with zero input and certain nonzero initial conditions $\vec{x}(t = 0) \neq 0$.

According to Eq. (7.3), the excitation of the system with the Dirac function

$$u(t) = \delta(t) \quad \circ\!\!-\!\!\bullet \quad U(s) = 1$$

yields

$$Y(s) = H(s) \cdot 1 = H(s)$$

in the Laplace domain. To calculate the response in the time domain, the partial fraction decomposition

$$Y(s) = H(s) \overset{!}{=} \sum_{\nu=0}^{N} \frac{K_{0,\nu}}{s^{\nu}} + \sum_{\nu=1}^{n_1} \frac{K_{\mathrm{r},\nu}}{s - p_{\mathrm{r},\nu}} + \sum_{\nu=1}^{n_2} \left(\frac{K_{\mathrm{c}1,\nu}}{s - p_{\mathrm{c},\nu}} + \frac{K_{\mathrm{c}2,\nu}}{s - p_{\mathrm{c},\nu}^{*}} \right)$$

$$(7.10)$$

is used. Here one assumes that the transfer function $H(s)$ is proper, i.e., $n \geq m$. The constants $K_{\mathrm{r},\nu}$ can be calculated as follows. Multiplying Eqs. (7.9) and (7.10) by $(s - p_{\mathrm{r},i})$ for a specific $i = 1, \ldots, n_1$ and setting $s = p_{\mathrm{r},i}$ leads to

$$K_{\mathrm{r},i} = \left[H(s) \, (s - p_{\mathrm{r},i}) \right]_{s = p_{\mathrm{r},i}}$$

$$= \frac{K \prod_{\nu=1}^{m_1} (p_{\mathrm{r},i} - z_{\mathrm{r},\nu}) \prod_{\nu=1}^{m_2} (p_{\mathrm{r},i} - z_{\mathrm{c},\nu})(p_{\mathrm{r},i} - z_{\mathrm{c},\nu}^{*})}{p_{\mathrm{r},i}^{N} \prod_{\nu=1,\nu\neq i}^{n_1} (p_{\mathrm{r},i} - p_{\mathrm{r},\nu}) \prod_{\nu=1}^{n_2} (p_{\mathrm{r},i} - p_{\mathrm{c},\nu})(p_{\mathrm{r},i} - p_{\mathrm{c},\nu}^{*})}.$$

The constants $K_{\mathrm{r},\nu}$ are always real, because in the denominator, the expression

$$(p_{\mathrm{r},i} - p_{\mathrm{c},\nu})(p_{\mathrm{r},i} - p_{\mathrm{c},\nu}^{*}) = (p_{\mathrm{r},i} - \mathrm{Re}\{p_{\mathrm{c},\nu}\})^{2} + (\mathrm{Im}\{p_{\mathrm{c},\nu}\})^{2}$$

is real, and the same applies to the numerator. A similar calculation yields the constants

$$K_{\mathrm{c}1,i} = \left[H(s) \, (s - p_{\mathrm{c},i}) \right]_{s = p_{\mathrm{c},i}},$$

$$K_{\mathrm{c}2,i} = \left[H(s) \, (s - p_{\mathrm{c},i}^{*}) \right]_{s = p_{\mathrm{c},i}^{*}},$$

and using the above-mentioned commutability property of the complex conjugate operator leads to

$$K_{\mathrm{c}2,i} = K_{\mathrm{c}1,i}^{*}.$$

The constant $K_{0,N}$ is obtained by multiplying by s^N; it reads

$$K_{0,N} = \left[H(s) \, s^N \right]_{s=0} = \frac{K \prod_{\nu=1}^{m_1}(-z_{\mathrm{r},\nu}) \prod_{\nu=1}^{m_2} |z_{\mathrm{c},\nu}|^2}{\prod_{\nu=1}^{n_1}(-p_{\mathrm{r},\nu}) \prod_{\nu=1}^{n_2} |p_{\mathrm{c},\nu}|^2}.$$

For the remaining constants $K_{0,i}$, a system of N linear equations is obtained by evaluating Eqs. (7.9) and (7.10) at N points $s = s_i$ that are different from the zeros and poles of the system. The constant $K_{0,0}$ is zero for strictly proper transfer functions $H(s)$, i.e., for $n > m$.

The transformation of Eq. (7.10) into the time domain

$$Y(s) = H(s) \quad \bullet\!\!-\!\!\circ \quad y(t) = h(t)$$

yields the impulse response

$$h(t) = K_{0,0}\, \delta(t) + \sum_{\nu=1}^{N} K_{0,\nu} \frac{t^{\nu-1}}{(\nu-1)!} + \sum_{\nu=1}^{n_1} K_{\mathrm{r},\nu}\, e^{p_{\mathrm{r},\nu} t}$$

$$+ \sum_{\nu=1}^{n_2} \left(K_{\mathrm{c}1,\nu}\, e^{p_{\mathrm{c},\nu} t} + K_{\mathrm{c}1,\nu}^{*}\, e^{p_{\mathrm{c},\nu}^{*} t} \right),$$

as Table A.4 shows ($\Theta(t)$ is omitted for the sake of simplicity). The elements of the last sum can be rewritten as

$$e^{\mathrm{Re}\{p_{\mathrm{c},\nu}\}t} \left[K_{\mathrm{c}1,\nu}\, e^{j\,\mathrm{Im}\{p_{\mathrm{c},\nu}\}t} + K_{\mathrm{c}1,\nu}^{*}\, e^{-j\,\mathrm{Im}\{p_{\mathrm{c},\nu}\}t} \right],$$

and the term of this expression in square brackets is equal to

$$|K_{\mathrm{c}1,\nu}| \left(e^{j\left(\mathrm{Im}\{p_{\mathrm{c},\nu}\}t + \angle K_{\mathrm{c}1,\nu}\right)} + e^{-j\left(\mathrm{Im}\{p_{\mathrm{c},\nu}\}t + \angle K_{\mathrm{c}1,\nu}\right)} \right),$$

where $|K|$ and $\angle K$ are the amplitude and phase of the complex number K, respectively. Altogether, the impulse response for $t \geq 0$ is

$$h(t) = K_{0,0}\, \delta(t) + \sum_{\nu=1}^{N} K_{0,\nu} \frac{t^{\nu-1}}{(\nu-1)!} + \sum_{\nu=1}^{n_1} K_{\mathrm{r},\nu}\, e^{p_{\mathrm{r},\nu} t} +$$

$$+ 2 \sum_{\nu=1}^{n_2} |K_{\mathrm{c}1,\nu}|\, e^{\mathrm{Re}\{p_{\mathrm{c},\nu}\}t}\, \cos\left(\mathrm{Im}\{p_{\mathrm{c},\nu}\}t + \angle K_{\mathrm{c}1,\nu}\right),$$

(7.11)

and it tends to zero as $t \to \infty$ if all poles have negative real parts.

7.1.4.2 Step Response

The response $y(t) = y_\Theta(t)$ to a step command

$$u(t) = \Theta(t)$$

can be calculated in an analogous way with $U(s) = 1/s$. An alternative is the use of the convolution integral

$$y_\Theta(t) = \int_0^t h(\tau)\,\Theta(t - \tau)\,\mathrm{d}\tau = \int_0^t h(\tau)\,\mathrm{d}\tau. \qquad (7.12)$$

With

$$\int_0^t e^{\delta\tau} \cos(\omega\tau + \varphi)\,\mathrm{d}\tau = \frac{\delta}{\delta^2 + \omega^2}\left(e^{\delta t} \cos(\omega t + \varphi) - \cos\varphi\right)$$

$$+ \frac{\omega}{\delta^2 + \omega^2}\left(e^{\delta t} \sin(\omega t + \varphi) - \sin\varphi\right)$$

$$= \frac{1}{\sqrt{\delta^2 + \omega^2}}\left(e^{\delta t} \cos\left(\omega t + \varphi - \arctan\frac{\omega}{\delta}\right)\right.$$

$$\left. - \cos\left(\varphi - \arctan\frac{\omega}{\delta}\right)\right),$$

the integration of (7.11) for $t \geq 0$ yields

$$y_\Theta(t) = K_{0,0} + \sum_{\nu=1}^{N} K_{0,\nu}\frac{t^\nu}{\nu!} + \sum_{\nu=1}^{n_1} \frac{K_{\mathrm{r},\nu}}{p_{\mathrm{r},\nu}}\left(e^{p_{\mathrm{r},\nu}t} - 1\right) +$$

$$+ \sum_{\nu=1}^{n_2} \frac{2|K_{\mathrm{c1},\nu}|}{|p_{\mathrm{c},\nu}|}\left(e^{\mathrm{Re}\{p_{\mathrm{c},\nu}\}t} \cos(\mathrm{Im}\{p_{\mathrm{c},\nu}\}t + \measuredangle K_{\mathrm{c1},\nu} - \measuredangle p_{\mathrm{c},\nu})\right.$$

$$\left. - \cos(\measuredangle K_{\mathrm{c1},\nu} - \measuredangle p_{\mathrm{c},\nu})\right). \qquad (7.13)$$

This calculation shows that the step response y_Θ will approach a finite value for large times t if and only if the conditions

$$N = 0, \quad p_{\mathrm{r},\nu} < 0, \quad \mathrm{Re}\{p_{\mathrm{c},\nu}\} < 0$$

are satisfied, i.e., all poles have negative real parts. Because the limit $\lim_{t\to\infty} y_\Theta(t)$ is then finite, the final value theorem can be applied:

$$\lim_{t\to\infty} y_\Theta(t) = \lim_{s\to 0} s\,Y_\Theta(s) = H(0).$$

The initial value theorem leads to

$$\lim_{t \to +0} y_\Theta(t) = K_{0,0} = \lim_{s \to \infty} s\, Y_\Theta(s) = \lim_{s \to \infty} H(s) = \begin{cases} 0 & \text{if } n > m, \\ K & \text{if } n = m. \end{cases}$$

For this reason, systems with $n = m$ are also said to have direct feedthrough. In contrast, strictly proper transfer functions with $n > m$ have a continuous output response at $t = 0$.

7.1.4.3   Frequency Response

An important test signal is the harmonic excitation

$$u(t) = \sin(\omega t) \quad \circ\!\!-\!\!\bullet \quad \frac{\omega}{s^2 + \omega^2}.$$

The amplitude of the test signal may be chosen arbitrarily because of the linearity property (7.2). If it is assumed that none of the poles of $H(s)$ is equal to $\pm j\omega$, the decomposition of the output response in the Laplace domain can be written as

$$Y(s) = H(s)\, \frac{\omega}{s^2 + \omega^2} = \frac{K_\omega}{s - j\omega} + \frac{K_\omega^*}{s + j\omega} + Y_{\text{trans}}(s),$$

where Y_{trans} has the same structure as the expression in Eq. (7.10), but with $K_{0,0} = 0$. A multiplication by $(s - j\omega)$ and the evaluation at $s = j\omega$ leads to

$$K_\omega = \left[H(s)\, \frac{\omega(s - j\omega)}{s^2 + \omega^2} \right]_{s=j\omega} = \left[H(s)\, \frac{\omega}{s + j\omega} \right]_{s=j\omega} = \frac{1}{2j} H(j\omega).$$

In the time domain, the output response reads

$$y(t) = K_\omega e^{j\omega t} + K_\omega^* e^{-j\omega t} + y_{\text{trans}}(t)$$

$$= H(j\omega) \frac{e^{j\omega t}}{2j} - H^*(j\omega) \frac{e^{-j\omega t}}{2j} + y_{\text{trans}}(t)$$

$$= |H(j\omega)|\, \sin(\omega t + \angle H(j\omega)) + y_{\text{trans}}(t).$$

If the transfer function $H(s)$ has only poles with negative real parts, the transient response y_{trans} will tend to zero, and $y(t)$ tends to a constant oscillation. The amplitude and phase of this oscillation with respect to the excitation $u(t)$ is determined by $H(j\omega)$, i.e., the value of the transfer function at $s = j\omega$. Because of the linearity property, this also applies to any shifted or scaled sinusoidal excitation. For this reason, the function $H(j\omega)$ depending on the frequency ω is called the

frequency response of the system $H(s)$ and is obtained by introducing $s = j\omega$ into $H(s)$.

