Chapter 11
Sampled Data Control Systems

In practice, most of the results delivered by control theory are realized by digital computers equipped with appropriate real-time facilities. Depending on how many control loops are implemented for a given application, the digital controller can be realized by various devices ranging from single-chip microcontrollers via single board controllers to PLCs or industrial PCs. The reliable network technology available today allows system developers to implement the controllers in a distributed topology, as well.

The digital realization of control algorithms also reflects the contemporary available computing technology. The control devices applied in industry integrate a number of open-loop and closed-loop control components as a single compact digital unit.

A simple scheme of a sampled data control system is shown in Fig. 11.1. Most of the processes control engineers deal with are continuous in nature. In this chapter it will be assumed that the control signal applied to the process (control input), as well as the process variable (output signal), are both continuous-time (CT) signals. While the input and output signals of a process are assumed to be continuous (called 'analog' in practice), the digital processing assumes the data is available in discrete-time (DT) form as a sequence of numbers. Consequently, the analog world represented by physical signals should be interfaced to the world of data used by digital computations. These interfaces are the *sampler* transforming the analog signals to discrete ones, and the *holder* transforming the discrete signals to analog ones. In practice, these interfaces are typically implemented by *analog to digital* (A/D) and *digital to analog* (D/A) converters, respectively. As far as Fig. 11.2 is concerned, observe that a real-time clock governs the operation of the digital computer to control the sampling and holding in a synchronized way. To distinguish the analog and discrete versions of the signals involved in sampled data control systems, the following notation will be used:

L. Keviczky et al., *Control Engineering*, Advanced Textbooks in Control and Signal Processing, https://doi.org/10.1007/978-981-10-8297-9_11

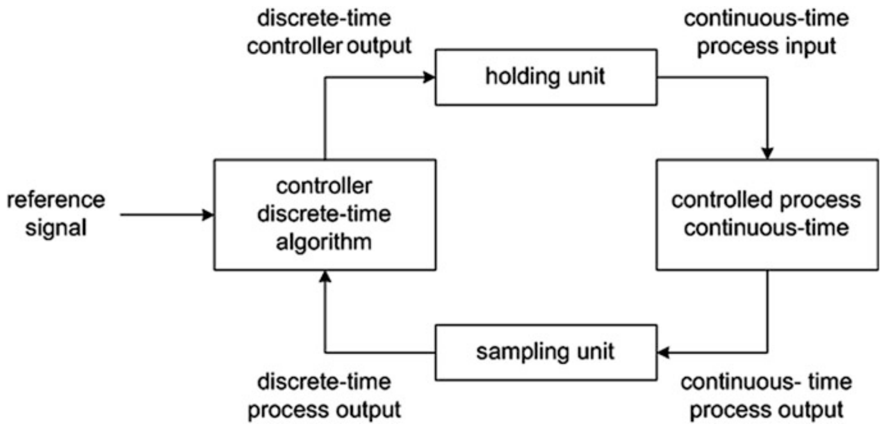


Fig. 11.1 Schematic diagram of closed-loop sampled data control systems

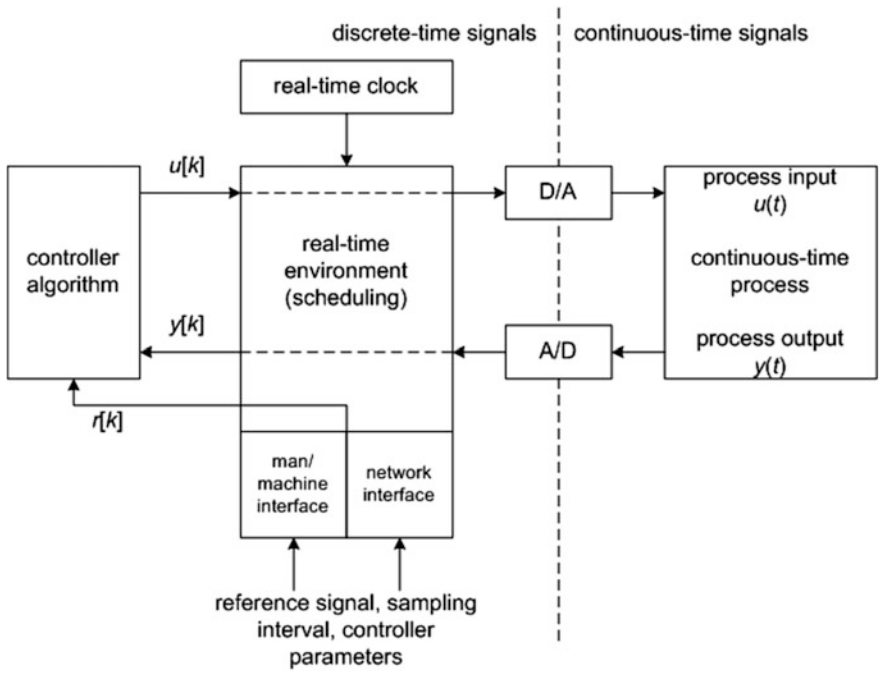


Fig. 11.2 Detailed set-up of sampled data control systems

- $x(t)$: CT signal
- $x[k] = x(kT_s)$: DT signal, where T_s denotes the sampling time and kT_s with $k = 0, 1, 2, \ldots$ assigns the sampling instants.

Comparing the sampled data control system in Fig. 11.2 to a CT closed-loop control system it can be seen that the control signal $u(t)$ is produced by a CT regulator $C(s)$, while the sequence of the control signal $u[k]$ is produced by a digital control algorithm running in a real-time digital environment. At each sampling instant the digital regulator carries out the following actions:

- Receive the sampled process variable and transmit the digitized data to the sampled data control algorithm.
- Receive the set value previously adjusted at a Man-Machine Interface (e.g., typed in) or delivered by a communication network (as a result of a calculation performed at a higher hierarchy level).
- Realize the digital control algorithm to calculate the digital control input $u[k]$ and send this digital value to the D/A converter.

The progress of the digital technologies surrounding the digital controller (intelligent sensors and actuators, advanced Man-Machine Interface devices, a wide range of cheap, though powerful and reliable network technologies) indicate that digital controllers will dominate the field over CT controllers in the future. A comparison of the continuous and digital control technologies can be summarized in favor of digital controllers as follows:

- The digital technology applied is more reliable and cheaper.
- Its flexibility is superior considering both the implementation and the variety of the control algorithms.
- Possible modifications and/or extensions are far easier to accomplish.
- Accuracy is kept constant over a long period of time.
- There are easy ways to deliver the set point value for the controller, to overwrite the controller's parameters, as well as to monitor the controller's operation.

There are, however, a few issues requiring special care:

- Between two samples the control system is left to operate in open-loop.
- The sampling rate should carefully be selected to be in harmony with the dynamics of the process and to comply with the capabilities of the real-time environment (performance and the number representation employed).
- The output of the digital controller (control signal) must be interpolated from a digital sequence to a CT function, thus the waveform of the control signal is restricted.
- Sampling introduces additional difficulties for the design (dead-time and unwanted dynamics).

11.1 Sampling

From the operation of a CT process information can be collected by observing the CT input/output signals of the process. In case this information is elaborated by a digital device, these CT signals will be represented by their samples. Assuming equidistant sampling (constant sampling time), an important result of systems theory helps us to decide whether a CT signal can be reconstructed from its samples or not. Namely, for band limited CT signals, SHANNON's sampling theorem requires that at least two samples must be available from the highest frequency component of the band-limited CT signal for a reconstruction. Concerning the importance of SHANNON's sampling theorem, we refer to the fact that this theorem constitutes the basis for the usability of a flexible, programmable, digital environment in CT signal processing (visualization, analysis in the frequency domain, etc.).

As was mentioned earlier, sampling is physically performed by A/D and D/A converters governed by the real-time clock of the controller. The A/D converter is driven by an analog signal $x(t)$ and produces a sequence $x[k]$ in a coded digital form. The operation of an A/D converter is commonly symbolized by a periodically closed switch (Fig. 11.3).

For various applications an A/D converter is selected according to several of the parameters attached to the converters. The speed of the conversion is characterized by the conversion time, which can be even less than $1\mu s$. Further parameters are the noise rejection and the resolution specifying the number of bits used for the digitized code (typically a value from 8 to 16).

Using again the notation introduced earlier, a sampling according to $f[k] = f(kT_s)$ is also called mathematical sampling. The sequence $f[k]$ can be derived by *impulse modulation* according to

$$f^*(t) = \sum_{n=-\infty}^{\infty} f(nT_s)\delta(t - nT_s), \tag{11.1}$$

where $f^*(t)$ is a sequence of DIRAC-impulses, and $f[k]$ is a sequence of numbers made out from the area of these DIRAC-impulses. Figure 11.4 shows the impulse modulation supposing $f(nT_s) \equiv 0$, $(n<0)$. Applying mathematical sampling, fundamental analysis can be carried out in the frequency domain. The analysis also points out the necessity of appropriate sampling according to the sampling theorem.

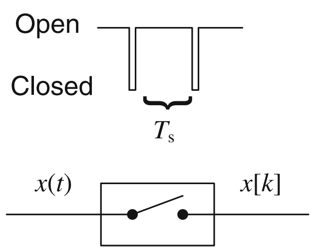


Fig. 11.3 Periodically controlled switch symbolizing the operation of the A/D converter

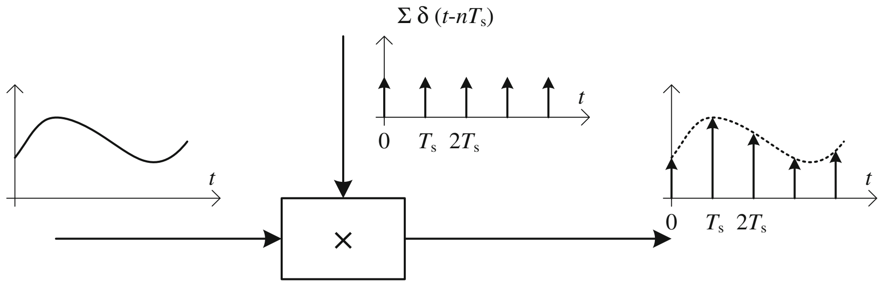


Fig. 11.4 Impulse modulation

Consider a CT signal $f(t)$ with a frequency spectrum of $F(j\omega)$. The theorem of signals and systems specifies the frequency spectrum of $f^*(t)$ to be

$$F^*(j\omega) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} F(j\omega + jk\omega_s), \tag{11.2}$$

where $\omega_s = 2\pi/T_s$ is the sampling radian frequency.

It is seen that the Fourier transform of the sampled signal is a sum of the side frequency components exhibited by $F(j\omega + jk\omega_s)$. The frequency folding phenomenon is avoided if $\omega_s \geq 2\omega_{max}$ (or $\omega_s/2 \geq \omega_{max}$) holds, where ω_{max} is the maximum frequency of the band limited $f(t)$ CT signal. The frequency $\omega_s/2$ is of great importance for the spectrum of the sampled CT signal, and is also called the Nyquist *frequency*. The phenomenon is presented by the Figs. 11.5, 11.6 and 11.7, where the spectrum of the CT signal, the spectrum of an appropriately sampled DT signal, as well as the spectrum of the inappropriately sampled signal are shown, respectively. Figure 11.8 explains how a non-existing (*alias*) signal may appear as a consequence of a slow sampling rate. In this example the slow sampling rate applied to the high frequency sinusoidal signal produces samples, which can be interpreted as samples of a non-existing, low frequency sinusoidal component. In a closed-loop control system, control actions to compensate this low frequency component are unnecessary and would only induce an additional disturbance in the closed-loop. To avoid aliasing, a low pass filter (also called an anti-aliasing filter) must be placed between the measured output signal and the A/D converter. The fundamental component of the spectrum of the DT signal is

