

## Computer Control Systems

*In this chapter we present the elements and the basic concepts of computer-controlled systems. The discretization and choice of sampling frequency will be first examined, followed by a study of discrete-time models in the time and frequency domains, discrete-time systems in closed loop and basic principles for designing digital controllers.*

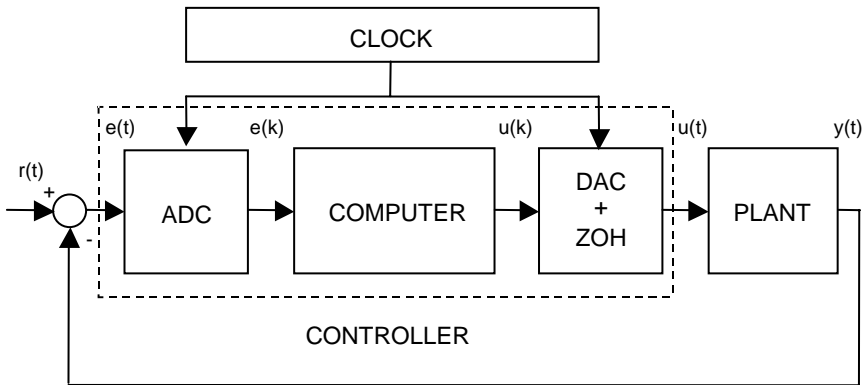
### 2.1 Introduction to Computer Control

The first approach for introducing a digital computer or a microprocessor into a control loop is indicated in Figure 2.1. The measured error between the reference and the output of the plant is converted into digital form by an analog-to-digital converter (ADC), at sampling instants  $k$  defined by the synchronization clock. The computer interprets the converted signal  $y(k)$  as a sequence of numbers, which it processes using a **control algorithm** and generates a new sequence of numbers  $\{u(k)\}$  representing the control. By means of a digital-to-analog converter (DAC), this sequence is converted into an analog signal, which is maintained constant between the sampling instants by a zero-order hold (ZOH). The cascade: ADC-computer-DAC should behave in the same way as an analog controller (PID type), which implies the use of a high sampling frequency but the algorithm implemented on the computer is very simple (we just do not make use of the potentialities of the digital computer!).

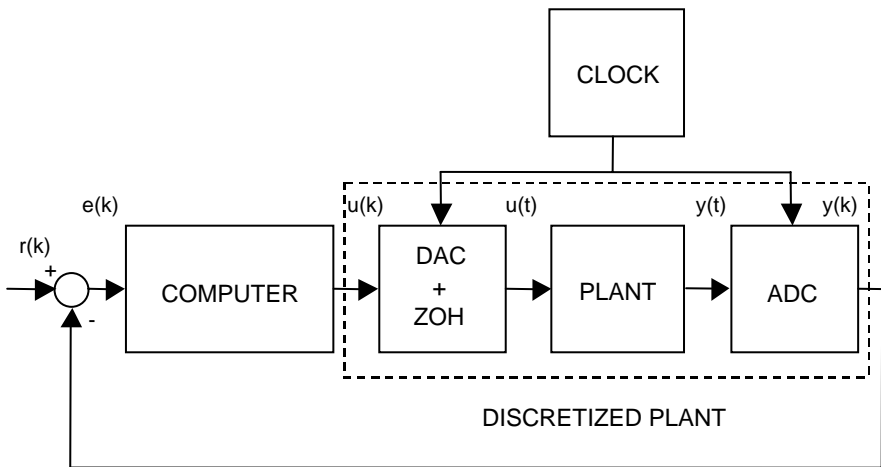
A second and much more interesting approach for the introduction of a digital computer or microprocessor in a control loop is illustrated in Figure 2.2 which can be obtained from Figure 2.1 by moving the reference-output comparator after the analog-to-digital converter. The reference is now specified in a digital way as a sequence provided by a computer.

In Figure 2.2 the set DAC - plant - ADC is interpreted as a discretized system, whose control input is the sequence  $\{u(k)\}$  generated by the computer, the output being the sequence  $\{y(k)\}$  resulting from the A/D conversion of the system output  $y(t)$ . This discretized system is characterized by a “**discrete-time model**”, which

describes the relation between the sequence of numbers  $\{u(k)\}$  and the sequence of numbers  $\{y(k)\}$ . This model is related to the continuous-time model of the plant.