There are two main reasons why the frequency response is important for feedback systems. First, $H(j\omega)$ can easily be measured by exciting the system with different frequencies ω, even if the transfer function $H(s)$ of the physical system is not known. Second, $H(j\omega)$ can be used for the stability analysis of the closed feedback loop with the Nyquist criterion (see Sect. 7.4.2).

So far, it has been assumed that $j\omega$ is not a pole of $H(s)$. Without further calculation, it can be reasoned that if $j\omega$ is a pole, $H(s)$ has a singularity at $H(j\omega)$ and the excitation with frequencies close to ω will lead to very large amplitudes. If the chosen frequency is exactly ω, this will result in a perfect resonance, and the oscillation at the output will grow without bound, although the input is a bounded signal.

7.1.4.4 General Input Function

In the previous sections, the Laplace transform was used to calculate specific output responses for SISO systems. In case of general input functions, **multiple-input and multiple-output (MIMO) systems**, or initial values, it is often more convenient to consider the state-space representation. In Sect. 2.8.6, it was shown that autonomous linear systems of differential equations

$$\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = A \cdot \vec{r} \qquad \vec{r}(0) = \vec{r}_0$$

have the solution (2.99),

$$\vec{r}(t) = e^{tA}\,\vec{r}_0,$$

where e^{tA} is the matrix exponential function. In the presence of an input vector $\vec{u}(t)$, the system is no longer autonomous in general. The input may be a control effort or a disturbance such as a noise signal. In Sect. 7.1.2, the Laplace domain solution of a system with inputs was given by Eq. (7.8) as

$$\vec{X}(s) = (sI - A)^{-1}\vec{x}(0) + (sI - A)^{-1}B\vec{U}(s).$$

Comparing this with the solution $\vec{r}(t)$ of the autonomous system, it is apparent that

$$(sI - A)^{-1} \quad \bullet\!\!-\!\!\circ \quad e^{tA}$$

must hold, i.e., we have found the Laplace transform of the matrix exponential function. Transforming $\vec{X}(s)$ into the time domain thus leads to

$$\vec{x}(t) = e^{tA}\vec{x}(0) + \int_0^t e^{(t-\tau)A}B\vec{u}(\tau)\,\mathrm{d}\tau.$$

It can be shown that the matrix exponential function has the following properties, similar to those of an ordinary exponential function (cf. [11]):

- series representation: $e^{tA} = I + \sum_{\nu=1}^{\infty} A^{\nu} \frac{t^{\nu}}{\nu!}$
- inverse: $\left(e^{tA}\right)^{-1} = e^{-tA}$
- multiplication: $e^{t_2 A} e^{t_1 A} = e^{(t_2 + t_1) A}$

7.1.5 Stability

In Sects. 2.8.6 and 2.8.10, it was shown that a linear autonomous system is asymptotically stable if and only if all eigenvalues of the system matrix A have negative real parts, i.e., are situated in the OLHP. Equivalently, the same holds for the roots of the characteristic equation. Asymptotic stability for autonomous systems implies that a trajectory that starts at some initial value will tend to a fixed point.

For a system with nonzero inputs $\vec{u}(t)$, this definition may not be sufficient. The input can be a persistent disturbance with a certain amplitude that prevents the system from approaching the fixed point. For a feedback system, it is, however, necessary that the states $\vec{x}(t)$ or the output $\vec{y}(t)$ remain bounded. This motivates the following definition:

Definition 7.1. A dynamical system

$$\frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t} = \vec{v}_1(\vec{x}(t), \vec{u}(t)), \quad \vec{y}(t) = \vec{v}_2(\vec{x}(t), \vec{u}(t))$$

with input $\vec{u}(t)$, states $\vec{x}(t)$, and output $\vec{y}(t)$ is assumed to be in equilibrium for $t = t_0$ with arbitrary real t_0, i.e., $\vec{x}(t_0) = \vec{x}_F$, where $\vec{x}_F$ is a fixed point. This fixed point is said to be **bounded-input bounded-output (BIBO) stable** if for every finite c_1 with $\|\vec{u}(t)\| < c_1$ for $t \geq t_0$, there exists a finite c_2 such that $\|\vec{y}(t)\| \leq c_2$ for $t \geq t_0$.

(See, e.g., Ludyk [11, Definition 3.37, p. 159].)

The step response (7.13) shows that $y_\Theta(t)$ is bounded if all poles of $H(s)$ have negative real parts. Because of Eq. (7.12), this is also true if

$$|y_\Theta(t)| = \left| \int_0^t h(\tau) \, \mathrm{d}\tau \right| \leq \int_0^t |h(\tau)| \, \mathrm{d}\tau \leq \int_0^{\infty} |h(\tau)| \, \mathrm{d}\tau \leq c_2 < \infty$$

holds, i.e., if the impulse response $h(t)$ is absolutely integrable.

In general, the following theorem holds.

Theorem 7.2. *An LTI SISO system is BIBO stable if and only if the following (equivalent) conditions are satisfied:*

- *The transfer function $H(s)$ has only poles with negative real parts.*
- *the impulse response $h(t)$ is absolutely integrable.*

(See, e.g., Ludyk [11, Theorems 3.39 and 3.40, p. 160].)

In addition, there is a close relationship between BIBO and asymptotic stability. The transfer function $H(s)$ can be written as

$$H(s) = C(sI - A)^{-1}B = \frac{C \text{ adj}(sI - A) \, B}{\det(sI - A)},$$

where $\text{adj}(A)$ denotes the adjugate[3] matrix of A. Thus, the poles of $H(s)$ are obtained by calculating the roots of the characteristic equation

$$\det(sI - A) = 0,$$

and these are identical to the eigenvalues of A. However, due to pole–zero cancelations, the poles are, in general, a subset of the eigenvalues of A, i.e., not every eigenvalue is a pole of $H(s)$. If A has only eigenvalues with negative real parts, the system is asymptotically stable, and this always implies that the poles have negative real parts. This consideration leads to the following theorem:

Theorem 7.3. *An LTI system that is asymptotically stable is also BIBO stable, but a BIBO stable system is not always asymptotically stable.*

(See, e.g., Ludyk [Theorem 3.41, p. 160][11].)

7.2 Standard Closed Loop

The block diagram in Fig. 7.2 is called the standard feedback loop. It has one input and one output and is thus also called a **single-input single-output (SISO)** system.

The feedback system can be described by the following equations:

$$Y(s) = X_{d2}(s) + H_p(s)\Big[X_{d1}(s) + U(s)\Big],$$

$$U(s) = H_c(s)X_e(s),$$

$$X_e(s) = Y_r(s) - H_m(s)\Big[X_{d3}(s) + Y(s)\Big].$$

Solving these equations for the output $Y(s)$ leads to

---

[3]The cofactor matrix of A is a matrix that consists of the (i, k) minors of A multiplied by the factor $(-1)^{i+k}$. The adjugate matrix of A is the transpose of the cofactor matrix of A.
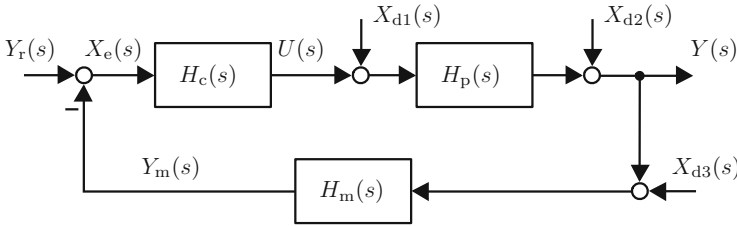
Fig. 7.2 Standard feedback loop: transfer functions of process $H_p(s)$, controller $H_c(s)$, and measurement $H_m(s)$. The signals are reference $Y_r(s)$, control error $X_e(s)$, input $U(s)$, output $Y(s)$, and disturbances $X_{d1}(s)$, $X_{d2}(s)$, and $X_{d3}(s)$

$$Y(s) = H_{ry}(s)\Big[Y_r(s) - H_m(s)X_{d3}(s)\Big] + H_{dy}(s)\Big[H_p(s)X_{d1}(s) + X_{d2}(s)\Big]$$
(7.14)

with the reference to output transfer function

$$H_{ry}(s) = \frac{H_p(s)H_c(s)}{1 + H_p(s)H_c(s)H_m(s)}$$
(7.15)

and the disturbance to output transfer function

$$H_{dy}(s) = \frac{1}{1 + H_p(s)H_c(s)H_m(s)}.$$
(7.16)

A **unity feedback system** has $H_m(s) = 1$, and in this case, the disturbance to output transfer function

$$H_{dy}(s) = \frac{1}{1 + H_p(s)H_c(s)}$$

is also called the **sensitivity function**, and the reference to output transfer function

$$H_{ry}(s) = \frac{H_p(s)H_c(s)}{1 + H_p(s)H_c(s)}$$

is the **complementary sensitivity function**. Note that $H_{dy}(s) + H_{ry}(s) = 1$.

Usually, the process transfer function $H_p(s)$ has to be determined in a separate modeling step before the analysis or the design of the feedback loop. The modeling can be based on analytical equations if the underlying physical principles are well known. If this is not the case, measurements may be used for a system identification. In both cases, modeling assumptions have to be made to limit the complexity of the system. Often, nonlinearities in the feedback loop are linearized, and high-frequency dynamics are omitted.
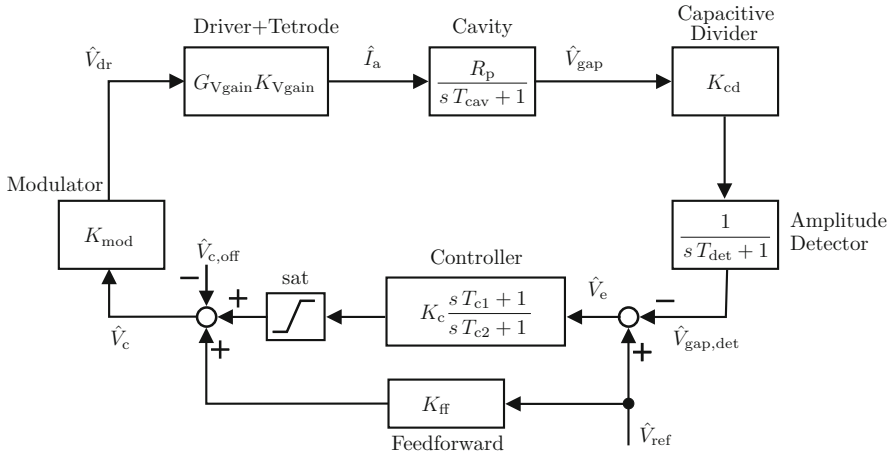
Fig. 7.3 Model of the amplitude feedback loop

7.3   Example: Amplitude Feedback

As a realistic example of a feedback loop, the amplitude feedback control of a ferrite-loaded cavity will be considered. The feedback is needed to hold the amplitude $\hat{V}_{\text{gap}}$ of the RF voltage close to a given reference value $\hat{V}_{\text{gap,ref}}$. In our example, the cavity feedback loop behaves highly nonlinearly with respect to the RF frequency f_{RF} and the reference amplitude $\hat{V}_{\text{ref}}$. In the following, the operating point

$$f_{\text{RF}} = 3 \text{ MHz}, \quad \hat{V}_{\text{gap,ref}} = 2 \text{ kV},$$

will be considered. A model of the feedback loop was obtained in [12] based on measurements, and the corresponding block diagram is shown in Fig. 7.3. In the following, only amplitudes of RF signals are used, not the RF signals themselves.