$$F^*(j\omega) = \frac{1}{T_s} F(j\omega). \tag{11.3}$$

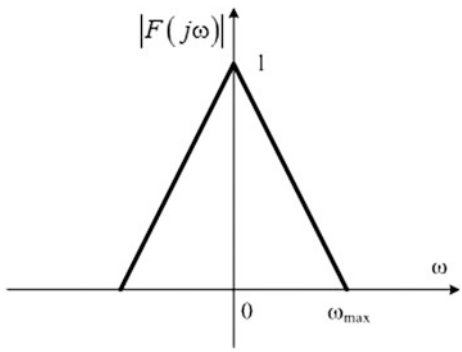


Fig. 11.5 Spectrum of a *CT* signal

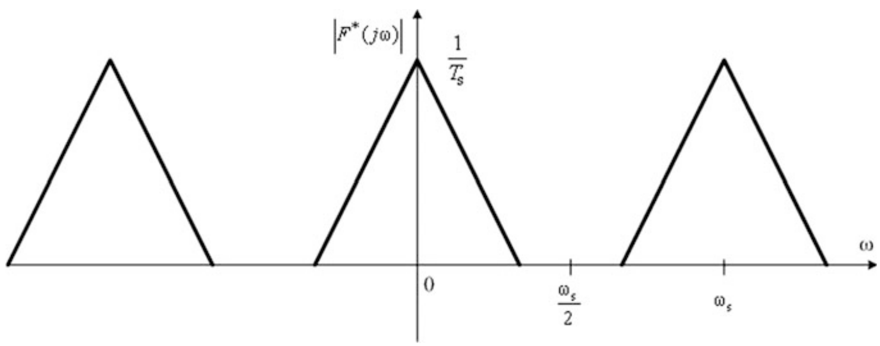


Fig. 11.6 Spectrum of a DT signal assuming an appropriate sampling rate

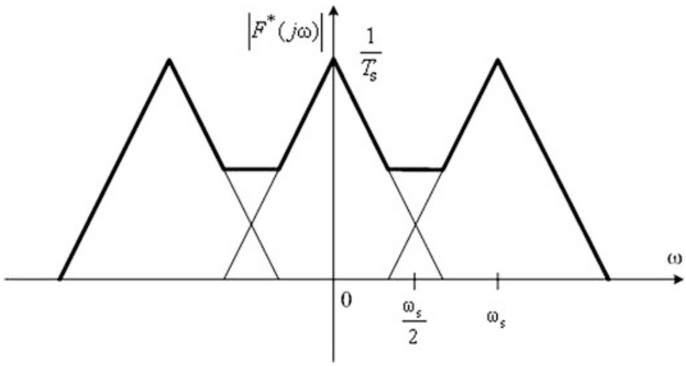


Fig. 11.7 Spectrum of a DT signal assuming an inappropriately low sampling rate

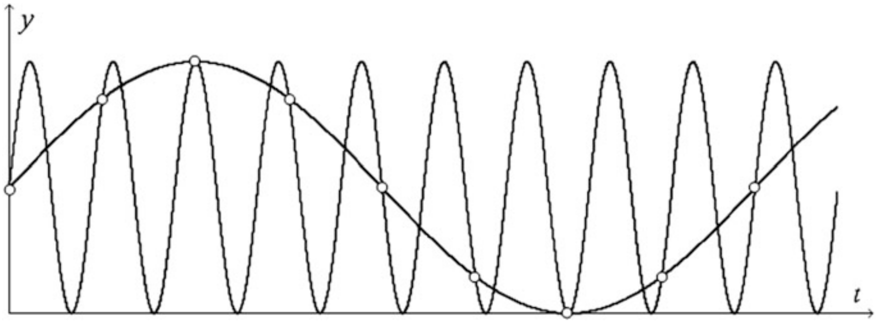
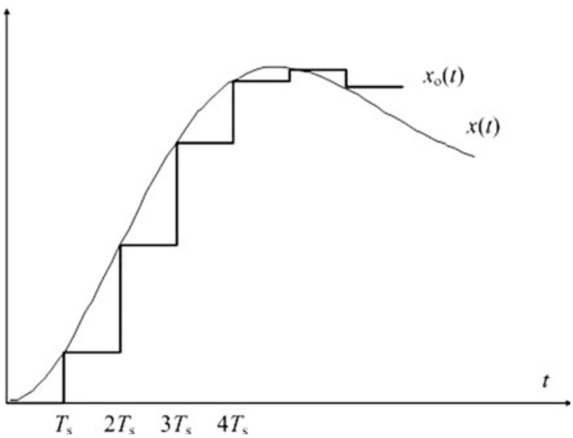


Fig. 11.8 Appearance of the non-existing component according to the low sampling rate

11.2 Holding

Signals in a sampled data control loop form a hybrid system simultaneously presenting CT and DT signals. From a systems engineering point of view a, CT process and the resulting DT signal processing in the controller should be interfaced to each other. Sampling performs the continuous-discrete transformation. However, a functional unit to perform the discrete-continuous transformation should inevitably be inserted into the loop to ensure closed-loop operation. The holding unit is also a controlled element to receive and decode a digital input signal, as well as to produce a CT approximation between two samples. As far as the nature of the approximation is concerned (constant, linear, quadratic, etc.), there is no general requirement. However, the easy realization of the constant approximation (*zero-order holding*) has become a practical standard. Applying a zero-order holding (ZOH) unit results in a staircase waveform for the CT output (see Fig. 11.9). Note that **MATLAB**® offers the application of the *stairs* function to perform a ZOH discrete-continuous transformation.

Fig. 11.9 Application of a ZOH unit



As mentioned earlier, the signal dynamics between two sampling instants is not limited to be a constant, but it could follow a first- or second-order course. In the case of first-order holding, the output of the holding unit is determined by a straight line defined by the last two sampled data (see Fig. 11.10). As in practice zero order holding satisfies all the requirements, holding with higher-order approximations will not be considered in the sequel.

The mathematical description of the ZOH unit should comply with the operation shown by Fig. 11.11. Consider the impulse response of the ZOH unit: $1(t) - 1(t - T_s)$. As the unit step response is produced by an integrator driven by a Dirac impulse and the delay by T_s can be taken into account by a transfer function of e^{-sT_s}, Fig. 11.12 allows seeing that the transfer function of the ZOH unit is

$$W_{\mathrm{ZOH}}(s) = \frac{1}{s} - \frac{e^{-sT_s}}{s} = \frac{1 - e^{-sT_s}}{s} \tag{11.4}$$

Once its transfer function is known, the frequency function of the ZOH unit can easily be determined (see A.11.3 of Appendix A.11.1). Some fundamental frequency domain properties of the ZOH unit, however, can be discovered using the following Taylor series approximation in the low-frequency domain:

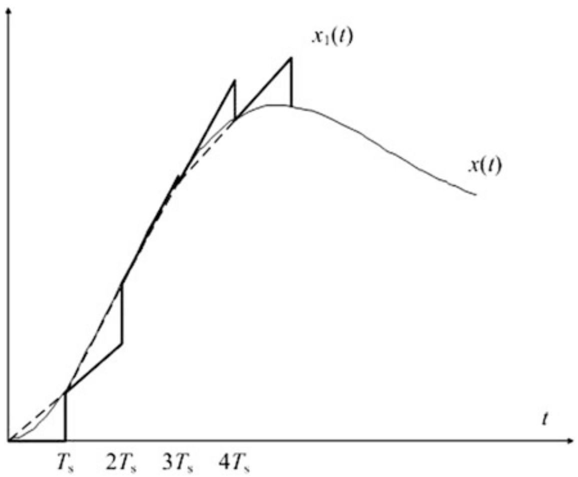
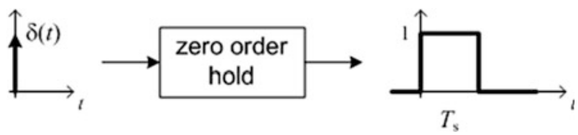


Fig. 11.10 Application of a first order holding unit

Fig. 11.11 Impulse response of the ZOH unit



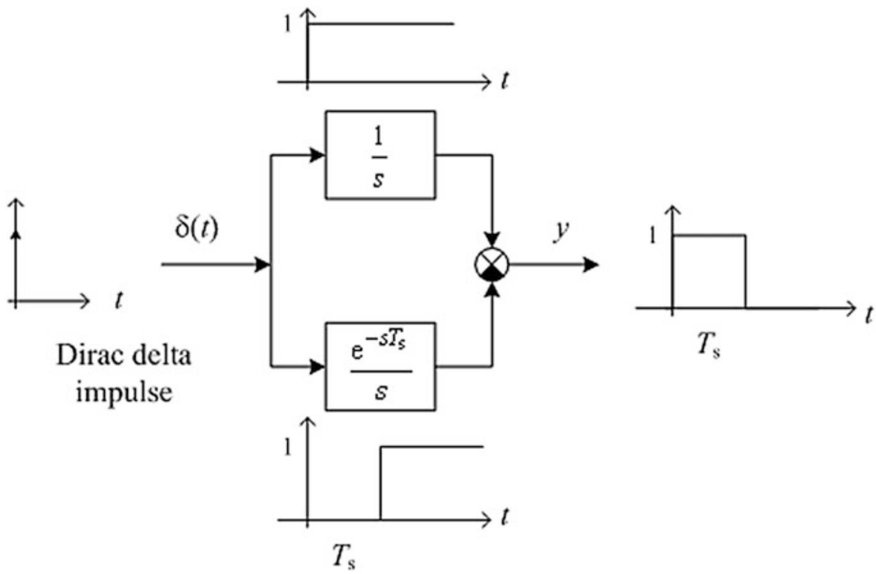


Fig. 11.12 Impulse response components of the ZOH unit

$$
\begin{aligned}
W_{\mathrm{ZOH}}(s) &= \frac{1-e^{-sT_{\mathrm{s}}}}{s} \cong \frac{1-\left(1-sT_{\mathrm{s}}+\frac{s^2T_{\mathrm{s}}^2}{2}-\cdots\right)}{s} \approx T_{\mathrm{s}}\left(1-\frac{sT_{\mathrm{s}}}{2}+\cdots\right)\\
&\approx T_{\mathrm{s}}e^{-sT_{\mathrm{s}}/2}
\end{aligned}
\tag{11.5}
$$

According to the discussion of the spectral properties of sampled data systems in Sect. 11.1, the frequency function of a CT system built up as a series connection of a ZOH unit and a CT process given by a transfer function $H(s)$ can well be approximated as

$$
H_{\mathrm{d}}(j\omega) = \frac{1}{T_{\mathrm{s}}}\tilde{H}(j\omega) \approx \frac{1}{T_{\mathrm{s}}}T_{\mathrm{s}}e^{-j\omega T_{\mathrm{s}}/2}H(j\omega) = e^{-j\omega T_{\mathrm{s}}/2}H(j\omega), \tag{11.6}
$$

where $\tilde{H}$ is the approximate transfer function of the joint ZOH unit and the given CT system. The low-frequency approximation indicates that the ZOH unit inserts a delay by $T_{\mathrm{s}}/2$ into the loop. As pointed out earlier, the appearance of any delay in the loop is an unwanted effect in the control loop both considering the stability of the closed-loop system and the quality of its transient response. It is to be emphasized, however, that the application of a holding unit is an absolutely inevitable element in the control loop to interface the DT and CT signal domains.

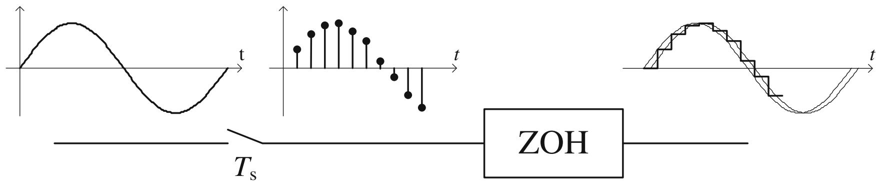


Fig. 11.13 Scheme to demonstrate the delay of the ZOH unit

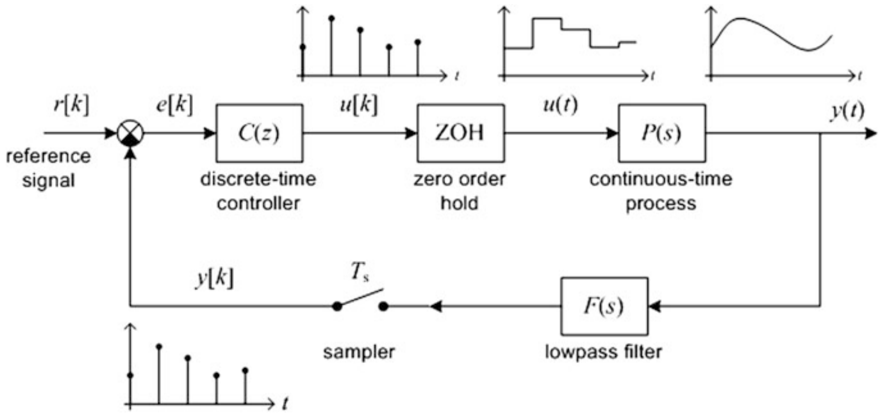


Fig. 11.14 Detailed block diagram of a closed-loop sampled data control system

The physical arrangement in Fig. 11.13 verifies the validity of the above approximation. Samples of a sinusoidal CT signal drive a ZOH unit. The output of the ZOH unit is a staircase signal whose fundamental harmonic component has also been drawn together with the original sinusoidal CT signal. The exact phase delay between these two sinusoidal signals depends on the sampling time. However, the approximation according to a delay of $T_s/2$ is rather convincing.