**Figure 2.1.** Digital realization of an « analog » type controller



**Figure 2.2.** Digital control system

This approach offers several advantages. Among these advantages here we recall the following:

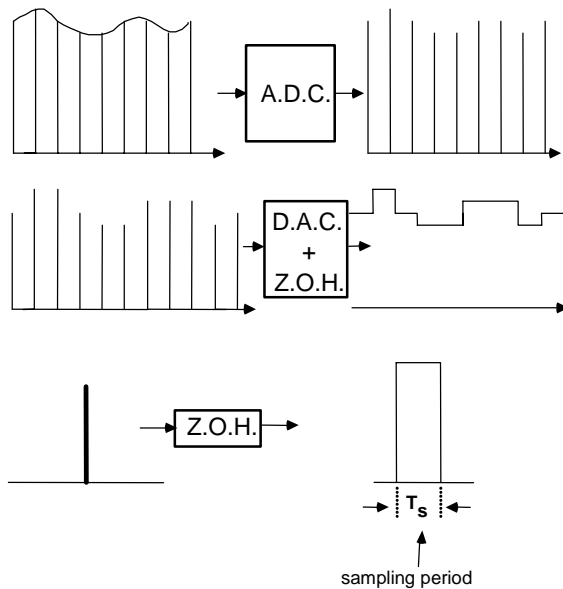
1. The sampling frequency is chosen in accordance with the “bandwidth” of the continuous-time system (it will be much lower than for the first approach).

2. Possibility of a direct design of the control algorithms tailored to the discretized plant models.
3. Efficient use of the computer since the increase of the sampling period permits the computation power to be used in order to implement algorithms which are more performant but more complex than a PID controller, and which require a longer computation time.

*In fact, if one really wants to take advantage of the use of a digital computer in a control loop, the “language” must also be changed. This may be achieved by replacing the continuous-time system models by discrete-time system models, the continuous-time controllers by digital control algorithms, and by using dedicated control design techniques.*

The changing over to this new “language” (discrete-time dynamic models) makes it possible to use various high performing control strategies which cannot be implemented by analog controllers.

The operating details of the ADC (analog-to-digital converter), the DAC (digital-to-analog converter) and the ZOH (zero-order hold) are illustrated in Figure 2.3.



**Figure 2.3.** Operation of the analog-to-digital converter (ADC), the digital-to-analog converter (DAC) and the zero-order hold (ZOH)

The analog-to-digital converter implements two functions:

1. Analog signal sampling: this operation consists in the replacement of the continuous signal with a sequence of values equally spaced in the time domain (the temporal distance between two values is the sampling period), as these values correspond to the continuous signal amplitude at sampling instants.
2. Quantization: this is the operation by means of which the amplitude of a signal is represented with a discrete set of different values (quantized values of the signal), generally coded with a binary sequence.

The general use of high-resolution A/D converters (where the samples are coded with 12 bits or more) allows one to consider the quantification effects as negligible, and this assumption will hold in the following. Quantization effects will be taken into account in Chapter 8.

The digital-analog converter (DAC) converts at the sampling instants a discrete signal, digitally coded, in a continuous signal.

The zero-order hold (ZOH) keeps constant this continuous signal between two sampling instants (sampling period), in order to provide a continuous-time signal.

## 2.2 Discretization and Overview of Sampled-data Systems

### 2.2.1 Discretization and Choice of Sampling Frequency

Figure 2.4 illustrates the discretization of a sinusoid of frequency  $f_0$  for several sampling frequencies  $f_s$ .

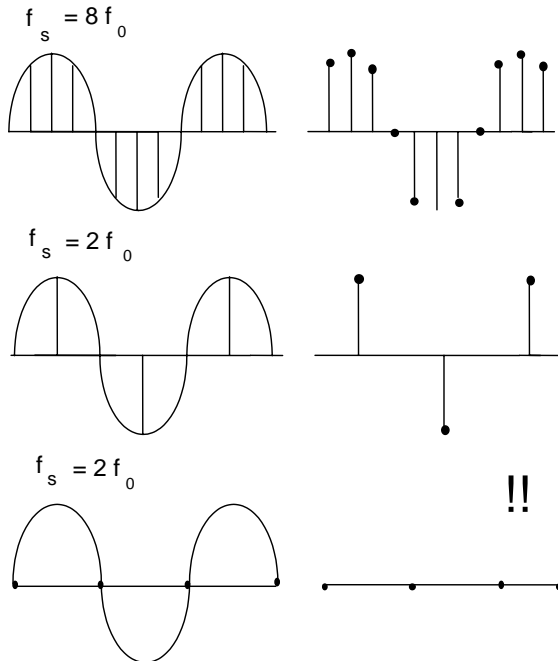
It can be noted that, for a sampling frequency  $f_s = 8f_0$ , the continuous nature of the analog signal is unaltered in the sampled signal.