The feedback loop consists of the following subcomponents:

- The **cavity** is driven by the anode current with the amplitude $\hat{I}_{\text{a}}$. The amplitude of the resulting gap voltage $\hat{V}_{\text{gap}}$ acts approximately as a first-order system (PT$_1$) with respect to $\hat{I}_{\text{a}}$ (see also Appendix A.7.1). The "gain" is equal[4] to $R_{\text{p}} \approx 2700 \,\Omega$, and the time constant is $T_{\text{cav}} \approx 4 \,\mu\text{s}$. The set points are $\hat{V}_{\text{gap}} = 2 \text{ kV}$ and $\hat{I}_{\text{a}} = 0.75 \text{ A}$.
- A **capacitive divider** is used to downscale the gap voltage of one-half the gap with a factor of 1000. This has no significant influence on the time constants in the loop. With respect to the total gap voltage, the scaling is $K_{\text{cd}} = 1/2000$.

---

[4]Due to the output impedance of the tetrode, this value is about one-half the pure cavity impedance specified in Table 4.1 on p. 198.
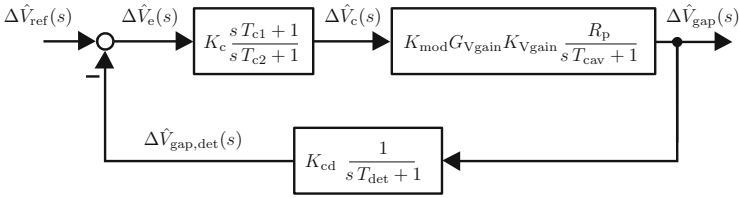
Fig. 7.4 Small-signal model of the amplitude feedback loop

- An **amplitude detector** with time constant $T_{det} = 5\,\mu s$ is used to obtain the amplitude $\hat{V}_{gap,det}$. This amplitude is then compared to the reference $\hat{V}_{ref}$. The set points are $\hat{V}_{gap,det} = 1\,V$ and $\hat{V}_{ref} = 1.04\,V$.
- The parameters of the **controller** are $K_c = 14.9$, $T_{c1} = 17.2\,\mu s$, and $T_{c2} = 487.2\,\mu s$. A saturation limit **sat** follows that limits the control output to $\pm 7.23\,V$. The offset voltage is $\hat{V}_{c,off} = 0.2\,V$. In the **feedforward** loop, the gain is $K_{ff} = 0.6$. According to these values, the set point of the control effort is $\hat{V}_c = 1.02\,V$.
- The (amplitude) **modulator** produces a sinusoidal signal modulated with $\hat{V}_c$. The sinusoidal signal with initial amplitude $0.316\,V$ ($0\,dBm$) is damped with a factor of $-12.2\,dB$; this corresponds to a factor of 0.245 for the voltage amplitude. Altogether, the modulator can be modeled as a gain $K_{mod} = 0.316 \cdot 0.245$. Hence, the set point of the driving voltage is $\hat{V}_{dr} = 79\,mV$.
- The gains of the **driver and tetrode amplifiers** depend on the RF frequency and the amplitude of the gap voltage. For the chosen setting, we have $G_{Vgain} \approx 27\,S$ and $K_{Vgain} \approx 0.35$.

Signal time delays with a magnitude of about $1\,\mu s$ are neglected in the following. However, they would be important for larger feedback gains.

The given set-point values were obtained by choosing $\hat{V}_{gap} = 2\,kV$. Because the stationary gain of the cavity transfer function is R_p, the necessary anode current amplitude equals $\hat{I}_a = \hat{V}_{gap}/R_p$. All other set-point values in the feedback loop follow accordingly. This results in a reference $\hat{V}_{ref}$ that is slightly higher than $\hat{V}_{gap,det}$ and thus in a stationary control error $\hat{V}_e = 40\,mV$. This steady-state error could be avoided by introducing an integral controller in the loop. However, it is also possible to adjust the reference in such a way that the desired value $\hat{V}_{gap}$ is reached, as has been done in this case.

The system is nonlinear due to the saturation function. This function and the offset values $\hat{V}_{ref}$ and $\hat{V}_{c,off}$ can be neglected if only small deviations with respect to the set point are considered. This leads to the linearized feedback loop in standard notation, as shown in Fig. 7.4 with amplitude error $\Delta\hat{V}_{gap} = \hat{V}_{gap} - \hat{V}_{gap,ref}$ and reference $\Delta\hat{V}_{ref} = 0$. Similarly, all other values are defined relative to their set-point values, e.g., the relative control effort is $\Delta\hat{V}_c = \hat{V}_c - 1.02\,V$.

A calculation of the reference to output transfer function according to (7.15) yields

$$H_{\text{ry}}(s) = \frac{K_{\text{c}}K_{\text{mod}}G_{\text{Vgain}}K_{\text{Vgain}}R_{\text{p}}(sT_{\text{c1}} + 1)(sT_{\text{det}} + 1)}{(sT_{\text{det}} + 1)(sT_{\text{c2}} + 1)(sT_{\text{cav}} + 1) + K_{\text{c}}K_{\text{mod}}G_{\text{Vgain}}K_{\text{Vgain}}R_{\text{p}}(sT_{\text{c1}} + 1)K_{\text{cd}}}.$$

A zero-pole-gain representation of this transfer function can be obtained by a numerical calculation of the poles and zeros. The gain is equal to the ratio of the factors of the highest order in s in the numerator and denominator. For the amplitude loop, these orders are s^2 and s^3, respectively, and the gain is

$$K = \frac{K_{\text{c}}K_{\text{mod}}G_{\text{Vgain}}K_{\text{Vgain}}R_{\text{p}}T_{\text{c1}}T_{\text{det}}}{T_{\text{det}}T_{\text{c2}}T_{\text{cav}}} = 2.6 \cdot 10^8 \; s^{-1}.$$

The resulting zero-pole-gain representation is

$$H_{\text{ry}}(s) = \frac{\Delta \hat{V}_{\text{gap}}(s)}{\Delta \hat{V}_{\text{ref}}(s)} = 2.6 \cdot 10^8 \; s^{-1} \; \frac{(s - z_1)(s - z_2)}{(s - p_1)(s - p_2)(s - p_3)}$$

with zeros

$$z_1 = -5.81 \cdot 10^4 \; s^{-1}, \quad z_2 = -2 \cdot 10^5 \; s^{-1},$$

and poles

$$p_1 = -2.42 \cdot 10^4 \; s^{-1}, \quad p_{2,3} = -(2.14 \pm j \; 1.44) \cdot 10^5 \; s^{-1}.$$

Thus, the closed-loop system is BIBO stable. The pole p_1 is closest to the imaginary axis and dominates the dynamics of the feedback. The dominating pole corresponds to a closed-loop bandwidth and a time constant of

$$\omega_{\text{ry}} = -p_1 = 2.42 \cdot 10^4 \; s^{-1} \quad \Rightarrow \quad T_{\text{ry}} = -\frac{1}{p_1} \approx 40 \, \mu s.$$

The absolute values of the remaining poles are larger by an order of magnitude. They are thus negligible for a first rough evaluation of the closed-loop dynamics.

7.4 Analysis and Stability

The closed-loop transfer function

$$H_{\text{ry}}(s) = \frac{b_0 + b_1 s + b_2 s^2}{a_0 + a_1 s + a_2 s^2 + s^3}$$

can be obtained from the given open-loop transfer function using only basic manipulations. The calculation of the poles p_i from the characteristic equation

$$0 = a_0 + a_1 s + a_2 s^2 + s^3$$

is a more complex task, and numerical computations are necessary for higher-order systems in general. For a stability analysis, one may, however, not be interested in the exact values of the poles, but only in the decision whether all poles have negative real parts. There are several **stability criteria** that can be applied without solving the characteristic equation directly. The Hurwitz and Nyquist criteria will be presented in the next sections.

7.4.1 Routh–Hurwitz Stability Criterion

The **Routh–Hurwitz criterion** is a necessary and sufficient condition for the roots of the polynomial

$$a_0 + a_1 s + \ldots + a_{n-1} s^{n-1} + s^n \tag{7.17}$$

to have only negative real parts, in which case the polynomial is then called a **Hurwitz polynomial**. The criterion is of particular interest if the coefficients a_i contain undetermined parameters. An example of such a parameter is the controller gain in the feedback loop. With the Routh–Hurwitz criterion, inequalities in these parameters can then be obtained for the closed loop to be stable.

A first necessary condition is given by the following theorem:

Theorem 7.4. *If the polynomial* (7.17) *is Hurwitz, then it has only positive coefficients* $a_i > 0$, $i = 0, 1, \ldots, n-1$.

(See, e.g., Ludyk [11, Theorem 3.43, p. 161].)

This enables a first simple test whether a polynomial can be Hurwitz. If any of the coefficients is missing, i.e., $a_i = 0$, or any a_i is negative, there will be roots with nonnegative real part, and the polynomial is not Hurwitz.

A necessary and sufficient condition is presented by the Hurwitz criterion. It uses the $\nu \times \nu$ Hurwitz determinants

$$H_\nu := \det \begin{pmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \ldots & a_{n-2\nu+1} \\ 1 & a_{n-2} & a_{n-4} & \ldots & a_{n-2\nu+2} \\ 0 & a_{n-1} & a_{n-3} & \ldots & a_{n-2\nu+3} \\ 0 & 1 & a_{n-2} & \ldots & a_{n-2\nu+4} \\ 0 & 0 & a_{n-1} & \ldots & a_{n-2\nu+5} \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \ldots & a_{n-\nu} \end{pmatrix}, \tag{7.18}$$

where the coefficients a_i in the matrix with an index $i < 0$ are set to zero. As an example, the first three determinants for a polynomial with degree $n \geq 5$ are

$$H_1 := a_{n-1},$$

$$H_2 := \det \begin{pmatrix} a_{n-1} & a_{n-3} \\ 1 & a_{n-2} \end{pmatrix},$$

$$H_3 := \det \begin{pmatrix} a_{n-1} & a_{n-3} & a_{n-5} \\ 1 & a_{n-2} & a_{n-4} \\ 0 & a_{n-1} & a_{n-3} \end{pmatrix}.$$

The Hurwitz criterion is given by the following theorem:

Theorem 7.5. *The polynomial* (7.17) *is Hurwitz if and only if the Hurwitz determinants H_ν defined by* (7.18) *are positive for $\nu = 1, \ldots, n$.*

(See, e.g., Gantmacher [13].)
A simplified version of this theorem needs only half the determinants:

Theorem 7.6. *Suppose that all the coefficients of the polynomial* (7.17) *are positive. For odd n, the polynomial is Hurwitz if and only if the Hurwitz determinants $H_2, H_4, \ldots, H_{n-1}$ are positive. For even n, the polynomial is Hurwitz if and only if the Hurwitz determinants $H_3, H_5, \ldots, H_{n-1}$ are positive.*