Selecting the sampling time is certainly a new aspect of DT design in comparison with simple CT systems. This issue will be discussed later on. However, selecting a sampling frequency ω_s, a Nyquist frequency of $\omega_N = \omega_s/2$ is also selected, which will suggest the application of a proper low-pass filter to avoid frequency folding. That is, the low-pass filter should pass a highest frequency component ω_{max} to satisfy the $\omega_N \geq \omega_{max}$ condition. A detailed block diagram of a closed-loop sampled data control system is drawn in Fig. 11.14. The time-domain behavior of the signals involved are also shown.

11.3 Description of Discrete-Time Signals, the z-Transformation and the Inverse z-Transformation

The *z-transformation* is a widely used way to describe DT signals and systems. Given a sequence $f[k]$ $(k = 0, 1, 2, \ldots)$ of data, its z-transform is defined by the infinite series

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = f[0] + z^{-1} f[1] + z^{-2} f[2] + \cdots, \tag{11.7}$$

where z is the complex valued operator of the transformation. Note that $f[k]$ is assumed to be a positive-time-function, i.e., $f[k] \equiv 0$, $(k < 0)$. Though $\mathcal{Z}\{f[k]\}$ is a function of z^{-1}, in practice the notation $F(z) = \mathcal{Z}\{f[k]\}$ is used for z-transforms.

The *region of convergence* (*ROC*) in the complex plane is given by a circle of radius R_1. The infinite series in (11.7) is assumed to be convergent outside of the circle given by R_1, i.e., $F(z) = f[0] + z^{-1} f[1] + z^{-2} f[2] + \cdots$ is convergent for $|z| > R_1$.

Discussing CT closed-loop control systems, the application of the Laplace transformation turned out to be a very useful tool both for analysis and design, provided that inverse Laplace transformation capabilities are also available. In a similar way inverse z-transformation techniques are to be developed for DT system analysis and design. An analytical expression for the inverse z-transformation is given by the following integral:

$$f[k] = \mathcal{Z}^{-1}\{F(z)\} = \frac{1}{2\pi j} \oint_{R_2} F(z) z^{k-1} dz, \tag{11.8}$$

where the integration runs along the circle R_2 around the origin in the complex plane with a radius allowing all the poles of $F(z)z^{k-1}$ to be within the circle. The inversion integral above is of theoretical importance. The proof of (11.8) is given in A.11.2 of Appendix A.5.

11.3.1 Basic Properties of the z-Transformation

Consider a DT signal $f[k] (k = 0, 1, 2, \ldots)$ with the z-transform of $F(z) = \mathcal{Z}\{f[k]\}$.

Multiplication by a constant coefficient

If c is a constant coefficient, then the z-transform of $g[k] = cf[k] (k = 0, 1, 2, \ldots)$ is

$$\mathcal{Z}\{g[k]\} = \sum_{k=0}^{\infty} z^{-k} g[k] = cf[0] + z^{-1}cf[1] + z^{-2}cf[2] + \cdots = c\sum_{k=0}^{\infty} z^{-k} g[0]$$
$$= cF(z). \tag{11.9}$$

Linearity

If c_1 and c_2 are constant coefficients, we have

$$\mathcal{Z}\{c_1 f_1[k] + c_2 f_2[k]\} = \sum_{k=0}^{\infty} z^{-k}(c_1 f_1[k] + c_2 f_2[k]) = c_1 F_1(z) + c_2 F_2(z), \tag{11.10}$$

which is considered as the linearity property for the z-transform: the z-transform of the linear combination of two signals is equal to the linear combination of the z-transforms of the signals involved.

Shift in the time-domain

Find the z-transform of $f[k-n]$, where $f[k-n]$ is derived from a DT signal $f[k]$ by a *delay* of n steps, assuming that n is a positive integer. For the delayed signal

$$\mathcal{Z}\{f[k-n]\} = z^{-n} F(z) \tag{11.11}$$

holds, as introducing $m = k - n$ and taking note of the fact that $f[m] \equiv 0$, $(m < 0)$,

$$\begin{aligned}
\mathcal{Z}\{f[k-n]\} &= \sum_{k=0}^{\infty} f[k-n] z^{-k} = z^{-n} \sum_{k=0}^{\infty} f[k-n] z^{-(k-n)} \\
&= z^{-n} \sum_{m=-n}^{\infty} f[m] z^{-m} = z^{-n} \sum_{m=0}^{\infty} f[m] z^{-m} = z^{-n} F(z)
\end{aligned}$$

follows. In a similar way, assuming n is a negative integer, a bit more involved relation can be derived for *advanced* signals:

$$\begin{aligned}
\mathcal{Z}\{f[k+n]\} &= \sum_{k=0}^{\infty} f[k+n] z^{-k} = z^{n} \sum_{k=0}^{\infty} f[k+n] z^{-(k+n)} \\
&= z^{n} \left\{ \sum_{k=0}^{\infty} f[k+n] z^{-(k+n)} + \sum_{k=0}^{n-1} f[k] z^{-k} - \sum_{k=0}^{n-1} f[k] z^{-k} \right\} \\
&= z^{n} \left\{ \sum_{k=0}^{\infty} f[k] z^{-k} - \sum_{k=0}^{n-1} f[k] z^{-k} \right\} = z^{n} \left\{ F(z) - \sum_{k=0}^{n-1} f[k] z^{-k} \right\} \\
&= z^{n} F(z) - z^{n} f[0] - z^{n-1} f[1] - \cdots - z f[n-1].
\end{aligned}$$

Multiplication by a^k
Given $F(z)$ as the z-transform of $f[k]$ the z-transform of the DT signal $a^k f[k]$ is

$$\mathcal{Z}\left\{a^k f[k]\right\} = \sum_{k=0}^{\infty} a^k f[k] z^{-k} = \sum_{k=0}^{\infty} f[k]\left(a^{-1}z\right)^{-k} = F\left(a^{-1}z\right). \tag{11.12}$$

Observe that the above derivation even allows a to be complex. Also, introducing $a = e^{-b}$, the multiplication rule can be expressed in the following form:

$$\mathcal{Z}\left\{e^{-bk} f[k]\right\} = \sum_{k=0}^{\infty} e^{-bk} f[k] z^{-k} = \sum_{k=0}^{\infty} f[k]\left(e^{b}z\right)^{-k} = F\left(e^{b}z\right). \tag{11.13}$$

11.3.2 The z-Transformation of Elementary Time Series

In the sequel, the z-transforms of a few elementary DT *signals* will be derived.
Unit impulse: $f[k] \equiv 1, \quad \mathrm{k} = 0$, otherwise $f[k] \equiv 0$

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = 1 + \sum_{k=1}^{\infty} z^{-k} 0 = 1. \tag{11.14}$$

Unit step: $f[k] \equiv 1 \ (k = 0, 1, 2, \ldots)$ and $f[k] \equiv 0, \ (k < 0)$
Apply the relation valid for the sum of a geometric series provided the *ROC* $|z| > 1$:

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = \sum_{k=0}^{\infty} z^{-k} 1 = \frac{1}{1 - z^{-1}} = \frac{z}{z-1}. \tag{11.15}$$

Unit ramp: $f[k] \equiv kT_{\mathrm{s}} \ (k = 0, 1, 2, \ldots)$
Again, provided the *ROC* is $|z| > 1$, we have

$$\begin{aligned}\mathcal{Z}\{f[k]\} &= \sum_{k=0}^{\infty} z^{-k} f[k] = \sum_{k=0}^{\infty} z^{-k} kT_{\mathrm{s}} = T_{\mathrm{s}} z^{-1}\left(1 + 2z^{-1} + 3z^{-2} + \cdots\right) \\ &= T_{\mathrm{s}} z^{-1} \frac{1}{\left(1 - z^{-1}\right)^2} = \frac{T_{\mathrm{s}} z}{(z-1)^2}\end{aligned} \tag{11.16}$$

Power function: $f[k] \equiv a^k \ (k = 0, 1, 2, \ldots)$ where a is a complex constant.
Provided the *ROC* is $|z| > |a|$, $\mathcal{Z}\{f[k]\}$ turns out to be

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = \sum_{k=0}^{\infty} z^{-k} a^k = \frac{1}{1 - az^{-1}} = \frac{z}{z-a}. \tag{11.17}$$

Applying the rule developed earlier for the multiplication by a^k, the above relation can be derived in a straightforward way:

$$\mathcal{Z}\{a^k\} = \mathcal{Z}\{a^k 1[k]\} = \left.\frac{z}{z-1}\right|_{z:=a^{-1}z} = \frac{a^{-1}z}{a^{-1}z-1} = \frac{1}{1-az^{-1}}.$$

Exponential function: $f[k] \equiv e^{-akT_s}$ $(k = 0, 1, 2, \ldots)$

Using the relation derived for the power function, assuming the *ROC* is $|z| > |e^{-aT_s}|$,

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = \sum_{k=0}^{\infty} z^{-k} e^{-akT_s} = \frac{1}{1-e^{-aT_s}z^{-1}} = \frac{z}{z-e^{-aT_s}}. \tag{11.18}$$

Sinusoidal function: $f[k] \equiv \sin(\omega k T_s)$ $(k = 0, 1, 2, \ldots)$

Applying the relation obtained earlier for the exponential function with

$$\sin(\omega k T_s) = \frac{1}{2j}\left(e^{j\omega k T_s} - e^{-j\omega k T_s}\right) \tag{11.19}$$

and assuming the *ROC* is $|z| > 1$, the following relation can be derived:

$$Z\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} \sin(\omega k T_s) = \frac{z\sin(\omega T_s)}{z^2 - 2z\cos(\omega T_s) + 1}. \tag{11.20}$$

The results derived so far, together with some extensions are summarized in Table 11.1.

Table 11.1 Laplace and z-transforms of some functions

$F(s)$	$f(t)$	$f[k] = f(kT_s)$	$F(z)$
1	$\delta(t)$	$\delta[k]$	1
$\frac{1}{s}$	$1(t)$	$1[k]$	$\frac{z}{z-1}$
$\frac{1}{s+a}$	e^{-at}	e^{-akT_s}	$\frac{z}{z-e^{-aT_s}}$
$\frac{1}{s^2}$	t	kT_s	$\frac{T_s z}{(z-1)^2}$
$\frac{1}{(s+a)^2}$	te^{-at}	$kT_s e^{-akT_s}$	$\frac{T_s e^{-aT_s} z}{(z-e^{-aT_s})^2}$
$\frac{1}{s^3}$	t^2	$(kT_s)^2$	$\frac{T_s^2 z(z+1)}{(z-1)^3}$
$\frac{b-a}{(s+a)(s+b)}$	$e^{-at} - e^{-bt}$	$e^{-akT_s} - e^{-bkT_s}$	$\frac{\left(e^{-aT_s} - e^{-bT_s}\right)z}{(z-e^{-aT_s})(z-e^{-bT_s})}$
$\frac{\omega}{s^2+\omega^2}$	$\sin(\omega t)$	$\sin(\omega k T_s)$	$\frac{z\sin(\omega T_s)}{z^2 - 2z\cos(\omega T_s) + 1}$
$\frac{s}{s^2+\omega^2}$	$\cos(\omega t)$	$\cos(\omega k T_s)$	$\frac{z^2 - z\cos(\omega T_s)}{z^2 - 2z\cos(\omega T_s) + 1}$

As far as the application of the this table is concerned an important comment should be added here. Table 11.1 contains four columns. However, an unambiguous mapping exists only for the pairs of $f(t) \leftrightarrow F(s)$ and $f[k] \leftrightarrow F(z)$. More precisely, given a CT signal $f(t)$, the table always indicates a z-transform $F(z)$, however, given a z-transform by $F(z)$, only a CT signal can be obtained, whose samples are only defined at the sampling instants. In such a way a number of CT signals can be derived with diverse intersampling behavior.

11.3.3 The Inverse z-Transformation

Assume that $F(z)$ is the z-transform of the DT signal $f[k]$. Here we ask the question of how to find the DT signal $f[k]$ once $F(z)$ is given. This task is called the inverse z-transformation. In theory, the answer to our question has already been formulated by

$$f[k] = \mathcal{Z}^{-1}\{F(z)\} = \frac{1}{2\pi j} \oint_{R_2} F(z) z^{k-1} dz. \tag{11.21}$$

In practice one of the following three options are used:

Polynomial division

Assume $F(z) = 10z/[(z-1)(z-0.2)]$ then divide the polynomial in the numerator by the polynomial in the denominator and read the samples of $f[k]$ $(k = 0, 1, 2, \ldots)$ as the coefficients obtained along the division:

$$(10z):\left(z^2 - 1.2z + 0.2\right) = f[0] + z^{-1} f[1] + z^{-2} f[2] + \cdots$$

Thus for the first few samples $f[0] = 0, f[1] = 10, f[2] = 12, f[3] = 12.4$ are obtained. The method is not quite efficient. In addition there are *CAD* tools leading to numerical results in a far more efficient way.