For the sampling frequency  $f_s = 2f_0$ , if the sampling is carried out at instants  $2\pi f_0 t$  other than multiples of  $\pi$ , a periodic sampled signal is still obtained. However if the sampling is carried out at the instants where  $2\pi f_0 t = n\pi$ , the corresponding sampled sequence is identically zero.

If the sampling frequency is decreased under the limit of  $f_s = 2f_0$ , a periodic sampled signal still appears, but its frequency differs from that of the continuous signal ( $f = f_s - f_0$ ).

In order to reconstruct a continuous signal from the sampled sequence, the sampling frequency must verify the condition (Nyquist's theorem):

$$f_s > 2f_{max} \quad (2.2.1)$$



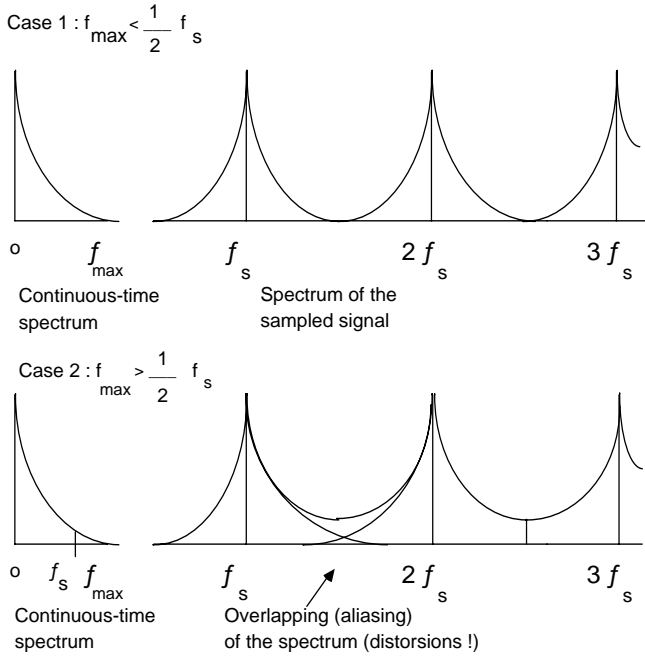
**Figure 2.4.** Sinusoidal signal discretization

in which  $f_{max}$  is the maximum frequency to be transmitted. The frequency  $f_s = 2f_{max}$  is a theoretical limit; in practice, a higher sampling frequency must be chosen.

The existence of a maximum limit for the frequency that may be converted without distortion, for a given sampling frequency, is also understandable when it is observed that the sampling of a continuous-time signal is a “magnitude modulation” of a “carrier” frequency  $f_s$  (analogy with the magnitude modulation in radio transmitters). The modulation effect may be observed in the replication of the spectrum of the modulating signal (in our case the continuous signal) around the sampling frequency and its multiples.

The spectrum of the sampled signal, if the maximum frequency of the continuous signal ( $f_{max}$ ) is less than  $(1/2)f_s$ , is represented in the upper part of Figure 2.5.

The spectrum of the sampled signal, if  $f_{max} > (1/2)f_s$ , is represented in the lower part of Figure 2.5. The phenomenon of overlapping (aliasing) can be observed. This corresponds to the appearance of distortions. The frequency  $(1/2)f_s$ , which defines the maximum frequency ( $f_{max}$ ) admitted for a sampling with no distortions, is known as “Nyquist frequency” (or Shannon frequency).

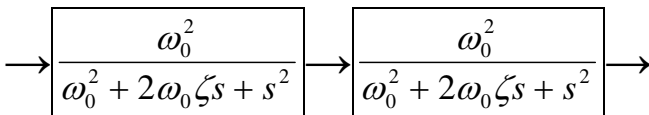


**Figure 2.5.** Spectrum of a sampled signal

For a given sampling frequency, in order to avoid the *folding* (aliasing) of the spectrum and thus of the distortions, the analog signals must be filtered prior to sampling to ensure that:

$$f_{\max} < \frac{1}{2} f_s \tag{2.2.2}$$

The filters used are known as “anti-aliasing filters”. A good anti-aliasing filter must have a minimum of two cascaded second-order cells ( $f_{\max} \ll (1/2) f_s$ ). An example of an anti-aliasing filter of this type is given in Figure 2.6. These filters must introduce a large attenuation at frequencies higher than  $(1/2) f_s$ , but their bandwidth must be higher than the required bandwidth of the closed loop system (generally higher than open loop system bandwidth). Circuits of this type (or more complex) are currently available.