(See, e.g., Gantmacher [13].)
Consider as an example the amplitude feedback introduced in Sect. 7.3. The denominator of the closed-loop transfer function reads

$$a_0 + a_1 s + a_2 s^2 + s^3$$

with

$$a_0 = 1.6129 \cdot 10^{15}\,\mathrm{s}^{-3}, \quad a_1 = 7.6901 \cdot 10^{10}\,\mathrm{s}^{-2}, \quad a_2 = 4.5205 \cdot 10^5\,\mathrm{s}^{-1}.$$

In the following, the physical units of these coefficients will be ignored to avoid confusion with the Laplace variable s. Since all coefficients are positive, this polynomial with $n = 3$ is Hurwitz, because

$$H_2 = \det \begin{pmatrix} a_2 & a_0 \\ 1 & a_1 \end{pmatrix} = a_2 a_1 - a_0 = 3.3150 \cdot 10^{16} > 0.$$

Now assume that the feedback gain K_c in the loop of Fig. 7.4 is a free parameter. As a consequence, the coefficients a_0 and a_1 become parameter-dependent:

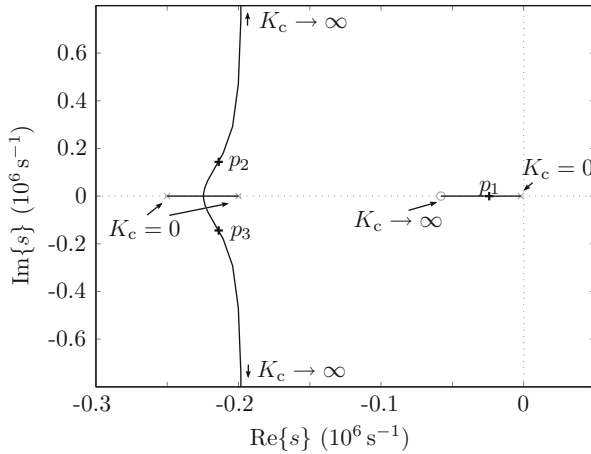$$a_0 = 1.0136 \cdot 10^{14} K_c + 1.0263 \cdot 10^{14}, \quad a_1 = 1.7435 \cdot 10^9 K_c + 5.0924 \cdot 10^{10}.$$

Fig. 7.5 Root locus of the amplitude feedback. The poles p_1, p_2, and p_3 are obtained for $K_c = 14.9$

The Hurwitz criterion now leads to the conditions

$$a_0 > 0 \quad \Rightarrow \quad K_c > -1.01,$$
$$a_1 > 0 \quad \Rightarrow \quad K_c > -29.21,$$
$$H_2 > 0 \quad \Rightarrow \quad K_c > -33.37.$$

Thus, the feedback loop is stable for $K_c > -1.01$. Due to the stability of the open-loop system, the closed-loop system obviously remains stable even if the feedback gain is slightly negative. A positive feedback gain K_c, however, is the typical case for the amplitude control. Figure 7.5 shows the closed-loop poles in the complex s-plane as a function of the positive gain $K_c > 0$. This type of diagram is also referred to as a the **root locus**. For $K_c = 0$, the closed-loop poles are equal to the open-loop poles

$$p_1(0) = -\frac{1}{T_{c2}}, \quad p_2(0) = -\frac{1}{T_{det}}, \quad p_3(0) = -\frac{1}{T_{cav}}$$

that are obtained from the open-loop transfer function (cf. Fig. 7.4)

$$H_{open}(s) = \frac{\Delta \hat{V}_{gap,det}(s)}{\Delta \hat{V}_e(s)}.$$

For increasing K_c, the closed-loop pole p_1 moves to the left toward the open-loop zero

$$z_1(0) = -\frac{1}{T_{c1}}$$

of $H_{\text{open}}(s)$, whereas the poles p_2 and p_3 approach each other and for a certain K_c between 0 and 14.9, a complex conjugate pole pair arises. The root locus indicates that the closed loop remains stable also for higher $K_c \to \infty$, because all three branches of the root locus remain in the OLHP. Since the branches of the root locus are the positions of the closed-loop poles,[5] the closed loop is stable. This is in agreement with the result of the Hurwitz criterion.

Please note that for a practical implementation, very large feedback gains K_c would not be recommendable for several reasons:

- For sufficiently large gains, the complex pair $p_{2,3}$ dominates the dynamics of the loop, resulting in an unacceptable oscillatory behavior.
- Large gains may increase disturbances, especially the measured noise.
- The feedback of the real system may become unstable for very large gains due to unmodeled high-frequency dynamics and delays.

7.4.2 Bode Plots and Nyquist Criterion

The Hurwitz stability criterion is based on the characteristic equation, i.e., on the denominator polynomial of the closed-loop transfer function. The **Bode plots** and the **Nyquist criterion** are approaches that are different in the sense that they rely on the open-loop transfer function

$$H_{\text{open}}(s) := H_c(s)\ H_p(s)\ H_m(s)$$

of the standard feedback loop; cf. Fig. 7.2. Consider as an example the system

$$H_{\text{open}}(s) = \frac{K\left(1 - \frac{s}{z_1}\right)}{s^N \left(1 - \frac{s}{p_1}\right)\left(1 - \frac{s}{p_2}\right)\left(1 - \frac{s}{p_2^*}\right)}. \tag{7.19}$$

This system is assumed to have a real zero $z_1 \neq 0$, a real pole $p_1 \neq 0$, a complex pole pair p_2 and p_2^*, and N poles at $s = 0$. The frequency response of $H_{\text{open}}(s)$ is given by

$$H_{\text{open}}(j\omega) = \frac{K\left(1 - \frac{j\omega}{z_1}\right)}{(j\omega)^N \left(1 - \frac{j\omega}{p_1}\right)\left(1 - \frac{j\omega}{p_2}\right)\left(1 - \frac{j\omega}{p_2^*}\right)}.$$

---

[5]The root locus is usually obtained by a numerical calculation of the closed-loop poles for different values of the gain.

The complex pole pair can also be written as

$$\left(1 - \frac{j\omega}{p_2}\right)\left(1 - \frac{j\omega}{p_2^*}\right) = 1 - \frac{\omega^2}{p_2 p_2^*} - j\omega\left(\frac{1}{p_2} + \frac{1}{p_2^*}\right) = 1 - \frac{\omega^2}{|p_2|^2} - j\omega\frac{2\mathrm{Re}\{p_2\}}{|p_2|^2}.$$

In a Bode diagram, the amplitude and phase of H_{open} are plotted versus the frequency $\omega > 0$. A logarithmic scale is used, which has the advantage that the multiplication of two transfer functions is equivalent to the sum of their Bode diagrams. The amplitude of H_{open} in decibels (dB) is calculated as

$$|H_{\mathrm{open}}(j\omega)|_{\mathrm{dB}} := 20 \log_{10} |H_{\mathrm{open}}(j\omega)|.$$

In our example, using the properties of the logarithmic function leads to

$$|H_{\mathrm{open}}(j\omega)|_{\mathrm{dB}} = 20 \log_{10} |K| + 20 \log_{10} \sqrt{1 + \frac{\omega^2}{z_1^2}} - 20N \log_{10} \omega -$$

$$- 20 \log_{10} \sqrt{1 + \frac{\omega^2}{p_1^2}} - 20 \log_{10} \sqrt{\left(1 - \frac{\omega^2}{|p_2|^2}\right)^2 + \left(\frac{2\omega\mathrm{Re}\{p_2\}}{|p_2|^2}\right)^2}.$$

This expression is the sum of five components. The first is the constant

$$|H_1(j\omega)|_{\mathrm{dB}} := 20 \log_{10} |K|.$$

The second function is due to the zero and can be approximated by two asymptotes:

$$|H_2(j\omega)|_{\mathrm{dB}} := 20 \log_{10} \sqrt{1 + \frac{\omega^2}{z_1^2}} \approx \begin{cases} 0\,\mathrm{dB} & \text{for } \omega \ll |z_1|, \\ 3\,\mathrm{dB} & \text{for } \omega = |z_1|, \\ 20 \log_{10} \omega - 20 \log_{10} |z_1| & \text{for } \omega \gg |z_1|. \end{cases}$$

$$(7.20)$$

The N-fold integrator leads to

$$|H_3(j\omega)|_{\mathrm{dB}} := -20N \log_{10} \omega.$$

For the pole p_1, the result is similar to the case of zero z_1, but with opposite signs:

$$|H_4(j\omega)|_{\mathrm{dB}} := -20 \log_{10} \sqrt{1 + \frac{\omega^2}{p_1^2}} \approx \begin{cases} 0\,\mathrm{dB} & \text{for } \omega \ll |p_1|, \\ -3\,\mathrm{dB} & \text{for } \omega = |p_1|, \\ -20 \log_{10} \omega + 20 \log_{10} |p_1| & \text{for } \omega \gg |p_1|. \end{cases}$$

Finally, the pole pair has the following asymptotes:

$$|H_5(j\omega)|_{\text{dB}} := -20\log_{10}\sqrt{\left(1 - \frac{\omega^2}{|p_2|^2}\right)^2 + \left(\frac{2\omega\text{Re}\{p_2\}}{|p_2|^2}\right)^2}$$

$$\approx \begin{cases} 0\,\text{dB} & \text{for } \omega \ll |p_2|, \\ -40\log_{10}\omega + 40\log_{10}|p_2| & \text{for } \omega \gg |p_2|. \end{cases}$$

The phase of H_{open} is given by

$$\measuredangle H_{\text{open}}(j\omega) = \sum_{i=1}^{5} \measuredangle H_i(j\omega).$$

The phases $\measuredangle H_i(j\omega)$ can be approximated by asymptotes in a similar way as shown for the amplitudes. For example, the zero leads to the phase

$$\measuredangle H_2(j\omega) = \measuredangle\left(1 - \frac{j\omega}{z_1}\right) = \begin{cases} \approx 0 & \text{for } \omega \to 0, \\ +\frac{\pi}{4} & \text{for } \omega = |z_1| \text{ and } z_1 < 0, \\ \approx +\frac{\pi}{2} & \text{for } \omega \to \infty \text{ and } z_1 < 0, \\ -\frac{\pi}{4} & \text{for } \omega = |z_1| \text{ and } z_1 > 0, \\ \approx -\frac{\pi}{2} & \text{for } \omega \to \infty \text{ and } z_1 > 0. \end{cases}$$

Figure 7.6 shows the Bode plots of the transfer functions $H_i(j\omega)$ with their asymptotes for a system with $N = 1$, positive gain K, and with the zero and poles in the OLHP, i.e., a stable system. The following observations can be made:

- The gain $H_1(j\omega) = K$ leads to an amplitude shift of the open-loop transfer function H_{open}.
- The zero $z_1 > 0$ raises the amplitude and phase; cf. $H_2(j\omega)$. At the frequency $\omega = |z_1|$, the amplitude is close to 3 dB, and the phase equals $\pi/4$. For large frequencies, the amplitude increases with 20 dB per (frequency) decade and the phase approaches $\pi/2$.
- The amplitude of the integrator $H_3(j\omega)$ tends to infinity for small frequencies. This fact enables steady-state accuracy for the closed loop with regard to stepwise disturbances. However, the phase of $-\pi/2$ may lead to stability problems in some cases. This can be shown with the Nyquist stability criterion, which will be presented below.
- The pole p_1 has the opposite effect to that of the zero z_1. For large frequencies, the amplitude slope is -20 dB per decade, and the phase approaches $-\pi/2$.
- For small or large frequencies, the complex pole pair acts as a double pole at $\omega = |p_2|$. However, for frequencies close to $|p_2|$, a resonance may occur. This means that $|H_5(j\omega)|$ may become considerably larger than 1. The frequency at
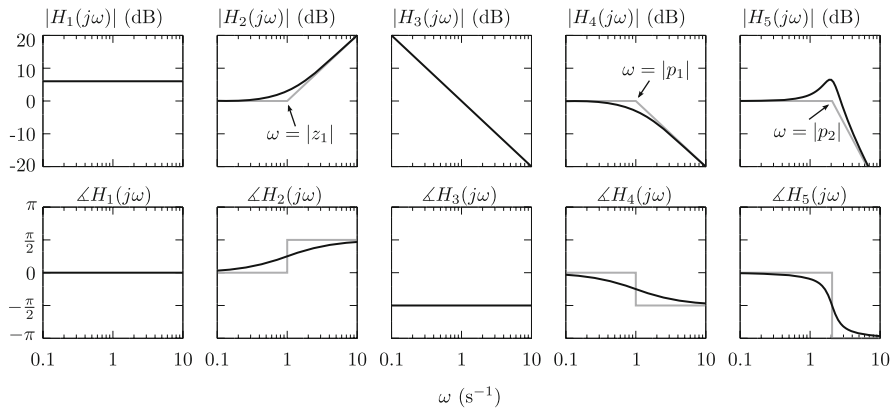
Fig. 7.6 Bode plots of the open-loop transfer function (7.19) for $N = 1$, $K = 2$, $z_1 = -1$, $p_1 = -1$, and $p_2 = -\frac{1}{2} + 2j$

which the maximum of $|H_5(j\omega)|$ occurs can be calculated analytically, and it reads

$$\omega_{\text{res}} = \sqrt{\text{Im}\{p_2\}^2 - \text{Re}\{p_2\}^2} \approx 1.94,$$

i.e., $\text{Im}\{p_2\} > \text{Re}\{p_2\}$ is a necessary condition for a resonance. Disturbances or input signals with frequencies close to ω_{res} will be amplified significantly in the open loop. A resonance in the open loop may be one reason why feedback is necessary. Feedback can provide additional damping, so that the resonance is not present in the closed-loop frequency response.

For the Bode plot of the system with the transfer function H_{open}, the Bode plots of the subsystems H_i have to be combined. As already shown, this simply corresponds to the sum of the amplitude and phase plots due to the use of a logarithmic scale. This also applies to the asymptotes. To sketch the asymptotes of the Bode plot of H_{open}, it is therefore possible to proceed as follows. First, the break points are calculated as the absolute value of the zeros and poles, i.e., $\omega = |z_i(0)|$ and $\omega = |p_i(0)|$. The argument 0 for both z_i and p_i emphasizes that the open-loop zeros and poles are used. Next, one begins with the asymptote of the N-fold integrator H_3. This asymptote is a line with slope $-20N$ dB per decade (of the frequency ω) that crosses the point with amplitude $20\log_{10}(K)$ at $\omega = 1\,\text{s}^{-1}$. For $N = 0$, the Bode plot begins with a horizontal asymptote. One then proceeds to higher frequencies, changing the slope of the asymptote at every break point. For a single pole, the slope changes by -20 dB per decade; for a single zero, by 20 dB per decade; and for multiple poles or zeros, accordingly with the multiple of these slopes. For the phase plot, one begins with a horizontal asymptote of $-N\frac{\pi}{2}$. At the break points, the asymptote is changed stepwise with $\frac{-\pi}{2}$ for a single pole, $\frac{\pi}{2}$ for a zero, and a multiple of $\frac{\pi}{2}$ for multiple poles or zeros. For the amplitude feedback, this procedure leads to
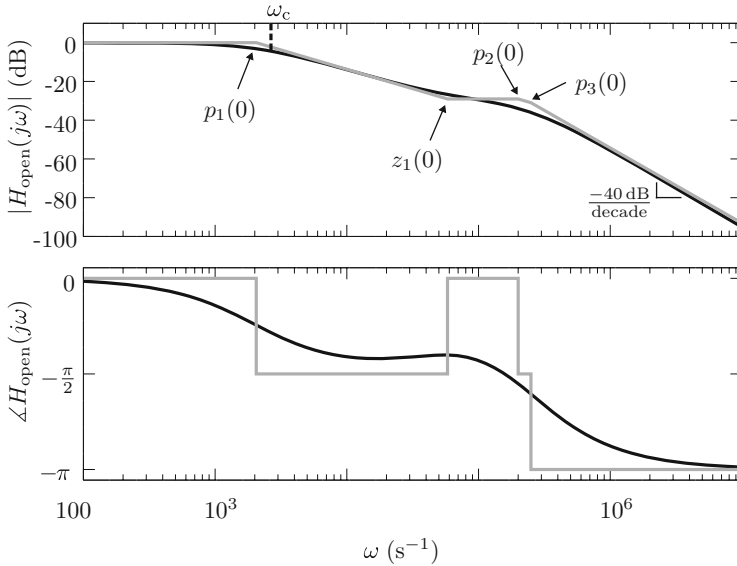
Fig. 7.7 Bode plot of the amplitude feedback example

the asymptotes as shown in Fig. 7.7 for $K_c = 1$. The exact Bode plot is shown as a solid black curve. The static open-loop gain equals

$$H_{\text{open}}(j\omega = 0) = K_{\text{mod}} G_{\text{Vgain}} K_{\text{Vgain}} R_p K_{\text{cd}} = 0.99$$

for $K_c = 1$. At $\omega = p_1(0)$, the first pole leads to a negative slope of $-20\,\text{dB}$ per decade. Next, the zero $z_1(0)$ raises the slope to zero, before the two remaining poles finally lead to a slope or **cutoff rate** of $-40\,\text{dB}$ per decade. The phase begins at zero and drops to

$$-\frac{\pi}{2} + \frac{\pi}{2} - \frac{\pi}{2} - \frac{\pi}{2} = -\pi$$

for large frequencies.

The frequency at which the amplitude drops by $-3\,\text{dB}$ is called the **cutoff frequency**. It is denoted by $\omega_c = 2004\frac{1}{\text{s}}$ in Fig. 7.7 and is also called the **bandwidth** of the open-loop transfer function [1].

Because the Bode plot contains all information about the open loop, there is a unique correspondence between this diagram and the transfer function $H_{\text{open}}(s)$. If the open loop is stable, the Bode plot can be obtained by measuring the frequency response $H_{\text{open}}(j\omega)$. An equivalent diagram that is very useful for determining the stability of the closed loop is the **Nyquist plot**. It is obtained by plotting the curve

$$H_{\text{open}}(j\omega) = \text{Re}\{H_{\text{open}}(j\omega)\} + j\,\text{Im}\{H_{\text{open}}(j\omega)\}$$
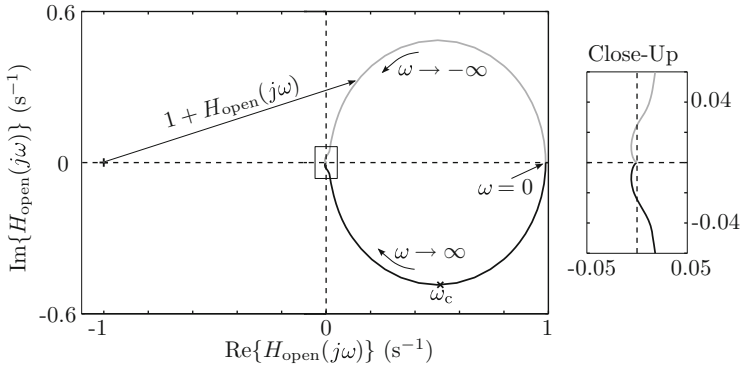
Fig. 7.8 Nyquist plot of the amplitude feedback example

in the complex plane for $\omega \in \mathbb{R}$. The Nyquist plot of the amplitude feedback example is shown in Fig. 7.8.

Due to

$$H_{\text{open}}(-j\omega) = H^*_{\text{open}}(j\omega),$$

the part of the Nyquist plot for negative frequencies ω is always axially symmetric to the part for positive frequencies. For this reason, the Nyquist plot is usually analyzed for only positive frequencies. From the discussion of the Bode plot, it is already known that the Nyquist plot begins at $H_{\text{open}}(j0) = 0.99$ and approaches the origin for large ω. Also, the phase approaches $-\pi$, as can be observed from the closeup view in Fig. 7.8. The vector

$$1 + H_{\text{open}}(j\omega)$$

points from $-1 + j0$ to the Nyquist plot, as shown in Fig. 7.8. Its behavior is essential for the stability of the closed loop. If we follow this vector from $\omega = 0$ to $\omega \to \infty$, we can define the change of its argument as
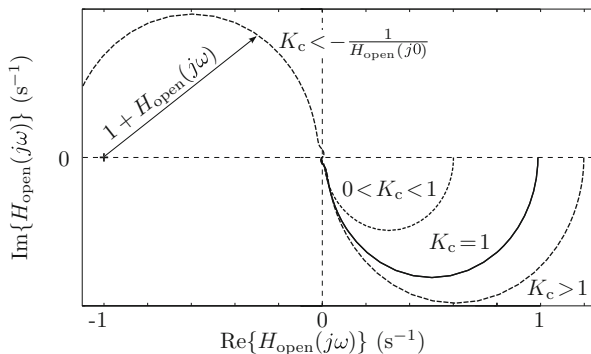
$$\Delta\varphi_{\text{Nyquist}} := \lim_{\omega \to \infty} \measuredangle\left(1 + H_{\text{open}}(j\omega)\right) - \measuredangle\left(1 + H_{\text{open}}(j0)\right). \tag{7.21}$$

The **general Nyquist stability criterion** can now be used to determine the stability of the closed loop:

Theorem 7.7. *The closed loop is asymptotically stable if and only if the continuous change of the argument as defined in Eq. (7.21) is equal to*

$$\Delta\varphi_{\text{Nyquist}} = n_{\text{unstable}}\pi + n_{\text{critical}}\frac{\pi}{2},$$

Fig. 7.9 Nyquist plot for different gains K_c

where n_{unstable} *is the number of (unstable) open-loop poles in the ORHP and* n_{critical} *is the number of open-loop poles on the imaginary axis.*

(See, e.g., Unbehauen [14, p. 156].)

Only the continuous change in the argument is considered. If, for example, the Nyquist plot consists of several branches due to open-loop poles on the imaginary axis, then $\Delta\varphi_{\text{Nyquist}}$ can be determined for each branch separately, and the total change is the sum of these results.

Since the amplitude feedback system in our example contains only stable open-loop poles, a necessary and sufficient condition for stability is

$$\Delta\varphi_{\text{Nyquist}} = 0,$$

as is the case for $K_c = 1$ in Fig. 7.8. Changing the gain K_c will only scale the Nyquist plot, as shown in Fig. 7.9. For positive gains $K_c > 0$, the closed loop will always be stable, because $\Delta\varphi_{\text{Nyquist}} = 0$. In the case of negative K_c, the Nyquist plot is also rotated by 180°, and the critical point $-1 + j0$ is crossed for

$$K_c = -\frac{1}{H_{\text{open}}(j0)} = -1.01$$

and the change in the argument is $\Delta\varphi_{\text{Nyquist}} = +\pi$. Thus, the closed loop is unstable for $K_c < -1.01$, a result already obtained with the Hurwitz criterion.