Calculating the DT *impulse response*

Consider $F(z)$ as an impulse response function (see later on the pulse transfer function) between a DT input signal with a z-transform of $U(z)$ and a response with a z-transform of $Y(z)$:

$$Y(z) = F(z)U(z). \tag{11.22}$$

Applying a unit impulse input with $U(z) = 1$, then for the z-transform of the response $Y(z) = F(z)$. The calculated samples obviously match the values obtained by polynomial division.

Partial Fractional Expansion (PFE)

The key point here is to decompose $F(z)$ to a sum of components existing in tables of z-transform pairs. Observe the structure of $F(z)$ in the rightmost column of Table 11.1. To ensure the appearance of z in the numerator set up the *PFE* form as follows:

$$\frac{F(z)}{z} = \frac{c_o}{z} + \frac{c_1}{z - p_1} + \frac{c_2}{z - p_2} + \cdots + \frac{c_n}{z - p_n}. \tag{11.23}$$

In this decomposition simple real poles have been assumed. As an example, consider

$$\frac{F(z)}{z} = \frac{12.5}{z-1} - \frac{12.5}{z-0.2} \quad \text{or} \quad F(z) = \frac{12.5z}{z-1} - \frac{12.5z}{z-0.2},$$

which lead to

$$f[k] = \mathcal{Z}^{-1}\{F(z)\} = \mathcal{Z}^{-1}\left\{\frac{12.5z}{z-1} - \frac{12.5z}{z-0.2}\right\} = 12.5\left(1 - 0.2^k\right),\ k = 0, 1, 2, \ldots$$

Observe that unlike the first two methods, the *PFE* delivers $f[k]$ in analytical form, so its value can be easily computed for arbitrary $k \geq 0$.

Handling complex poles

Setting up the *PFE* form becomes more involved in the case of complex poles of $F(z)$ even for a simple complex pole pair.

Example 11.1 Consider

$$F(z) = \frac{z^3 + 1}{z^3 - z^2 - z - 2}.$$

Following the *PFE* procedure,

$$\frac{F(z)}{z} = -\frac{0.5}{z} + \frac{0.643}{z-2} + \frac{c}{z-p} - \frac{\bar{c}}{z-\bar{p}},$$

is obtained, where $c = 0.429 + j0.0825$ and $p = -0.5 - j0.866$, and $\bar{c}$ and $\bar{p}$ denote their complex conjugate values, respectively. Evaluating the last two terms as one second-order component, two sinusoidal components can be reconstructed based on the last two rows in Table 11.1. Combining these two sinusoidal components into one single sinusoidal form yields

$$\begin{aligned} f[k] &= -0.5\delta[k] + 0.643(2)^k + c(p)^k + \bar{c}(\bar{p})^k \\ &= -0.5\delta[k] + 0.643(2)^k + \cos\left(\frac{4\pi}{3} + 10.89^\circ\right) \end{aligned}$$

by using the trigonometric identity

$$c(p)^k + \bar{c}(\bar{p})^k = 2C\sigma^k \cos(\Omega k + \Theta),$$

where $c = Ce^{j\Theta}$ and $p = \sigma e^{j\Omega}$. ■

Handling multiple poles

Example 11.2 The discussion concerning repeated poles will be reduced to a numerical example. Consider

$$F(z) = \frac{6z^3 + 2z^2 - z}{z^3 - z^2 - z + 1}.$$

The structure of the *PFE* is governed by the poles of $F(z)$. In this case $p_{1,2} = 1$ and $p_3 = -1$, i.e., $p_{1,2} = 1$ is a double pole. Accordingly,

$$\frac{F(z)}{z} = \frac{5.25}{z-1} + \frac{3.5}{(z-1)^2} + \frac{0.75}{z+1}.$$

Observe that in the *PFE* form the repeated pole shows up in single *and* double forms, as well. The inverse z-transformation then leads to

$$f[k] = 5.25(1)^k + 3.5k + 0.75(-1)^k, \quad k = 0, 1, 2, \ldots.$$ ■

General method for single poles

If the Laplace-transform of a *CT* signal has the form $F(s) = \mathcal{F}_{\mathrm{z}}(s)/\mathcal{F}_{\mathrm{p}}(s)$, where $\mathcal{F}_{\mathrm{z}}(s)$ and $\mathcal{F}_{\mathrm{p}}(s)$ are the numerator and denominator of $F(s)$, respectively, then the z-transform of the sampled signal is given by

$$F(z) = \sum_{i=1}^{n} \frac{\mathcal{F}_{\mathrm{z}}(p_i)}{\mathcal{F}'_{\mathrm{p}}(p_i)} \frac{z}{z - e^{p_i T_s}} = \sum_{i=1}^{n} \frac{\mathcal{F}_{\mathrm{z}}(p_i)}{\mathcal{F}'_{\mathrm{p}}(p_i)} \frac{1}{1 - e^{p_i T_s} z^{-1}}. \tag{11.24}$$

For single poles the simple 'cover up' technique to determine the residues can be used

$$\frac{1}{\mathcal{F}'(p_i)} = \lim_{s \to p_i} (s - p_i) \frac{1}{\mathcal{F}_{\mathrm{p}}(s)} \tag{11.25}$$

to ease the evaluation of $F(z)$. The application of this method assumes that the denominator is available in factored form. One component of the sum can be calculated by covering up $s - p_i$ and substituting $s = p_i$ in the rest.

11.3.4 Initial and Final Value Theorems

Once $F(z)$ is given, the initial and final value theorems provide tools to determine $f[0]$ and $\lim\limits_{k\to\infty} f[k]$ in a direct way, without performing the inverse z-transformation.

Initial value theorem $(k \to 0)$

Starting from the definition of the z-transform

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = f[0] + z^{-1} f[1] + z^{-2} f[2] + \cdots$$

it can be seen that $f[0]$ can be easily determined as $z \to \infty$:

$$f[0] = \lim_{z\to\infty} F(z). \tag{11.26}$$

Final value theorem $(k \to \infty)$

The basic idea is to separate the 'last' element of the sequence $f[0]$ by subtracting the original and the delayed sequences:

$$\begin{aligned} \lim_{k\to\infty} f[k] &= \lim_{z\to 1}\{F(z) - z^{-1}F(z)\} \\ &= (f[0] + f[1] + f[2] + \cdots) - (f[-1] + f[0] + f[1] + \cdots). \end{aligned}$$

Assuming positive-time-functions $(f[k] = 0,\ k < 0)$ results in

$$\lim_{k\to\infty} f[k] = \lim_{z\to 1}\{\left(1 - z^{-1}\right)F(z)\}. \tag{11.27}$$

The final value theorem can only be applied if the signal has a steady state value.

11.4 Description of Sampled Data Systems in the Discrete-Time and in the Operator and Frequency Domain

Discussing CT systems, the need for an abstract system description for closed-loop analysis and design was seen. Considering sampled data systems, the value of the control input and that of the samples of the output only change at the sampling instants. It seems to be reasonable then to create a mathematical model to describe the system behavior only at the sampling instants. The process of describing the behavior of a sampled CT system will be referred to as discretization. As a starting point, the *state-space* description will be discussed, then the methods of the *impulse response function* (pulse transfer function) and the *difference equation* will be derived. This triple is in complete harmony with the triple used for CT systems (state-space, transfer function, differential equation).

11.4.1 The State-Space Model

Assume the *LTI* state-space model of the CT system to be discretized along with a sampling time of T_s (see 3.10):

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t) \\ y(t) &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t) + du(t) \end{aligned} \tag{11.28}$$

The solution of the state equation as discussed earlier for CT systems (3.2.1) with the initial time t_o and the initial state vector $\boldsymbol{x}(t_o)$ is

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}(t-t_o)}\boldsymbol{x}(t_o) + \int_{t_o}^{t} e^{\boldsymbol{A}(t-\tau)}\boldsymbol{b}u(\tau)\mathrm{d}\tau = e^{\boldsymbol{A}(t-t_o)}\boldsymbol{x}(t_o) + \left[\int_{t_o}^{t} e^{\boldsymbol{A}(t-\tau)}u(\tau)\mathrm{d}\tau\right]\boldsymbol{b}. \tag{11.29}$$

Assuming that the discretized model will contain a ZOH unit (in other words the continuous control input will be a staircase function) perform the integration from kT_s to $(k+1)T_s$:

$$\begin{aligned} \boldsymbol{x}(kT_s + T_s) &= e^{\boldsymbol{A}T_s}\boldsymbol{x}(kT_s) + \int_{kT_s}^{kT_s+T_s} e^{\boldsymbol{A}(kT_s+T_s-\tau)}\boldsymbol{b}u(\tau)\mathrm{d}\tau \\ &= e^{\boldsymbol{A}T_s}\boldsymbol{x}(kT_s) + \boldsymbol{b}u(kT_s)\int_{kT_s}^{kT_s+T_s} e^{\boldsymbol{A}(kT_s+T_s-\tau)}\mathrm{d}\tau \\ &= e^{\boldsymbol{A}T_s}\boldsymbol{x}(kT_s) + \left[u(kT_s)\int_0^{T_s} e^{\boldsymbol{A}\lambda}\mathrm{d}\lambda\right]\boldsymbol{b} \end{aligned}$$

The above manipulations used the fact that $u(\tau) = \text{constant}$ within each sampling period $([kT_s \le \tau < (k+1)T_s])$, that is, $u(\tau) = u(kT_s)$. Furthermore, to simplify the evaluation of the integral, $\lambda = kT_s + T_s - \tau$ has been introduced. Using the standard notations $\boldsymbol{x}(kT_s + T_s) = \boldsymbol{x}[k+1]$, $\boldsymbol{x}(kT_s) = \boldsymbol{x}[k]$ and $u(kT_s) = u[k]$, the above equation can be rewritten as

$$\boldsymbol{x}[k+1] = e^{\boldsymbol{A}T_s}\boldsymbol{x}[k] + \left[\int_0^{T_s} e^{\boldsymbol{A}\lambda}\mathrm{d}\lambda\right]\boldsymbol{b}u[k]. \tag{11.30}$$

Introducing the parameter matrices for the discretized model

$$\boldsymbol{F} = e^{\boldsymbol{A}T_s} \quad \text{and} \quad \boldsymbol{g} = \int_0^{T_s} e^{\boldsymbol{A}\lambda}\mathrm{d}\lambda\,\boldsymbol{b} \tag{11.31}$$

the following discretized state-model is obtained: $\boldsymbol{x}[k+1] = \boldsymbol{F}\boldsymbol{x}[k] + \boldsymbol{g}u[k]$. Note that if $\boldsymbol{A}$ is invertible, then the integration in the expression for $\boldsymbol{g}$ can be carried out and this leads to the closed-form formula $\boldsymbol{g} = \boldsymbol{A}^{-1}\left(e^{\boldsymbol{A}T_s} - \boldsymbol{I}\right)\boldsymbol{b}$. The output equation of the CT state-model of $y(t) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(t) + du(t)$ can simply be sampled by substituting $t = kT_s$:

$$y(kT_s) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}(kT_s) + du(kT_s)$$

or, to emphasize the DT nature of the model,

$$y[k] = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}[k] + du[k]. \tag{11.32}$$

The DT state-model can be summarized as follows. The DT *state difference equation*

$$\boldsymbol{x}[k+1] = \boldsymbol{F}\boldsymbol{x}[k] + \boldsymbol{g}u[k] \tag{11.33}$$

and the DT *output equation* according to (11.32) form the DT state-model for $k = 0, 1, 2, \ldots$. Comparing the DT state-model with the CT state-model, (11.31) shows how to derive the parameter matrices of the DT state difference equation from the parameter matrices of the CT state-model, while the parameters of the output equation ($\boldsymbol{c}^{\mathrm{T}}$ and d) are identical for the DT and CT state-models. The matrix $\boldsymbol{F}$ will be called the state transition matrix, in particular assuming zero excitation, it governs the transition between $\boldsymbol{x}[k]$ and $\boldsymbol{x}[k+1]$ according to

$$\boldsymbol{x}[k+1] = \boldsymbol{F}\boldsymbol{x}[k]. \tag{11.34}$$

Example 11.3 To study the nature of the operations while discretizing a system and the influence of the sampling time on these operations, consider a second order example, namely a double integrator. Select the state variables as shown in Fig. 11.15.