7.4.3   Time Delay

If the feedback loop contains a considerable time delay T_d, this can be taken into account in the Laplace transform of the open loop $H_{\text{open}}(s)$. If, for example, the measurement of the output $y(t)$ is delayed, this leads to

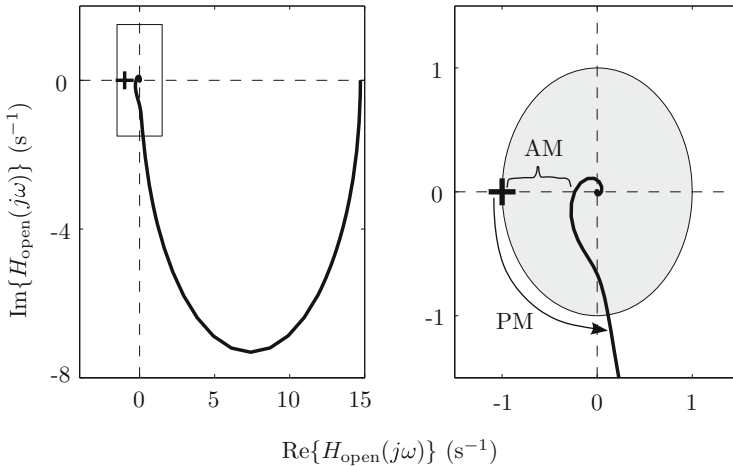$$y_{\text{delay}}(t) = y(t - T_d).$$

Fig. 7.10 Nyquist plot (*left*) and closeup (*right*) of the amplitude feedback with delay $T_d = 5\,\mu s$ and definition of the amplitude margin (*AM*) and phase margin (*PM*)

Due to the shift theorem of the Laplace transform, every open loop with a single delay can be expressed by

$$H_{\text{open,delay}}(s) = H_{\text{open}}(s)\, e^{-T_d s}.$$

The consequence of the exponential function is that the characteristic equation of the closed loop is no longer an algebraic equation, but a transcendental one. The number of poles becomes infinite, and the stability analysis is thus more involved. Fortunately, the Nyquist criterion can still be applied [15]. For the frequency response,

$$|H_{\text{open,delay}}(j\omega)| = |H_{\text{open}}(j\omega)|$$

$$\measuredangle H_{\text{open,delay}} = \measuredangle H_{\text{open}} - T_d \omega$$

holds, i.e., the delay leads to a faster decrease of the phase, but does not affect the amplitude. Figure 7.10 shows the Nyquist plot of the amplitude feedback with the nominal feedback gain of $K_c = 14.9$ and an additional time delay of $T_d = 5\,\mu s$. This time delay is a worst-case scenario for signal transit times due to a distance of about 100 m between the cavity and the LLRF unit [12]. The closeup shows that the closed loop is still stable, but not for arbitrary $K_c > 0$. The Nyquist plot crosses the horizontal axis at -0.237. Increasing the gain K_c by a factor of

$$AM = 20 \log\left(\frac{1}{0.237}\right) = 12.5\,\text{dB}$$

will therefore lead to a crossing of the critical point $-1 + j0$ and to instability. This factor is called the **amplitude margin** and is a measure for variations in the amplitude of the process transfer function that can be tolerated. For larger amplitude margins, the feedback is more robust against such variations. In addition, Fig. 7.10 shows that the Nyquist plot crosses the unit circle at an angle of about $-83°$. The frequency of this crossing is $\omega = 34.2 \cdot 10^3 \, \text{s}^{-1}$. The **phase margin**

$$PM = 180° - 83° = 97°$$

is defined as the distance to the critical point in terms of the phase, i.e., the tolerable variation in the phase of the process transfer function. A simple estimate[6] shows that an additional time delay of $T_d = 50 \, \mu\text{s}$ would lead to a phase decrease of

$$\omega T_d \approx 34.2 \cdot 10^3 \, \text{s}^{-1} \cdot 50 \, \mu\text{s} \approx 98°,$$

i.e., the feedback will remain stable for time delays up to this order of magnitude.

7.4.4   Steady-State Accuracy

The standard closed loop in Fig. 7.2 on p. 341 is said to have no **steady-state error** if

$$x_e(\infty) := \lim_{t \to \infty} x_e(t) = \lim_{t \to \infty} (y_r(t) - y_m(t)) = 0$$

is guaranteed, i.e., if the measured value converges to the reference value. From Fig. 7.2, the following expression for the steady-state error can be obtained:

$$X_e(s) = \frac{1}{1 + H_p(s)H_c(s)H_m(s)} \, Y_r(s) - \frac{H_p(s)H_m(s)}{1 + H_p(s)H_c(s)H_m(s)} \, X_{d1}(s) -$$

$$- \frac{H_m(s)}{1 + H_p(s)H_c(s)H_m(s)} \, (X_{d2}(s) + X_{d3}(s)). \tag{7.22}$$

In the following, it is assumed that all transfer functions in this expression are stable, i.e., have only poles in the OLHP. In this case, we can use the final-value theorem for Laplace transforms (cf. Sect. 2.2). Without disturbances, this leads to

$$x_e(\infty) = \lim_{s \to 0} \left(s \, X_e(s) \right) = \lim_{s \to 0} \left(\frac{s \, Y_r(s)}{1 + H_p(s)H_c(s)H_m(s)} \right) = \lim_{s \to 0} \left(\frac{s \, Y_r(s)}{1 + H_{\text{open}}(s)} \right).$$

---

[6]Because the amplitude does not depend on the time delay, the crossing of the unit circle always occurs at the same frequency.

It is now particularly important which type of reference signal $y_r(t)$ is assumed. For a step function, we have $Y_r(s) = K/s$ and[7]

$$x_e(\infty) = \begin{cases} \frac{K}{1+H_{open}(0)} & \text{for } |H_{open}(0)| < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

This shows that an integrator $(1/s)$ in the feedback loop—in the controller, the process, or the measurement transfer function—is sufficient for a vanishing steady-state error. For other reference signals, this may not be sufficient. For example, a ramp signal $(1/s^2)$ requires at least two integrators in the transfer functions of the feedback loop. However, too many integrators may lead to stability problems, because each integrator lowers the phase of the open-loop transfer function by $-\pi/2$.

   If significant disturbances are present, it is usually necessary that the integrator be contained in the controller, as can be seen from the other transfer functions in Eq. (7.22). Assuming that the process and measurement transfer functions have no integrator, $H_p(0)$ and $H_m(0)$ are finite, and an integral controller will lead to $x_e(\infty) = 0$ for stepwise disturbances.

7.5   Feedback Design

### 7.5.1   Tradeoff Between Performance and Robustness

The transfer function $H_p(s)$ in Fig. 7.2 on p. 341 usually describes the physical behavior of the real process only approximately. Reasons for model errors can be nonlinearities, dependence on time or operating conditions, and unmodeled high-frequency dynamics. In many cases, the model errors may be described by parameter variations in the numerator and denominator of the transfer function $H_p(s)$. These variations will lead to a change in performance of the closed-loop control. To estimate this effect, the **sensitivity function**

$$H_s := \frac{\partial H_{ry}}{\partial H_p} \frac{H_p}{H_{ry}}$$

is defined as the relative change of the closed-loop transfer function $H_{ry}(s)$ with respect to variations of the process transfer function $H_p(s)$. With Eq. (7.15), this leads to

---

[7]Note that $1 + H_{open}(0) = 0$ is impossible, since that would imply that $s = 0$ would be a pole, and this has been excluded by considering stable transfer functions.

Fig. 7.11 Sensitivity
function of the amplitude
feedback ($T_d = 0$)
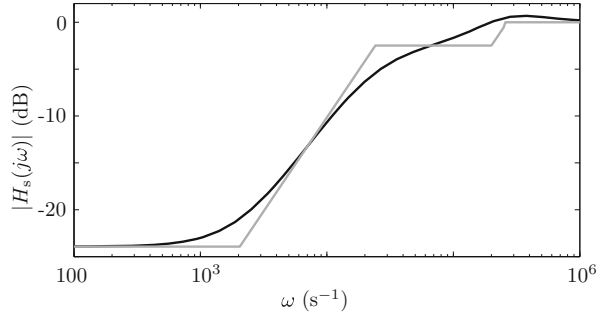
$$H_s = \frac{\partial \left(\frac{H_p H_c}{1 + H_p H_c H_m} \right)}{\partial H_p} \frac{H_p}{H_{ry}} = \frac{H_c(1 + H_p H_c H_m) - H_p H_c^2 H_m}{(1 + H_p H_c H_m)^2} \frac{H_p(1 + H_p H_c H_m)}{H_p H_c}$$

and finally to the sensitivity function

$$H_s(s) = \frac{1}{1 + H_p(s) H_c(s) H_m(s)}.$$

This is exactly the disturbance-to-output transfer function $H_{dy}(s)$ (cf. Eq. (7.16))
that was derived from Fig. 7.2. It is apparent that a sufficiently large feedback gain
$|H_c|$ will lead to both a small sensitivity $|H_s|$ and a good disturbance rejection.
However, a large feedback decreases the amplitude margin AM in many cases
and may lead to instability. This shows that a tradeoff between performance and
robustness specifications is usually necessary. Please note that for the open-loop
system, $H_c = 0$, and the sensitivity equals 1. For the closed-loop system, $|H_s|$ also
approaches 1 for large frequencies, because for most practical cases, $|H_p H_c H_m|$
tends to zero.

For our amplitude feedback example, the sensitivity function is equal to

$$H_s(s) = \frac{(s - z_1)(s - z_2)(s - z_3)}{(s - p_1)(s - p_2)(s - p_3)}$$

with

$$z_1 = -2.5 \cdot 10^5 \, \text{s}^{-1}, \quad z_2 = -2 \cdot 10^5 \, \text{s}^{-1}, \quad z_3 = -2.05 \cdot 10^3 \, \text{s}^{-1},$$
$$p_1 = -2.42 \cdot 10^4 \, \text{s}^{-1}, \quad p_{2,3} = (-2.14 \pm j \, 1.44) \cdot 10^5 \, \text{s}^{-1}.$$

Its amplitude $|H_s(j\omega)|$ is shown in Fig. 7.11. In contrast to $|H_{\text{open,delay}}(j\omega)|$, the
amplitude of the sensitivity function depends on the time delay.

The sensitivity shows that the amplitude feedback rejects disturbances or noise
with frequency components up to about 10 kHz. The closed loop is also less
sensitive with respect to model variations than the open loop in this frequency range.

However, the sensitivity is not zero for $\omega \to 0$. This implies that the closed loop does not reject DC offsets completely and may thus have a steady-state error. This can be shown as follows. From the standard feedback loop, the control error can be calculated as

$$X_e(s) = Y_r(s) - H_m(s) \cdot Y(s)$$

$$= Y_r(s) - H_m(s) \cdot \frac{H_c(s)\,H_p(s)}{1 + H_m(s)\,H_c(s)\,H_p(s)} \cdot Y_r(s)$$

$$= \frac{1}{1 + H_m(s)\,H_c(s)\,H_p(s)} \cdot Y_r(s).$$

If we assume that the closed loop is stable and the reference signal is equal to a unit step, i.e., $Y_r = 1/s$, then the final value of the control error is given by

$$\lim_{t \to \infty} x_e(t) = \lim_{s \to 0} \left(s \cdot \frac{1}{1 + H_m(s)\,H_c(s)\,H_p(s)} \cdot \frac{1}{s} \right)$$

$$= \frac{1}{1 + H_m(0)\,H_c(0)\,H_p(0)}.$$

Thus, the value of the sensitivity function for $\omega = 0$ is equal to the relative steady-state error of the closed-loop system. For the amplitude feedback loop, a value of 6.4%, or $-23.9\,\text{dB}$, is obtained. This steady-state error will also be apparent in the simulation results in the next section.