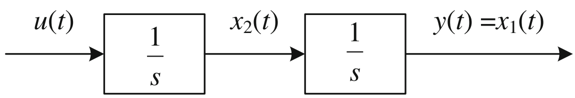


Fig. 11.15 CT double integrator

The CT state equations:

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = u(t)$$

The CT state-model:

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t),$$

hence

$$\boldsymbol{F} = e^{\boldsymbol{A}T_\mathrm{s}} = e^{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} T_\mathrm{s}}.$$

To determine the above matrix exponential evaluate the infinite series (see 3.20 and 3.26)

$$e^{\boldsymbol{A}T_\mathrm{s}} = \boldsymbol{I} + \boldsymbol{A}T_\mathrm{s} + \frac{1}{2}\boldsymbol{A}^2 T_\mathrm{s}^2 + \cdots$$

note that $\boldsymbol{A}^k = 0 \, (k \geq 2)$, thus

$$e^{\boldsymbol{A}T_\mathrm{s}} = \boldsymbol{I} + \boldsymbol{A}T_\mathrm{s} + \frac{1}{2}\boldsymbol{A}^2 T_\mathrm{s}^2 + \cdots = \boldsymbol{I} + \boldsymbol{A}T_\mathrm{s} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & T_\mathrm{s} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & T_\mathrm{s} \\ 0 & 1 \end{bmatrix}$$

and

$$\boldsymbol{g} = \int_0^{T_\mathrm{s}} e^{\boldsymbol{A}\lambda} \mathrm{d}\lambda \, \boldsymbol{b} = \int_0^{T_\mathrm{s}} \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} \mathrm{d}\lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \int_0^{T_\mathrm{s}} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \mathrm{d}\lambda = \begin{bmatrix} T_\mathrm{s}^2/2 \\ T_\mathrm{s} \end{bmatrix}.$$

In the case of higher order systems the application of *CAD* tools is advised. ■

Given the initial state $\boldsymbol{x}[0]$, the solution of the state difference equation is well known from the theory of "Signals and Systems":

$$\boldsymbol{x}[k] = \boldsymbol{F}^k \boldsymbol{x}[0] + \sum_{m=0}^{k-1} \boldsymbol{F}^{k-m-1} \boldsymbol{g} u[m], \tag{11.35}$$

where the first term depends on the initial value of the state vector, while the second is a weighted sum of the input samples at $0, 1, \ldots, (k-1)$. It can be seen that in the above solution the $\boldsymbol{F}$ state transition matrix plays a key role. It will be shown later on that $\boldsymbol{F}$ has a fundamental role in determining other important system properties, like stability, as well.

11.4.2 Input-Output Models Based on the Shift Operator

The input-output models will be presented according to the operators used to describe the relation between the input and output samples. First an expressive modeling approach based on the application of the shift operator will be shown. This approach directly supports the discussion using the concept of difference equations. Just as the LAPLACE-transform played a fundamental role in the description of CT systems, the application of z-transforms will be useful for DT systems.

Discussing CT systems, it has been shown that an input-output model given by a differential equation or by a transfer function can be transformed into an infinite number of input-output equivalent state-models. On the contrary, input-output equivalent state-models always exhibit one single input-output model. This property is valid for DT systems, as well. To present this property, first introduce the *shift operator* according to

$$q\boldsymbol{x}[k] = \boldsymbol{x}[k+1] \quad (k = \ldots, -1, 0, 1, \ldots). \tag{11.36}$$

The function of the q operator is to advance a scalar- or vector-valued DT sequence $\boldsymbol{x}[k]$ by one single step. The repeated application of the shift operator leads to advancing the sequence by several steps. E.g., for m steps,

$$q^m\boldsymbol{x}[k] = \boldsymbol{x}[k+m] \quad (k{=}\ldots, -1, 0, 1,\ldots) \tag{11.37}$$

is to be applied. A delay can be realized by using the inverse of the operator q:

$$q^{-1}\boldsymbol{x}[k] = \boldsymbol{x}[k-1] \quad (k = \ldots, -1, 0, 1, \ldots) \tag{11.38}$$

or

$$q^{-m}\boldsymbol{x}[k] = \boldsymbol{x}[k-m] \quad (k{=}\ldots, -1, 0, 1,\ldots).$$

In the sequel, the operator q^{-1} will be referred to as the delay or shift operator. Apply q to the DT state-model:

$$\begin{aligned} \boldsymbol{x}[k+1] &= q\boldsymbol{x}[k] = \boldsymbol{F}\boldsymbol{x}[k] + \boldsymbol{g}u[k] \\ y[k] &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}[k] + du[k] \end{aligned} \tag{11.39}$$

Solving for $\boldsymbol{x}[k]$

$$\boldsymbol{x}[k] = (q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g}u[k] \tag{11.40}$$

allows rewriting the output equation as

$$y[k] = \boldsymbol{c}^{\mathrm{T}}(q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g}u[k] + du[k] = \left[\boldsymbol{c}^{\mathrm{T}}(q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g} + d\right]u[k]. \tag{11.41}$$

The dependence of the output sequence on the input sequence can be expressed by

$$y[k] = G(q)u[k] = \left[\boldsymbol{c}^{\mathrm{T}}(q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g} + d\right]u[k], \tag{11.42}$$

where $G(q)$, the *transfer function operator*, has been introduced. Analyzing $G(q)$, it can be seen that it is a rational function

$$G(q) = \boldsymbol{c}^{\mathrm{T}}(q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g} + d = \boldsymbol{c}^{\mathrm{T}}\frac{\mathbf{adj}(q\boldsymbol{I} - \boldsymbol{F})}{\det(q\boldsymbol{I} - \boldsymbol{F})}\boldsymbol{g} + d = \frac{\mathcal{B}(q)}{\mathcal{A}(q)}, \tag{11.43}$$

where $\det(q\boldsymbol{I} - \boldsymbol{F})$ and $\boldsymbol{c}^{\mathrm{T}}\mathbf{adj}(q\boldsymbol{I} - \boldsymbol{F})\boldsymbol{g}$ are polynomials in the operator q with real coefficients:

$$\begin{aligned} \mathcal{A}(q) &= \det(q\boldsymbol{I} - \boldsymbol{F}) = q^n + a_1 q^{n-1} + \cdots + a_n \\ \mathcal{B}(q) &= b_{\mathrm{o}} q^{n_{\mathrm{B}}} + b_1 q^{n_{\mathrm{B}}-1} + \cdots + b_{n_{\mathrm{B}}} \end{aligned} \tag{11.44}$$

As far as the degrees of the introduced polynomials $\mathcal{A}(q)$ and $\mathcal{B}(q)$ are concerned, n is the number of the state variables, while n_B is subject to the constraint $n_{\mathrm{B}} \leq n$. For $d = 0$, which holds for most systems, (11.43) is a strictly proper rational function, so it is reasonable to rewrite (11.44) for $n_{\mathrm{B}} = n - 1$ as

$$\begin{aligned} \mathcal{A}(q) &= \det(q\boldsymbol{I} - \boldsymbol{F}) = q^n + a_1 q^{n-1} + \cdots + a_n \\ &= q^n\left(1 + a_1 q^{-1} + \cdots + a_n q^{-n}\right) = q^n \mathcal{A}\left(q^{-1}\right) \\ \mathcal{B}(q) &= b_1 q^{n-1} + b_2 q^{n-2} + \cdots + b_n \\ &= q^n\left(b_1 q^{-1} + b_2 q^{-2} + \cdots + b_n q^{-n}\right) = q^n \mathcal{B}\left(q^{-1}\right) \end{aligned} \tag{11.45}$$

Similarly to the terminology introduced for CT systems, the roots of $\mathcal{A}(q) = 0$ will be called the DT poles and the roots of $\mathcal{B}(q) = 0$ will be called the DT zeros.

Based on the introduced polynomials and utilizing the time domain interpretation of the shift operator, another model of DT systems, namely the difference equation model, can be derived. As discussed earlier,

$$y[k] = G(q)u[k] = \frac{\mathcal{B}(q)}{\mathcal{A}(q)}u[k], \tag{11.46}$$

which can be written as

$$\mathcal{A}(q)y[k] = \mathcal{B}(q)u[k]. \tag{11.47}$$

Further substitution gives

$$\left(q^n + a_1 q^{n-1} + \cdots + a_n\right)y[k] = \left(b_1 q^{n-1} + b_2 q^{n-2} + \cdots + b_n\right)u[k] \tag{11.48}$$

and finally

$$y[k+n] + a_1 y[k+n-1] + \cdots + a_n y[k] = b_1 u[k+n-1] + b_1 u[k+n-2] + \cdots + b_n u[k].$$

Divide both sides of (11.48) by q^n:

$$\left(1 + a_1 q^{-1} + \cdots + a_n q^{-n}\right)y[k] = \left(b_1 q^{-1} + b_2 q^{-2} + \cdots + b_n q^{-n}\right)u[k] \tag{11.49}$$

and rearrange it

$$y[k] = b_1 u[k-1] + b_2 u[k-2] + \cdots + b_n u[k-n] - a_1 y[k-1] - \cdots - a_n y[k-n] \tag{11.50}$$

which is the input-output difference equation. Another interpretation suggests considering (11.50) as a *recursive formula* to calculate the output sample. It can be seen that to derive the recursive formula, the polynomials of (11.50) have been used in their form arranged by q^{-1}. Thus the pulse transfer function operator can equivalently be defined as follows:

$$G\left(q^{-1}\right) = \frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})} = \frac{\mathcal{B}(q^{-1})}{1 + \widetilde{\mathcal{A}}(q^{-1})}, \tag{11.51}$$

which is also called the filter form, for based on (11.51) the recursive formula by (11.50) can always be derived:

$$\begin{aligned} y[k] &= \boldsymbol{B}\left(q^{-1}\right)u[k] - \tilde{\boldsymbol{A}}\left(q^{-1}\right)y[k] \\ &= b_1 u[k-1] + b_1 u[k-2] + \cdots + b_n u[k-n] - a_1 y[k-1] - \cdots - a_n y[k-n] \end{aligned} \tag{11.52}$$

which is linear-in-the-parameters $\{a_i; b_i\}$ and ready to be implemented in a computing environment.

Example 11.4 Find the pulse transfer function operator $G(q)$ for the double integrator discussed earlier in Example 11.3. Apply $T_s = 1$:

$$G(q) = \boldsymbol{c}^{\mathrm{T}}(q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g} + d = [1 \quad 0]\begin{bmatrix} q-1 & -1 \\ 0 & q-1 \end{bmatrix}^{-1}\begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \frac{0.5(q+1)}{(q-1)^2}$$
$$= \frac{0.5q + 0.5}{q^2 - 2q + 1}.$$

Both poles of the system are equal to $p_{\mathrm{d}_{1,2}} = 1$, while the single zero is $z_{\mathrm{d}} = -1$. ■

11.4.3 Modeling Based on the z-Transformation

The z-transform of a DT sequence $f[k], (k = 0, 1, 2, \ldots)$ is defined as follows:

$$\mathcal{Z}\{f[k]\} = \sum_{k=0}^{\infty} z^{-k} f[k] = f[0] + z^{-1}f[1] + z^{-2}f[2] + \cdots, \tag{11.53}$$

It can be shown that the complex variable of the z-transform (z) is in a close relationship with the complex variable of the LAPLACE-transform (s). As a guiding principle, a discretization is looked for where the values attached to the sequence of impulses are identical to those obtained from sampling. The theory of hybrid systems calls this principle *impulse invariance*. Consider the mathematically sampled form of a CT signal $f(t)$:

$$f^*(t) = \sum_{m=0}^{\infty} f(mT_{\mathrm{s}})\delta(t - mT_{\mathrm{s}}), \tag{11.54}$$

then find the LAPLACE-transform of $f^*(t)$:

$$\mathcal{L}\{f^*(t)\} = \mathcal{L}\left\{\sum_{m=0}^{\infty} f(mT_{\mathrm{s}})\delta(t - mT_{\mathrm{s}})\right\} = f(0) + f(T_s)e^{-sT_{\mathrm{s}}} + f(2T_{\mathrm{s}})e^{-2sT_{\mathrm{s}}} + \cdots$$

Note that $f^*(t)$ is a DT signal. To comply with the impulse invariance principle

$$\mathcal{L}\{f^*(t)\} = \mathcal{Z}\{f[k]\} \tag{11.55}$$

is obtained, which leads to

$$f(0) + f(T_{\mathrm{s}})e^{-sT_{\mathrm{s}}} + f(2T_{\mathrm{s}})e^{-2sT_{\mathrm{s}}} + \cdots = f[0] + z^{-1}f[1] + z^{-2}f[2] + \cdots,$$

implying that

$$z = e^{sT_s}. \tag{11.56}$$

To emphasize the importance of the relation $z = e^{sT_s}$ find the stability region for DT systems. For CT systems, the left half-plane of the $s = \sigma + j\omega$ complex plane turned out to be the stability region concerning the poles of the CT system. The $z = e^{sT_s}$ maps this left half-plane into the unit disc in the complex z-plane. To see this, observe that $z = e^{sT_s}$ maps the borderline of the stability of the s-plane $s = j\omega\,(-\infty < \omega < \infty)$ to the unit circle. As the frequency increases the region of stability is to the left of the borderline in the s-plane, so will happen in the z plane. More exactly, sampling will map the fundamental band $-\pi/T_s < \omega < \pi/T_s$ in the s half-plane into the unit disc (bands outside the fundamental band repeat themselves). The two lines parallel to the negative real axis in the s-plane at $\omega = \pm j\pi/T_s$ will be transformed to one single line (the negative real axis of the z-plane). Poles according to the limit of the SHANNON sampling law will be transformed to the negative real axis of the z-plane.

The simple basic building blocks for DT models can be determined by approximating the CT operators while acting on the sequence of sampled data. Start with simple differentiation.

Backward difference:

Based on

$$\frac{\mathrm{d}y}{\mathrm{d}t} \approx \frac{y[k] - y[k-1]}{T_s} = \frac{1 - z^{-1}}{T_s} y[k] \tag{11.57}$$

the operator $H_D(s) = s$ for differentiation has a DT equivalent by,

$$G_D(z) = \frac{z-1}{T_s z} = \frac{1 - z^{-1}}{T_s}. \tag{11.58}$$

Forward difference:

$$\frac{\mathrm{d}y}{\mathrm{d}t} \approx \frac{y[k+1] - y[k]}{T_s} = \frac{z-1}{T_s} y[k] = \frac{1 - z^{-1}}{T_s z^{-1}} y[k], \tag{11.59}$$

suggests using

$$G_D(z) = \frac{z-1}{T_s} = \frac{1 - z^{-1}}{T_s z^{-1}}. \tag{11.60}$$

as the DT equivalent of the CT differentiation.

Consider now the simple integration by $H_{\mathrm{I}}(s) = 1/s$.

Right hand rectangular rule:

$$y(t) = \int_0^t u(\tau)\mathrm{d}\tau \approx \sum_{i=0}^{k} u(i)T_{\mathrm{s}} = T_{\mathrm{s}} \sum_{i=0}^{k} u(i) = y[k] = y[k-1] + T_{\mathrm{s}}u[k], \quad (11.61)$$

allows setting up the following recursive formula:

$$y[k] - y[k-1] = \left(1 - z^{-1}\right)y[k] = T_{\mathrm{s}}u[k]. \quad (11.62)$$

Consequently,

$$G_{\mathrm{I}}(z) = \frac{T_{\mathrm{s}}}{1 - z^{-1}} = \frac{T_{\mathrm{s}}z}{z - 1}. \quad (11.63)$$

Left hand rectangular rule:

$$y(t) = \int_0^t u(\tau)\mathrm{d}\tau \approx \sum_{i=0}^{k} u(i)T_{\mathrm{s}} = T_{\mathrm{s}} \sum_{i=0}^{k} u(i) = y[k] = y[k-1] + T_{\mathrm{s}}u[k-1], \quad (11.64)$$

leads to

$$y[k+1] - y[k] = (z-1)y[k] = T_{\mathrm{s}}u[k]. \quad (11.65)$$

Consequently,

$$G_{\mathrm{I}}(z) = \frac{T_{\mathrm{s}}}{z - 1} = \frac{T_{\mathrm{s}}z^{-1}}{1 - z^{-1}}. \quad (11.66)$$

To recapture the methodology used for discussing CT systems, Laplace-transforms were effectively used by introducing various transfer functions. For DT systems the z-transforms may play a similarly important role once the notion of *pulse transfer function* is introduced as the ratio of the z-transforms of two signals. A natural requirement here is that both the input and the output signal should be DT sequences. This means that the ZOH unit should also be taken into account. Figure 11.16 shows the proper arrangement driven by the DT input $u[k]$ and generating the DT output signal $y[k]$.

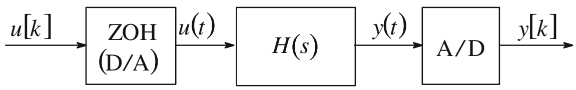


Fig. 11.16 A CT block extended by a ZOH and a sampler

The *pulse transfer function* of the discretized system (assuming zero initial conditions) is

$$G(z) = \frac{\mathcal{Z}\{y[k]\}}{\mathcal{Z}\{u[k]\}}. \tag{11.67}$$

For an arbitrary input the sampled response can be calculated by

$$y[k] = \mathcal{Z}\Big\{ y(t)|_{t=kT_\mathrm{s}} \Big\} = \mathcal{Z}\Big\{ \mathcal{L}^{-1}[Y(s)]_{t=kT_\mathrm{s}} \Big\} = \mathcal{Z}\Big\{ \mathcal{L}^{-1}[U(s)H(s)]_{t=kT_\mathrm{s}} \Big\}. \tag{11.68}$$

It can be seen that the above expression is input-dependent, in other words instead of a general DT model, only a DT model valid for a given class of inputs can be formulated. Aim at setting up a *Step Response Equivalent* (*SRE*) DT model valid for a unit step sequence of $u[k]$ resulting in a CT step response $u(t)$ input:

$$y[k] = \mathcal{Z}\left\{ \mathcal{L}^{-1}\left[\frac{H(s)}{s}\right]_{t=kT_\mathrm{s}} \right\}. \tag{11.69}$$

In this case the CT and DT models will exhibit identical outputs at the sampling instants. Observe that the LAPLACE-transform yields the step response of the CT system:

$$\mathcal{L}^{-1}\left\{\frac{H(s)}{s}\right\} = v(t), \tag{11.70}$$

whose samples determine the DT step response

$$v[k] = \mathcal{L}^{-1}\left[\frac{H(s)}{s}\right]_{t=kT_\mathrm{s}}, \tag{11.71}$$

thus the DT model can be written as

$$G(z) = \frac{\mathcal{Z}\{v[k]\}}{\mathcal{Z}\{u[k]\}} = \frac{\mathcal{Z}\{v[k]\}}{\frac{z}{z-1}} = \frac{z-1}{z}\mathcal{Z}\{v[k]\} = \left(1 - z^{-1}\right)\mathcal{Z}\{v[k]\}. \tag{11.72}$$

Note that the above relation can be found in some textbooks as

$$G(z) = \left(1 - z^{-1}\right)\mathcal{Z}\left\{\frac{H(s)}{s}\right\}. \tag{11.73}$$

The above form is rather expressive, however, it is not correct in the mathematical sense, since a LAPLACE-transform obviously has no z-transform. The correct procedure is to perform an inverse LAPLACE-transformation, then sampling the CT signal, a DT signal is obtained by substituting $t = kT_s$. In principle, the *SRE* transformation matches the case of using a ZOH unit, i.e., between the sampling instants there are always step-like excitations.

In a similar way, *Ramp Response Equivalent* (*RRE*) DT models can also be derived:

$$y[k] = \mathcal{Z}\left\{\mathcal{L}^{-1}\left[\frac{H(s)}{s^2}\right]_{t=kT_s}\right\}. \tag{11.74}$$

In principle, the *RRE* transformation matches the case of using a first order holding unit in the closed-loop. In this arrangement, there are always ramp-like excitations between two samples.

Example 11.5 Apply the results obtained so far for setting up a DT model discretizing a double integrator. The unit step response of the double integrator is

$$v(t) = \frac{t^2}{2} \quad (\mathrm{t} \geq 0), \tag{11.75}$$

or in a sampled form

$$v[k] = \frac{(kT_s)^2}{2} \quad (\mathrm{k} \geq 0), \tag{11.76}$$

z-transformation gives

$$\mathcal{Z}\{v[k]\} = \frac{T_s^2}{2} z(z+1)/(z-1)^3, \tag{11.77}$$

and finally

$$G(z) = \left(1 - z^{-1}\right)\mathcal{Z}\{v[k]\} = \frac{z-1}{z}\mathcal{Z}\{v[k]\} = \frac{z-1}{z}\frac{T_s^2}{2}\frac{z(z+1)}{(z-1)^3} = \frac{T_s^2}{2}\frac{(z+1)}{(z-1)^2}. \tag{11.78}$$

Specifically, $G(z) = (z+1)/2(z-1)^2 = (0.5z+0.5)/(z^2 - 2z+1)$ is obtained for $T_s = 1$. ■

Observe the formal coincidence of the forms derived by the pulse transfer operator $G(q)$ and by the z-transforms. For the relations of the z-transformation the interpretation of multiplication by z is advancing a sequence by one sample and the interpretation of multiplication by z^{-1} is delaying a sequence by one sample, in general it can be stated that $G(q)$ can be obtained by substituting $z = q$ in $G(z)$. Despite the formal matching keep it in mind that the shift operator and the variable of the z-transforms are different notions!

Repeating the derivation used for the shift operator it can be shown that $G(z) = \dfrac{\mathcal{B}(z)}{\mathcal{A}(z)} = \boldsymbol{c}^{\mathrm{T}} \dfrac{\mathbf{adj}(z\boldsymbol{I} - \boldsymbol{F})}{\det(z\boldsymbol{I} - \boldsymbol{F})} \boldsymbol{g} + d$. From this the characteristic equation is found to be $\det(z\boldsymbol{I} - \boldsymbol{F}) = 0$. For the stability of the DT model then $|z_i| < 1 \ (i = 1, 2, .., n)$ applies.

Example 11.6 Derive an analytical relation to determine the *SRE* DT model of the first order lag given by $H(s) = K/(1 + sT)$. Use the simplified expression of $G(z) = (1 - z^{-1})\mathcal{Z}\{H(s)/s\}$

$$\begin{aligned}
G(z) &= \left(1 - z^{-1}\right)\mathcal{Z}\left\{\frac{H(s)}{s}\right\} = \left(1 - z^{-1}\right)\mathcal{Z}\left\{\frac{K}{s(1 + sT)}\right\} = \left(1 - z^{-1}\right)\mathcal{Z}\left\{\frac{K}{s} - \frac{KT}{1 + sT}\right\} \\
&= \left(1 - z^{-1}\right)\frac{Kz}{z - 1} - \left(1 - z^{-1}\right)\frac{Kz}{z - e^{-T_s/T}} = K\left(1 - \frac{z - 1}{z - e^{-T_s/T}}\right) = K\frac{1 - e^{-T_s/T}}{z - e^{-T_s/T}} \\
&= \frac{K\left(1 - e^{-T_s/T}\right)z^{-1}}{1 - e^{-T_s/T}z^{-1}} = \frac{b_1 z^{-1}}{1 + a_1 z^{-1}}
\end{aligned} \tag{11.79}$$

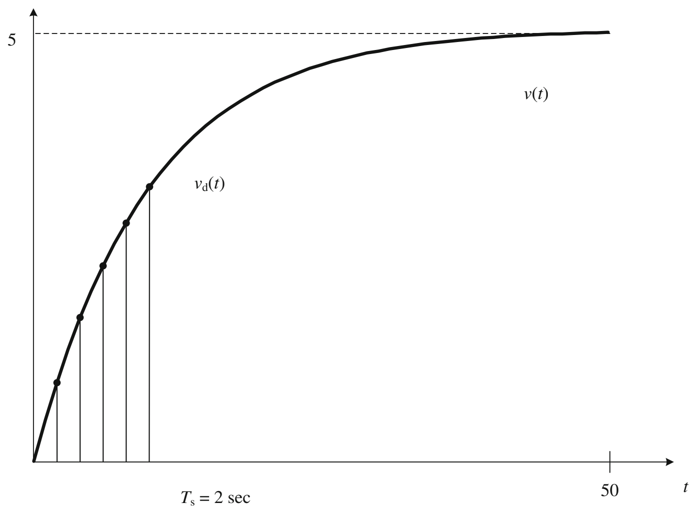


Fig. 11.17 CT and *SRE* DT step responses of a first order lag

Figure 11.17 shows the step response for $K = 5$ and $T = 10$. Check the initial and final value of the step response using the initial value and final theorems:

$$y[0] = \lim_{z\to\infty} Y(z) = \lim_{z\to\infty} \{U(z)G(z)\} = \lim_{z\to\infty} \left\{ \frac{z}{z-1} \frac{K(1-e^{-T_s/T})}{z-e^{-T_s/T}} \right\} = 0$$

and

$$\begin{aligned}\lim_{k\to\infty} y[k] &= \lim_{z\to 1}\left\{\left(1-z^{-1}\right)Y(z)\right\} = \lim_{z\to 1}\left\{\left(1-z^{-1}\right)U(z)G(z)\right\} \\ &= \lim_{z\to 1}\left\{\left(1-z^{-1}\right)\frac{z}{z-1}\frac{K\left(1-e^{-T_s/T}\right)}{z-e^{-T_s/T}}\right\} = K\end{aligned}$$