7.5.2 Design Goals and Specifications

The main design goals of feedback are stability, a fast dynamic response, disturbance rejection, a small tracking error, and robustness against parameter variations. In addition, the control effort should comply with the physical limitations of the process. There exist several parameters to describe these specifications quantitatively. In the time domain, the response to a step disturbance or reference signal is often considered, and the following quantities are used to describe the dynamic response:

- **Rise time**: transit time from 10% to 90% of the final value, i.e., of the output step size.
- Percentage of **overshoot**.
- **Settling time**: time after which the output stays inside a $\pm 5\%$ or $\pm 2\%$ interval around the final value.
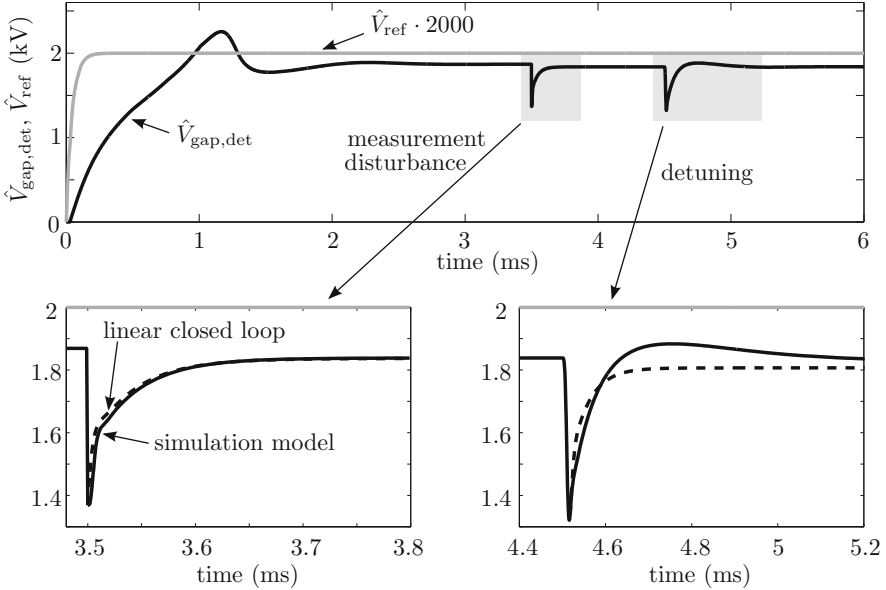- **Steady-state error** between the reference signal and the output.

Fig. 7.12 Performance of the amplitude feedback

The performance of the amplitude feedback example is shown in Fig. 7.12. The curve $\hat{V}_{gap,det}$ is obtained from a simulation model from [12], which is in good agreement with measurements. The reference signal $\hat{V}_{ref}$ is initially raised from zero to 1 V. Due to a prefilter with a time constant of 43 μs, the reference signal is raised not stepwise, but smoothly. The simulation model includes not only the amplitude feedback, but also a resonance frequency feedback to ensure that the cavity is in resonance. At the beginning of the simulation, the resonance frequency feedback has to settle and has a strong coupling with $\hat{V}_{gap}$. At $t \approx 3$ ms, both feedback loops have reached their equilibrium.

The amplitude feedback is excited at $t = 3.5$ ms with a stepwise disturbance of the measurement $\hat{V}_{gap,det}$. The dynamic response of the simulation model is compared to the response of the linear closed loop $H_{ry}(s)$ with $T_d = 0$ (Fig. 7.12, bottom left). This shows that the transfer function $H_{ry}(s)$ describes the behavior very well for small deviations from equilibrium. From the simulation results, a rise time of 73 μs, a 5% settling time of 103 μs, and a steady-state error of 6.4% are obtained.

At $t = 4.5$ ms, the cavity is detuned, so that the gap voltage drops by about 0.5 kV. This time, the simulation model shows a different behavior due to the interaction of the resonance frequency feedback with the amplitude feedback. This demonstrates that nested control loops are dynamically coupled in general. If the coupling is strong, it is necessary to take this into account during the analysis

and design of the feedback. Nested control loops can be described by MIMO or multivariable control systems [16].

In addition to the mentioned parameters, there also exist specifications in the frequency domain:

- **Resonant peak**: the maximum of the closed-loop frequency response $|H_{\mathrm{ry}}(j\omega)|$ indicates relative stability and is recommended to be between 1.1 and 1.5 [1].
- **Bandwidth**: the frequency at which $|H_{\mathrm{ry}}(j\omega)|$ has decreased by $-3\,\mathrm{dB}$ with respect to the zero-frequency value.
- **Cutoff rate**: the slope of $|H_{\mathrm{ry}}|$ at high frequencies.
- **Amplitude margin** and **phase margin** (cf. Sect. 7.4.3): an *AM* larger than $6\,\mathrm{dB}$ and a *PM* between $30°$ and $60°$ are regarded as a good tradeoff between robustness and performance [1].

In our example, the bandwidth of $H_{\mathrm{ry}}(j\omega)$ equals $30.3 \cdot 10^3\,\mathrm{s}^{-1}$ (which corresponds to $\Delta f = 4831\,\mathrm{Hz}$), and the cutoff-rate is $-20\,\mathrm{dB/decade}$.

7.5.3   PID Control

A general proper **PID control** algorithm is given by

$$H_{\mathrm{c}}(s) = \frac{U(s)}{X_{\mathrm{e}}(s)} = K_{\mathrm{P}} + K_{\mathrm{I}}\frac{1}{s} + K_{\mathrm{D}}\frac{s}{T_{\mathrm{D}}s + 1};$$

it is a combination of a proportional, an integral, and a derivative controller. The transfer function can also be written as

$$H_{\mathrm{c}}(s) = \frac{(K_{\mathrm{P}}T_{\mathrm{D}} + K_{\mathrm{D}})s^2 + (K_{\mathrm{P}} + K_{\mathrm{I}}T_{\mathrm{D}})s + K_{\mathrm{I}}}{s(T_{\mathrm{D}}s + 1)}; \tag{7.23}$$

it has two zeros and two poles. A pure derivative is obtained for $T_{\mathrm{D}} = 0$. However, this leads to an improper transfer function. In the time domain, the controller is described by the differential equation

$$T_{\mathrm{D}}\dot{u}(t) + u(t) = (K_{\mathrm{P}}T_{\mathrm{D}} + K_{\mathrm{D}})\dot{x}_e(t) + (K_{\mathrm{P}} + K_{\mathrm{I}}T_{\mathrm{D}})x_e(t) + K_{\mathrm{I}}\int_0^t x_{\mathrm{e}}(\tau)\,\mathrm{d}\tau.$$

In steady state, the control error x_{e} must be zero due to the integration.

The controller of the amplitude feedback example is of PDT$_1$ type. This can be shown as follows. A general **PDT$_1$ controller** can be written as

$$H_{\mathrm{c}}(s) = K_{\mathrm{P}} + \frac{K_{\mathrm{D}}s}{T_{\mathrm{D}}s + 1} = K_{\mathrm{P}}\frac{s\left(T_{\mathrm{D}} + \frac{K_{\mathrm{D}}}{K_{\mathrm{P}}}\right) + 1}{s\,T_{\mathrm{D}} + 1}.$$

With

$$K_c = K_P, \quad T_{c1} = T_D + \frac{K_D}{K_P}, \quad T_{c2} = T_D,$$

we obtain the amplitude controller that is shown in Fig. 7.4.

To design a general PID controller, it is necessary to determine the four degrees of freedom K_P, K_D, K_I, and T_D so that the specifications are met. If the open-loop system is stable, the two zeros of $H_c(s)$ may be used to compensate open-loop poles. The time constant T_D should not be chosen too small, because that would amplify high-frequency noise.

Several so-called **tuning rules** exist for the design of PI and PID controllers [5]. A simple tuning rule is described in [16] that is based on the approximation of the process transfer function with a first-order model

$$H_{\text{approx}}(s) = \frac{K}{T \cdot s + 1} \, e^{-T_d \cdot s},$$

with the gain K, the time constant T, and a time delay T_d. For a PI controller, the tuning rule is (cf. [16, p. 57])

$$K_P = \frac{1}{K} \frac{T}{T_{\text{tune}} + T_d}, \qquad K_I = \frac{K_P}{\min\{T, 4(T_{\text{tune}} + T_d)\}},$$

with a single tuning parameter T_{tune}. A small value of this parameter will lead to fast output performance, whereas a large value implies a high robustness and smaller values of the input. A typical tradeoff is the choice $T_{\text{tune}} = T_d$.

This tuning rule can be applied to the amplitude feedback loop example. From Fig. 7.4, the open-loop transfer function

$$\frac{\Delta \hat{V}_{\text{gap,det}}(s)}{\Delta \hat{V}_c(s)} = \frac{K_{\text{mod}} G_{\text{Vgain}} K_{\text{Vgain}} R_p K_{\text{cd}}}{(s \, T_{\text{cav}} + 1)(s \, T_{\text{det}} + 1)}$$

is obtained. For this type of transfer function, the following first-order approximation may be used; cf. [16, p. 58]:

$$K = K_{\text{mod}} G_{\text{Vgain}} K_{\text{Vgain}} R_p K_{\text{cd}} = 0.9877, \qquad T = T_{\text{det}} + \frac{1}{2} T_{\text{cav}} = 7 \, \mu\text{s},$$

$$T_d = \frac{1}{2} T_{\text{cav}} = 2 \, \mu\text{s}.$$

With $T_{\text{tune}} = T_d$ as the choice of the tuning parameter, the coefficients of the resulting PI controller are

$$K_P = 1.7718, \qquad K_I = 2.5312 \cdot 10^5 \, \frac{1}{\text{s}}.$$

The settling time of the linear amplitude feedback with this controller is $16.4\,\mu s$ for a 5% interval around the set point. This is considerably faster than the PDT_1 controller. Furthermore, the PI controller leads to a zero steady-state error. Note, however, that for the design in this section, we have neglected any interaction of the amplitude loop with the resonance frequency feedback loop.

For the practical implementation of a PID controller, some issues should be taken into account. If the process is stable, it is often sufficient to use a PI controller. Derivative action, i.e., $K_D \neq 0$, will lead to an increased sensitivity with respect to measurement noise. If the reference signal $y_r(t)$ contains steps and a derivative action is needed, it is usually better to use the measured output $y_m(t)$ as input of the derivative part of the controller instead of the control error $x_e(t)$; cf. [16, p. 56] and [5, p. 317]. One challenge for the integral action is the so-called **integrator windup** [5], a nonlinear effect.

We can illustrate this effect by means of Fig. 7.3. We assume that the controller has integral action and generates a value that exceeds the constraints of the subsequent saturation function. In this case, the output of the feedback will be a constant value as long as the saturation function is active. This may be interpreted as a feedback loop that is no longer closed, because the output of the controller does not depend on the control error. The integral controller will, however, continue to integrate the control error, and this may result in a poor overall feedback performance. Measures that prevent windup are known as **antiwindup**.

7.5.4 Stability Issues for Nonlinear Systems

As described in Sect. 7.1.3, almost every practical feedback system is, in fact, a nonlinear system

$$\frac{d\vec{x}(t)}{dt} = \vec{v}_1(\vec{x}(t), \vec{u}(t)), \tag{7.24a}$$

$$\vec{y}_m(t) = \vec{v}_2(\vec{x}(t)), \tag{7.24b}$$

where $\vec{y}_m$ is the output vector with the measured quantities of the process. A common approach is to calculate the linearization

$$\frac{d\Delta\vec{x}(t)}{dt} = A \cdot \Delta\vec{x}(t) + B \cdot \Delta\vec{u}(t), \tag{7.25a}$$

$$\Delta\vec{y}_m(t) = C \cdot \Delta\vec{x}(t) \tag{7.25b}$$

of the system for a certain equilibrium and to use it for the analysis or design of a linear controller so that the closed-loop behavior is stable. This approach has also been chosen in the previous sections. An important question that now arises is whether the linear controller will also be able to stabilize the nonlinear system.