It may be of interest to analyze the dependence of the coefficients b_1 and a_1 on T and T_s, respectively. Introducing the relative sampling rate as $x = T_s/T$ we get

$$\begin{aligned} b_1 &= K(1-e^{-x}) \\ a_1 &= -e^{-x} \end{aligned} \tag{11.80}$$

It is easy to check the steady-state gain: $G(z=1) = K$. The pole of $G(z)$ turns out to be $p_1^d = -a_1 = e^{-x}$. ■

Example 11.7 To derive the DT *SRE* model of a second order CT process is a far more tiresome procedure, though this relationship is frequently needed. For a process with a complex pole pair $p_{1,2}^c = a \pm jb$, the well-known expressions using the damping factor of ξ and the natural frequency ω_o are $a = -\xi\omega_o$ and $b = \omega_o\sqrt{1-\xi^2}$. The DT pair of complex conjugates takes the following form:

$$\left(z-p_1^d\right)\left(z-\bar{p}_1^d\right) = z^2 - \left[2e^{2aT_s}\cos(bT_s)\right]z + e^{2aT_s}, \tag{11.81}$$

■

11.4.4 Analysis of DT Systems in the Frequency Domain

The analysis of systems driven by sinusoidal signals forms a fundamental method of system analysis. Consider the response of a stable DT model by

$$G(z) = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \cdots + b_n}{z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n} = \frac{\mathcal{B}(z)}{\mathcal{A}(z)} \tag{11.82}$$

driven by a sequence of sinusoidal samples

$$u[k] = K\cos(\omega_o k)$$

From Table 11.1,

$$U(z) = \mathcal{Z}\{u[k]\} = K\frac{z^2 - z\cos(\omega_o T_s)}{z^2 + 2z\cos(\omega_o T_s) + 1},$$

and the z-transform of the output is

$$Y(z) = U(z)G(z) = K\frac{z^2 - z\cos(\omega_o T_s)}{z^2 + 2z\cos(\omega_o T_s) + 1}G(z) = K\frac{z^2 - z\cos(\omega_o T_s)}{z^2 + 2z\cos(\omega_o T_s) + 1}\frac{\mathcal{B}(z)}{\mathcal{A}(z)}$$

Further manipulations give

$$\frac{Y(z)}{z} = K\frac{z - \cos(\omega_o T_s)}{(z - e^{j\omega_o T_s})(z - e^{-j\omega_o T_s})}\frac{\mathcal{B}(z)}{\mathcal{A}(z)}$$

and the *PFE* turns out to be

$$\frac{Y(z)}{z} = \frac{\mathcal{V}(z)}{\mathcal{A}(z)} + \frac{c}{z - e^{j\omega_o T_s}} + \frac{\bar{c}}{z - e^{-j\omega_o T_s}}$$

where $\mathcal{V}(z)$ is a polynomial with degree less than n, while the coefficient c equals

$$c = \left[\left(z - e^{j\omega_o T_s}\right)\frac{Y(z)}{z}\right]_{z=e^{j\omega_o}} = \cdots = \frac{K}{2}G\left(e^{j\omega_o T_s}\right).$$

As $\bar{c}$ is the complex conjugate of c,

$$Y(z) = \frac{z\mathcal{V}(z)}{\mathcal{A}(z)} + \frac{K}{2}\left[\frac{zG(e^{j\omega_o T_s})}{z - e^{j\omega_o T_s}} + \frac{z\bar{G}(e^{j\omega_o T_s})}{z - e^{j\omega_o T_s}}\right]$$

is the sum of a transient $y_{tr}[k]$ and a steady-state component $y_{ss}[k]$:

$$y[k] = y_{tr}[k] + y_{ss}[k]. \tag{11.83}$$

Stability requires $\lim_{k\to\infty} y_{tr}[k] = 0$ and the steady-state component is

$$y_{ss}[k] = K\left|G\left(e^{j\omega_o T_s}\right)\right|\cos(\omega_o T_s k + \Theta), \tag{11.84}$$

where

$$G\left(e^{j\omega_o T_s}\right) = \left|G\left(e^{j\omega_o T_s}\right)\right| e^{j\Theta}. \tag{11.85}$$

The sinusoidal excitation causes the steady-state response of a stable DT system to be sinusoidal with the same frequency but an amplitude of

$$G(z)\big|_{z=e^{j\omega_o T_s}} = G\left(e^{j\omega_o T_s}\right) = \left|G\left(e^{j\omega_o T_s}\right)\right| e^{j\Theta}. \tag{11.86}$$

The above result is in harmony with the result obtained earlier for CT systems. The results are obviously different, however, the DT and CT systems share identical properties.

Based on the discussion of DT systems so far, it is rather general that in the expression of the $G(z)$ pulse transfer function and that of the transfer function operator $G(q)$ involve the variables z^{-1} and q^{-1}, rather than z and q, respectively. To derive the filter forms $G(z^{-1})$ and $G(q^{-1})$ can be derived from $G(z)$ and $G(q)$ by dividing both their numerator and the denominator by the highest power term. These forms are

$$\begin{aligned} G\left(z^{-1}\right) &= \frac{\mathcal{B}(z^{-1})}{\mathcal{A}(z^{-1})} = \frac{b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} \\ &= \frac{\left(b_1 + b_2 z^{-1} + \cdots + b_n z^{-(n-1)}\right) z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} \end{aligned} \tag{11.87}$$

or

$$\begin{aligned} G\left(q^{-1}\right) &= \frac{\mathcal{B}(q^{-1})}{\mathcal{A}(q^{-1})} = \frac{b_1 q^{-1} + b_2 q^{-2} + \cdots + b_n q^{-n}}{1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n}} \\ &= \frac{\left(b_1 + b_2 q^{-1} + \cdots + b_n q^{-(n-1)}\right) q^{-1}}{1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n}} \end{aligned} \tag{11.88}$$

Note that if the transfer function of the CT system is strictly proper, than the *SRE* transformation results in transfer functions $G(z)$ and $G(q)$ with pole excess one. The CT time-delay e^{-sT_d} will be represented by

$$z^{-d} \quad ; \quad q^{-d} \tag{11.89}$$

where

$$d = \mathrm{int}\left\{\frac{T_d}{T_s}\right\} \tag{11.90}$$

with $d = \mathrm{int}\{\ldots\}$ meaning the separation of the integer part. z^{-d} corresponds to the general delay relationship by (11.11). The z^{-1} and q^{-1} terms in the expression of

$G(z^{-1})$ and $G(q^{-1})$, respectively, are not part of the z^{-d} and q^{-d} terms representing the time-delay. Thus for a process containing an actual time-delay it is reasonable to use the pulse transfer function

$$G\left(z^{-1}\right) = \frac{b_1 + b_2 z^{-1} + \cdots + b_n z^{-(n-1)}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} z^{-(d+1)}, \tag{11.91}$$

which is the *SRE* transform of the general CT system given by

$$H(s) = k \frac{\prod_{i=1}^{m} \left(s - z_i^{\mathrm{c}}\right)}{\prod_{i=1}^{n} \left(s - p_i^{\mathrm{c}}\right)} e^{-sT_{\mathrm{d}}}; \quad m < n. \tag{11.92}$$

In the CT state-model form the gain of the direct constant channel is zero $(d = 0)$.

11.4.5 Transformation of Zeros

The allocation of poles and zeros of an *nth* order $G(z)$ pulse transfer function is an interesting issue. Assume the *SRE* transformation is used for the discretization. Based on the example elaborated earlier for a first order lag, it can be seen that the transformation of the poles follows $p_i^{\mathrm{d}} = e^{-x_i}$, where $x_i = T_{\mathrm{s}}/T_i$ and the pole of the CT system is $-1/T_i$. (The relation is valid for complex conjugate poles as well, however, it is reasonable to handle those poles as complex conjugate pairs. Anyway, the transformation relationship will be far more involved.) The exponential relationship will map stable CT poles to stable DT poles, for the complete CT stability region (the left half-plane) is transformed into the stable DT region (the unit disc). Also, unstable CT poles (in the right half-plane) will be transformed into unstable DT poles (outside the unit disc). The transformation of the zeros, however, does not follow this pattern. Note first that $G(z)$ always has $(n - 1)$ zeros. Provided that the CT system has only m stable zeros, then m out of the $(n - 1)$ DT zeros are transformed approximately by $z_i^{\mathrm{d}} \approx e^{-x_i}$ (the exact relationship is far more involved). The rest of the zeros—there are $(n - m - 1)$ of them—follow a different law, as those zeros have no CT counterparts. Typically, these (unmatched) zeros are located on the real axis. If the pole excess is larger than one, even for minimum phase systems (having stable zeros), the unstable zeros (out of the unmatched zeros) should be accounted for. In short, even minimum-phase CT systems may turn into non-minimum phase DT systems after the *SRE* transformation.

Similarly, an important observation is that a fractional time-delay (a delay which is not an integer multiple of the sampling time) may result in unstable zeros. As unstable zeros play an important role in the DT controller design procedure, it should be noted that nonminimum-phase CT systems represent a very special class. However, DT nonminimum-phase systems are rather common.

11.5 Structural Properties of State Equations

The state-space description of a sampled data system consists of the state difference equation

$$\boldsymbol{x}[k+1] = \boldsymbol{F}\boldsymbol{x}[k] + \boldsymbol{g}u[k] \tag{11.93}$$

and

$$y[k] = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}[k] + du[k]. \tag{11.94}$$

Introducing a shift element the realization of the system is shown in Fig. 11.18.

The state difference equation describes the development of the states of the system, whereas the output equation gives how the output depends on the state variables and incidentally directly on the input. The pulse transfer function of the system (supposing zero initial conditions) is given by

$$G(z) = \frac{\mathcal{Z}\{y[k]\}}{\mathcal{Z}\{u[k]\}} = \boldsymbol{c}^{\mathrm{T}}(z\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g} + d. \tag{11.95}$$

Similarly to CT systems, an infinite number of state space representations can be derived from a given pulse transfer function, which provide the same output signal for a given input signal (therefore these representations are called input-output equivalent descriptions). To verify this statement, in the sequel several different input-output equivalent state space descriptions will be derived from a given pulse transfer function.

Example 11.8 Consider the pulse transfer function

$$G(z) = \frac{b'_0 z^3 + b'_1 z^2 + b'_2 z + b'_3}{z^3 + a_1 z^2 + a_2 z + a_3} = \frac{b_1 z^2 + b_2 z + b_3}{z^3 + a_1 z^2 + a_2 z + a_3} + d, \tag{11.96}$$

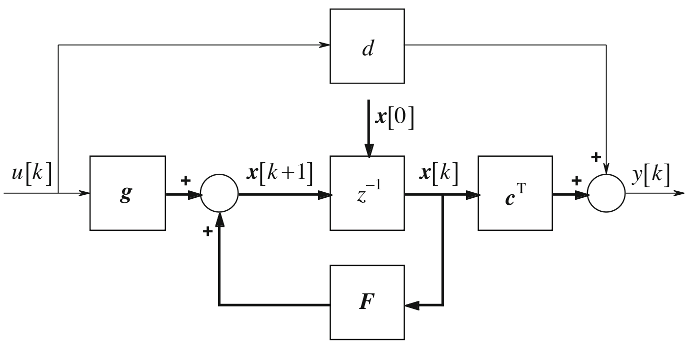


Fig. 11.18 Block-diagram representing the state equation of a discrete system

Transform it to the form

$$G(z) = \frac{b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} + d. \tag{11.97}$$

Then

$$Y(z) = \frac{b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} U(z) + dU(z). \tag{11.98}$$

Observe that the value of d is non-zero only if the degrees of the numerator and of the denominator of $G(z)$ are identical (the degree of the numerator can be at most the same as that of the denominator).

Let us realize first the relationship

$$Y(z) = \left[\left(b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}\right) / \left(1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}\right)\right] U(z),$$

then the value of $dU(z)$ has to be added to the output. By cross multiplication

$$Y(z) + a_1 z^{-1} Y(z) + a_2 z^{-2} Y(z) + a_3 z^{-3} Y(z) = b_1 z^{-1} U(z) + b_2 z^{-2} U(z) + b_3 z^{-3} U(z) \tag{11.99}$$

is obtained. Rearranging the equation $Y(z)$ can be expressed as

$$Y(z) = [b_1 U(z) - a_1 Y(z)] z^{-1} + [b_2 U(z) - a_2 Y(z)] z^{-2} + [b_3 U(z) - a_3 Y(z)] z^{-3}. \tag{11.100}$$

The realization will provide a block-scheme, which besides the constants and the summation elements, contains shift blocks with pulse transfer function of z^{-1}, realizing a one step delay. As seen from the equation above, the signal $[b_1 U(z) - a_1 Y(z)]$ has to pass through one delay element, the signal $[b_2 U(z) - a_2 Y(z)]$ has to pass through two, whereas signal $[b_3 U(z) - a_3 Y(z)]$ has to go through three. Based on these considerations the block-diagram showing the realization is drawn in Fig. 11.19. The state variables can be chosen at the outputs of the shift elements. Thus the state equation can be written as

$$\begin{aligned} x_1[k+1] &= -a_3 x_3[k] + b_3 u[k] \\ x_2[k+1] &= x_1[k] - a_2 x_3[k] + b_2 u[k] \\ x_3[k+1] &= x_2[k] - a_1 x_3[k] + b_1 u[k] \\ y[k] &= x_3[k] + du[k] \end{aligned} \tag{11.101}$$

In the figure the time domain and z domain notations appear at the same time, thus the equations of the given realization can be directly "read": if the output of a shift element is $x_i[k]$, then its input is $x_i[k+1]$. It is easily seen in the figure that a

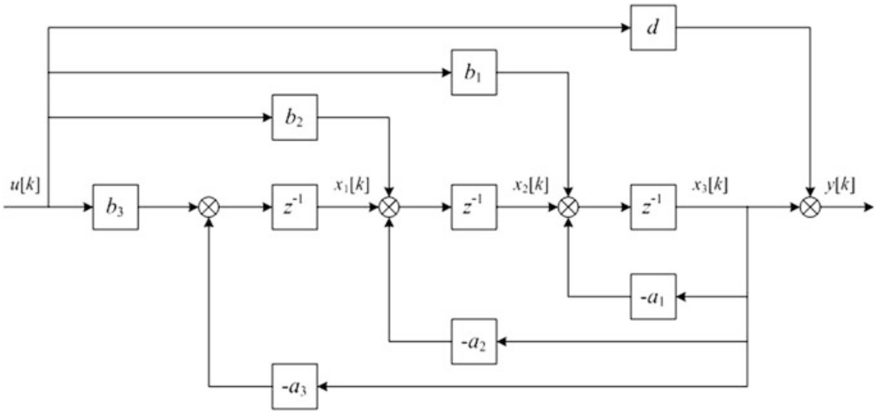


Fig. 11.19 Block-scheme of the observable canonical form

direct channel between $u[k]$ and $y[k]$ without dynamics exists only if $d \neq 0$. This means that an abrupt change in the input signal causes an immediate change in the output signal only if the degrees of the numerator and of the denominator of the pulse transfer function are identical. If this condition does not hold, then a change in the input signal affects the output signal only in later steps. Finally, arranging the equations derived above according to the state space form yields

$$\begin{aligned} \boldsymbol{x}[k+1] &= \begin{bmatrix} x_1[k+1] \\ x_2[k+1] \\ x_3[k+1] \end{bmatrix} = \begin{bmatrix} 0 & 0 & -a_3 \\ 1 & 0 & -a_2 \\ 0 & 1 & -a_1 \end{bmatrix} \boldsymbol{x}[k] + \begin{bmatrix} b_3 \\ b_2 \\ b_1 \end{bmatrix} u[k] = \bar{\boldsymbol{F}}_{\mathrm{o}} \boldsymbol{x}[k] + \bar{\boldsymbol{g}}_{\mathrm{o}} u[k] \\ y[k] &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{bmatrix} + du[k] = \bar{\boldsymbol{c}}_{\mathrm{o}}^{\mathrm{T}} \boldsymbol{x}[k] + du[k] \end{aligned} \tag{11.102}$$

Note that the derived state space description gives the so called *observable canonical form* (see the explanation of this phrase around the formula (3.55) in the CT case).

Derive now another realization of the subsystem by

$$Y(z) = \left[\left(b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}\right) / \left(1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}\right)\right] U(z)$$

(the realization of the dynamic subsystem will be extended by the $du[k]$ element afterwards). The method is the following: after cross multiplication an intermediate variable is defined; first this variable is created from the input signals, then the output is built using the intermediate $v[k]$ variable. Now the pulse transfer function is arranged as

$$\frac{Y(z)}{b_1z^{-1}+b_2z^{-2}+b_3z^{-3}}=\frac{U(z)}{1+a_1z^{-1}+a_2z^{-2}+a_3z^{-3}}=V(z). \tag{11.103}$$

Let us consider first the equation

$$\frac{U(z)}{1+a_1z^{-1}+a_2z^{-2}+a_3z^{-3}}=V(z), \tag{11.104}$$

which in the time domain provides the recursive relationship

$$v[k]=u[k]-a_1v[k-1]-a_2v[k-2]-a_3v[k-3]. \tag{11.105}$$

Regarding the realization let us start with three serially connected shift elements driven by the input $u[k]$ (Fig. 11.20). Coming back to the equation $V(z)=Y(z)/(b_1z^{-1}+b_2z^{-2}+b_3z^{-3})$ the difference equation in the time domain describing the output signal can be written as $y[k]=b_1v[k-1]+b_2v[k-2]+b_3v[k-3]$, thus the whole realization can be drawn as given in Fig. 11.21. Based on Fig. 11.21 the following equations can be written:

$$\begin{aligned} x_1[k+1]&=-a_1x_1[k]-a_2x_2[k]-a_3x_3[k]+u[k]\\ x_2[k+1]&=x_1[k]\\ x_3[k+1]&=x_2[k]\\ y[k]&=b_1x_1[k]+b_2x_2[k]+b_3x_3[k]+du[k] \end{aligned} \tag{11.106}$$

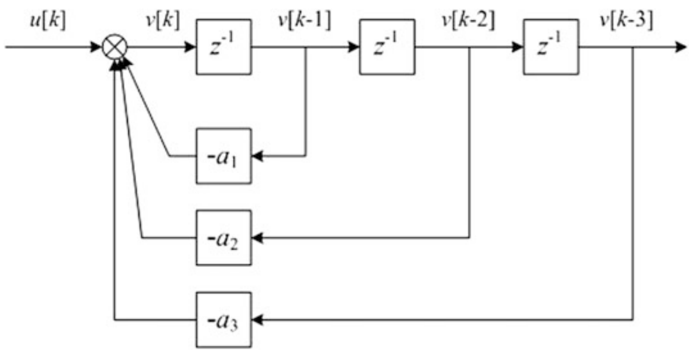


Fig. 11.20 Creating an intermediate variable

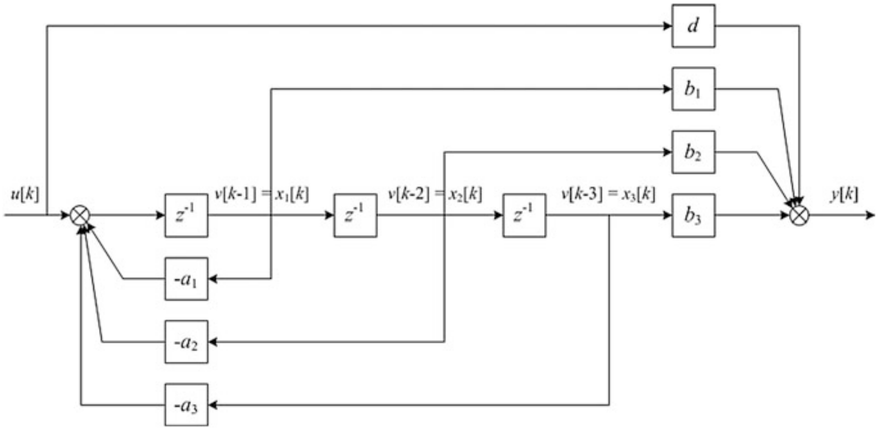


Fig. 11.21 Realization of the controllable canonical form

Thus the state space model is

$$\boldsymbol{x}[k+1] = \begin{bmatrix} x_1[k+1] \\ x_2[k+1] \\ x_3[k+1] \end{bmatrix} = \begin{bmatrix} -a_1 & -a_2 & -a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \boldsymbol{x}[k] + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u[k] = \boldsymbol{F}_{\mathrm{c}}\boldsymbol{x}[k] + \boldsymbol{g}_{\mathrm{c}}u[k]$$

$$y[k] = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{bmatrix} + du[k] = \boldsymbol{c}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{x}[k] + du[k] \tag{11.107}$$

This model is called the *controllable canonical form* (see the explanation of this phrase for the formula (3.46) in the CT case). ■

The solution of the DT state space system in the form $\boldsymbol{x}[k] = (q\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{g}u[k]$ directly shows that in the case of a diagonal state transition matrix $\boldsymbol{F}$ the matrix inversion can be avoided; therefore transforming the pulse transfer function to a form which results in a diagonal representation will lead to an advantageous structure. Generally let $G(z)$ be

$$G(z) = \frac{b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n-1} z^{-(n-1)}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} + d = \sum_{i=1}^{n} \frac{\beta_i z^{-1}}{1 + \alpha_i z^{-1}} + d, \tag{11.108}$$

then introducing the state variables by

$$x_i[k+1] = \beta_i u[k] - \alpha_i x_i[k] \tag{11.109}$$

the state difference equation will be

$$\boldsymbol{x}[k+1] = \begin{bmatrix} x_1[k+1] \\ x_2[k+1] \\ \vdots \\ x_n[k+1] \end{bmatrix} = \begin{bmatrix} -\alpha_1 & 0 & 0 & 0 \\ 0 & -\alpha_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & -\alpha_n \end{bmatrix} \begin{bmatrix} x_1[k] \\ x_2[k] \\ \vdots \\ x_n[k] \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} u[k]$$

$$= \boldsymbol{F}_{\mathrm{d}}\boldsymbol{x}[k] + \boldsymbol{g}_{\mathrm{d}}u[k] \tag{11.110}$$

and the output equation is given by

$$y[k] = [1 \quad 1 \quad \ldots \quad 1]\boldsymbol{x}[k] + du[k] = \boldsymbol{c}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{x}[k] + du[k]. \tag{11.111}$$

It was supposed that the poles of $G(z)$ are single and real. The derived form is also called *parallel canonical form*. The realization is shown in Fig. 11.22.

It can be easily seen that besides $\boldsymbol{F}_{\mathrm{d}} = \mathbf{diag}[-\alpha_1, -\alpha_2, \ldots, -\alpha_n]$ the state space model with

$$\begin{aligned} \boldsymbol{x}[k+1] &= \boldsymbol{F}_{\mathrm{d}}\boldsymbol{x}[k] + \boldsymbol{g}_{\mathrm{d}}u[k] \\ y[k] &= \boldsymbol{c}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{x}[k] + du[k] \end{aligned} \tag{11.112}$$

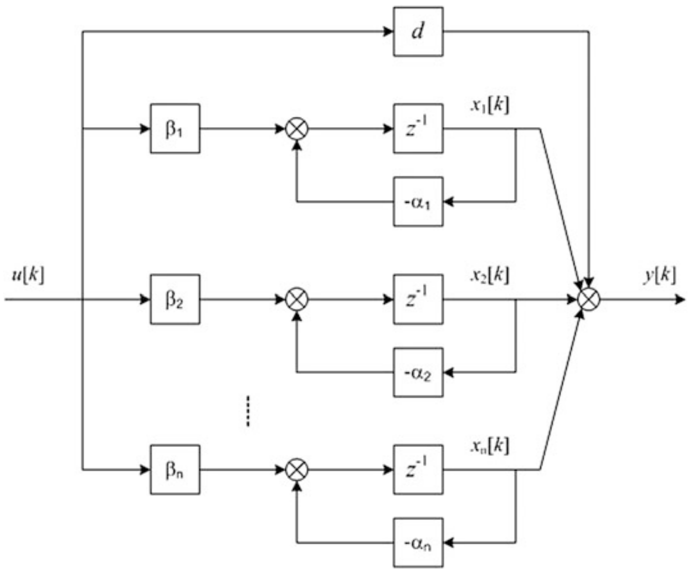


Fig. 11.22 Block-diagram of the parallel canonical form

gives a parallel canonical form, where $g_i c_i = g_i = \beta_i$ is fulfilled. (Comparing this with the CT forms (3.38), (3.39) it can be seen that in the DT form, simply $\gamma_i = 1$ was chosen.)

(In the state space forms—following the conventions—the gain of the direct constant channel was denoted by d, whose value is the same both for the CT and the DT case. As in the DT case its value generally is zero, it was not so disturbing, that the same letter was used also for the discrete dead-time.)