The stability theory of Lyapunov that was described in Sect. 2.8.5 is useful to obtain some conclusions concerning this question. In order to use the theory of Lyapunov, it is necessary to analyze the feedback loop in the time domain, because the frequency domain approach is in general not applicable to nonlinear systems.

Consider first a very general linear controller in state-space representation

$$\frac{\mathrm{d}\vec{x}_{\mathrm{c}}(t)}{\mathrm{d}t} = A_{\mathrm{c}} \cdot \vec{x}_{\mathrm{c}}(t) + B_{\mathrm{c}} \cdot \vec{x}_{\mathrm{e}}(t), \tag{7.26a}$$

$$\vec{u}_{\mathrm{c}}(t) = C_{\mathrm{c}} \cdot \vec{x}_{\mathrm{c}}(t) + D_{\mathrm{c}} \cdot \vec{x}_{\mathrm{e}}(t), \tag{7.26b}$$

where $\vec{x}_{\mathrm{e}} = \Delta\vec{y}_{\mathrm{r}} - \Delta\vec{y}_{\mathrm{m}}$ denotes the vector with measured control errors, $\vec{u}_{\mathrm{c}}$ is the actuator value that can be used as input to the process (i.e., $\Delta\vec{u} = \vec{u}_{\mathrm{c}}$), and $\vec{x}_{\mathrm{c}}$ contains the internal states of the controller. This type of controller is also known as a **dynamic output feedback**, because the controller has a dynamic structure and it uses the output vector $\vec{y}_{\mathrm{m}}$ as the only information about the process. This type of controller also contains the PID controller as a special case: rewriting the transfer function (7.23) as the sum of a constant and a remaining polynomial leads to

$$H_{\mathrm{c}}(s) = \frac{U_{\mathrm{c}}(s)}{X_{\mathrm{e}}(s)} = \left(K_{\mathrm{P}} + \frac{K_{\mathrm{D}}}{T_{\mathrm{D}}}\right) + \frac{\left(K_{\mathrm{I}} - \frac{K_{\mathrm{D}}}{T_{\mathrm{D}}^2}\right)s + \frac{K_{\mathrm{I}}}{T_{\mathrm{D}}}}{s^2 + \frac{1}{T_{\mathrm{D}}}s}.$$

Using the results of Sect. 7.1.2 and taking the additional direct feedthrough into account leads to the following state-space representation of the controller:

$$\frac{\mathrm{d}\vec{x}_{\mathrm{c}}(t)}{\mathrm{d}t} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{T_{\mathrm{D}}} \end{bmatrix} \cdot \vec{x}_{\mathrm{c}}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot x_{\mathrm{e}}(t),$$

$$u_{\mathrm{c}}(t) = \begin{bmatrix} \frac{K_{\mathrm{I}}}{T_{\mathrm{D}}} & \left(K_{\mathrm{I}} - \frac{K_{\mathrm{D}}}{T_{\mathrm{D}}^2}\right) \end{bmatrix} \cdot \vec{x}_{\mathrm{c}}(t) + \left(K_{\mathrm{P}} + \frac{K_{\mathrm{D}}}{T_{\mathrm{D}}}\right) \cdot x_{\mathrm{e}}(t).$$

This is a dynamic output feedback. Note that the case of a pure derivative controller ($T_{\mathrm{D}} = 0$) is not included in this representation.[8] Due to Eq. (7.26), the transfer function of the controller can be obtained by

$$H_{\mathrm{c}}(s) = C_{\mathrm{c}} \cdot (sI - A_{\mathrm{c}})^{-1} \cdot B_{\mathrm{c}} + D_{\mathrm{c}}.$$

Connecting the controller (7.26) with system (7.25) (i.e., by $\Delta\vec{u} = \vec{u}_{\mathrm{c}}$) leads directly to the following dynamics of the closed loop:

---

[8]This is, however, not a serious limitation, since a pure derivative would be both undesirable in the presence of noise and is not realizable on any physical hardware.

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\Delta\vec{x}(t)\\ \vec{x}_{\mathrm{c}}(t)\end{bmatrix}=\underbrace{\begin{bmatrix}A-B\cdot D_{\mathrm{c}}\cdot C & B\cdot C_{\mathrm{c}}\\ -B_{\mathrm{c}}\cdot C & A_{\mathrm{c}}\end{bmatrix}}_{A_{\mathrm{cl}}}\cdot\begin{bmatrix}\Delta\vec{x}(t)\\ \vec{x}_{\mathrm{c}}(t)\end{bmatrix}+\begin{bmatrix}B\cdot D_{\mathrm{c}}\\ B_{\mathrm{c}}\end{bmatrix}\cdot\Delta\vec{y}_{\mathrm{r}}(t).$$

$$(7.27)$$

We assume that the controller is designed properly, so that the closed-loop dynamics are stable. According to the results of Sect. 7.1.5, this is the case if A_{cl} has only eigenvalues with negative real parts.

After the controller design, the controller will be connected to the real nonlinear process. One possible choice for the input of the nonlinear system (7.24) is then

$$\vec{u}(t)=\vec{u}_{\mathrm{F}}+\vec{u}_{\mathrm{c}}(t)=\vec{u}_{\mathrm{F}}+C_{\mathrm{c}}\cdot\vec{x}_{\mathrm{c}}(t)+D_{\mathrm{c}}\cdot\vec{x}_{\mathrm{e}}(t),$$

where $\vec{u}_{\mathrm{F}}$ is a feedforward value that equals the input value at the equilibrium point $\vec{x}=\vec{x}_{\mathrm{F}}$. In other words,

$$\vec{v}_1(\vec{x}_{\mathrm{F}},\vec{u}_{\mathrm{F}})=0$$

is assumed, and the controller has only to correct deviations from the equilibrium. The control error is now given by

$$\vec{x}_{\mathrm{e}}=\vec{y}_{\mathrm{r}}-\vec{y}_{\mathrm{m}}=\vec{y}_{\mathrm{r}}-\vec{v}_2(\vec{x}).$$

These choices of the closed-loop connection lead to the following dynamics:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\vec{x}(t)\\ \vec{x}_{\mathrm{c}}(t)\end{bmatrix}=\begin{bmatrix}\vec{v}_1(\vec{x}(t),\vec{u}(t))\\ A_{\mathrm{c}}\cdot\vec{x}_{\mathrm{c}}(t)+B_{\mathrm{c}}\cdot(\vec{y}_{\mathrm{r}}(t)-\vec{v}_2(\vec{x}(t)))\end{bmatrix}. \qquad (7.28)$$

A linearization around $\vec{x}=\vec{x}_{\mathrm{F}}$, $\vec{u}=\vec{u}_{\mathrm{F}}$, $\vec{x}_{\mathrm{c}}=0$, and $\vec{y}_{\mathrm{r}}=\vec{v}_2(\vec{x}_{\mathrm{F}})$ leads to the same linear dynamics as Eq. (7.27). This is reasonable, because it means that the same result is obtained either by linearizing the nonlinear closed-loop dynamics or by using the linearization (7.25) of the open-loop system (7.24) to obtain the linear closed-loop model (7.27).

We already assumed that Eq. (7.27) is stable, and we can now use theorem 2.18. For $\Delta\vec{y}_{\mathrm{r}}=0$ and the previous assumption of a strictly stable matrix A_{cl} (the real parts of all eigenvalues are negative), the theorem can be applied to Eq. (7.28), and the consequence is a stable equilibrium of the nonlinear setup. This is an important motivation for using linear control design in many cases, even for systems that are practically nonlinear.

Note, however, that the linear system (7.27) is asymptotically stable in the global sense, i.e., for arbitrary initial values, whereas in general, the asymptotic stability of the nonlinear system (7.28) is given only in a local neighborhood around the equilibrium. This neighborhood, also called a **region of attraction**, may be so small that from a practical point of view, the equilibrium is in fact unstable. The size of

the region of attraction can be estimated using Lyapunov functions as defined in Sect. 2.8.5.

A nonzero reference value $\Delta \vec{y}_r \neq 0$ acts as an excitation. As long as it is not too large, the closed loop will be stable.

If further disturbances act on the system (7.24) or the model is inaccurate, this may lead to a steady-state error. In most cases, an integral controller will help to avoid such an error. A pure integral controller can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} x_c(t) = x_e(t),$$

$$u_c(t) = K_I x_c(t).$$

Therefore, $A_c = 0$, $B_c = 1$, $C_c = K_I$, and $D_c = 0$. The closed-loop dynamics for a SISO system are then

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \vec{x}(t) \\ x_c(t) \end{bmatrix} = \begin{bmatrix} \vec{v}_1(\vec{x}(t), u_F + K_I\, x_c(t)) \\ y_r - v_2(\vec{x}(t)) \end{bmatrix},$$

and from the bottom row, we have the equilibrium

$$\frac{\mathrm{d}}{\mathrm{d}t} x_c = 0 \quad \Rightarrow \quad y_r = v_2(\vec{x}) = y_m,$$

and the steady-state error will therefore tend to zero for stepwise reference signals.

References

1. W.S. Levine (ed.), *The Control Handbook: Control System Fundamentals*, 2nd edn. (CRC Press/Taylor & Francis Group, West Palm Beach/London, 2011)
2. J.W. Polderman, J.C. Willems, *Introduction to the Mathematical Theory of Systems and Control* (Springer, New York, 1998)
3. J. Doyle, B. Francis, A. Tannenbaum, *Feedback Control Theory* (Macmillan, New York, 1990)
4. E.D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems.* Textbooks in Applied Mathematics, Number 6, 2nd edn. (Springer, New York, 1998)
5. K.J. Aström, R.M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers* (Princeton University Press, Princeton, 2012)
6. P. Baudrenghien, Low-level RF, in *CAS - CERN Accelerator School: RF for Accelerators*, Ebeltoft, 8–17 June 2010, pp. 341–367
7. D. Lens, Modeling and control of longitudinal single-bunch oscillations in heavy-ion synchrotrons. Fortschrittberichte VDI, Reihe 8, Mess-, Steuerungs- und Regelungstechnik; Nr. 1209; Dissertation, Technische Universität Darmstadt, 2012
8. K.J. Aström, B. Wittenmark, *Computer-Controlled Systems* (Prentice Hall, Englewood, 1997)
9. O. Föllinger, *Laplace- und Fourier-Transformation* (Hüthig, Heidelberg, 1990)
10. O. Föllinger, *Regelungstechnik: Einführung in die Methoden und ihre Anwendung* (Hüthig Buch, Heidelberg, 1990)
11. G. Ludyk, *Theoretische Regelungstechnik 1* (Springer, Berlin, 1995)

12. U. Hartel, Modellierung des Regelungs- und Steuerungssystems einer Beschleunigungseinheit für Synchrotrons. Diplomarbeit, Technische Universität Darmstadt, Darmstadt, 2011
13. F.R. Gantmacher, *Matrizentheorie* (Deutscher Verlag der Wissenschaften, Berlin, 1986)
14. H. Unbehauen, *Regelungstechnik I*, 15. Auflage (Vieweg+Teubner Verlag, Wiesbaden, 2008)
15. O. Föllinger, Zur Stabilität von Totzeitsystemen, Regelungstechnik, S. 145–149 (1967)
16. S. Skogestad, I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design* (Wiley, London, 2005)