**Figure 2.6.** Anti-aliasing filter

In the case of very low frequency sampling, first a sampling at a higher frequency is carried out (integer multiple of the desired frequency), using an appropriate analog anti-aliasing filter. The sampled signal thus obtained is passed through a digital anti-aliasing filter followed by a frequency divider (decimation) thereby giving a sampled signal having the required frequency. This procedure is shown in Figure 2.7. It is also employed every time the frequency of data acquisition is higher than the sampling frequency chosen for the loop that must be controlled (the sampling frequency should be an integer divider of the acquisition frequency).

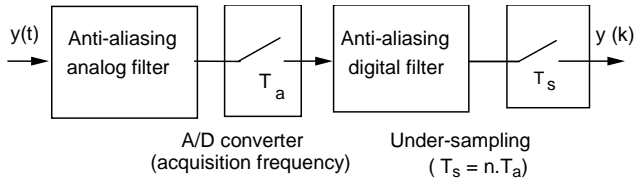


Figure 2.7. Anti-aliasing filtering with under-sampling

## 2.2.2 Choice of the Sampling Frequency for Control Systems

The sampling frequency for digital control systems is chosen according to the desired bandwidth of the closed loop system. Note that, no matter how the desired performances are specified, these can always be related to the closed loop system bandwidth.

*Example:* Let us consider the performances imposed in Section 1.1.6 on the step response (maximum overshoot 5%, rise time 2.75 s). The transfer function to be determined corresponds to the desired closed loop system transfer function. From the diagrams given in Figure 1.11 we have deduced that the closed loop transfer function must be a normalized second-order transfer function with  $\zeta=0.7$  and  $\omega_0=1 \text{ rad/s}$ . By immediately using the diagrams given in Figure 1.12, it can be observed that the bandwidth of the closed loop system is approximately equal to

$$f_B^{CL} = \frac{1}{2\pi} \text{ Hz}$$

The rule used to choose the sampling frequency in control systems is the following:

$$f_s = (6 \text{ to } 25) f_B^{CL} \quad (2.2.3)$$

where

$f_s$ : sampling frequency,  $f_B^{CL}$ : closed loop system bandwidth

Rule of Equation 2.2.3 is equally used in open loop, when it is desired to choose the sampling frequency in order to identify the discrete-time model of a plant. In this case  $f_B^{CL}$  is replaced by an estimation of the bandwidth of the plant.

For information purposes, Table 2.1 gives the sampling periods ( $T_s = 1/f_s$ ) used for the digital control of different types of plants.

The rule for choosing the sampling frequency given in Equation 2.2.3 can be connected to the transfer function parameters.

*First-order system*

$$H(s) = \frac{1}{1 + sT_0}$$

In this case the system bandwidth is

$$f_B = f_0 = \frac{1}{2\pi T_0}$$

(an attenuation greater than 3 db is introduced for frequencies higher than  $\omega_0 = 1/T_0 = 2\pi f_0$ ).

**Table 2.1.** Choice of the sampling period for digital control systems (indicative values)

Type of variable (or plant)	Sampling period (s)
Flow rate	1 – 3
Level	5 – 10
Pressure	1 – 5
Temperature	10 - 180
Distillation	10 - 180
Servo-mechanisms	0.001 - 0.05
Catalytic reactors	10 - 45
Cement plants	20 - 45
Dryers	20 – 45



By applying the rule of Equation 2.2.3 the condition for choosing the sampling period is obtained ( $T_s = 1/f_s$ ):

$$\frac{T_0}{4} < T_s < T_0 \quad (2.2.4)$$

This corresponds to the existence of two to nine samples on the rise time of a step response.

*Second- order system*

$$H(s) = \frac{\omega_0^2}{\omega_0^2 + 2\zeta\omega_0s + s^2}$$

The bandwidth of the second-order system depends on  $\omega_0$  and on  $\zeta$  (see Figure 1.12).

For example:

$$\zeta = 0.7 \Rightarrow f_B = \frac{\omega_0}{2\pi}$$

$$\zeta = 1 \Rightarrow f_B = \frac{0.6\omega_0}{2\pi}$$

By applying the rule of Equation 2.2.3, the following relations are obtained between the natural frequency  $\omega_0$  and the sampling period  $T_s$  :

$$0.25 \leq \omega_0 T_s \leq 1 \quad ; \quad \zeta = 0.7 \quad (2.2.5)$$

and

$$0.4 \leq \omega_0 T_s \leq 1.75 \quad ; \quad \zeta = 1 \quad (2.2.6)$$

The lower values correspond to the choice of a high sampling frequency and the upper values to the choice of a low sampling frequency.

For simplicity's sake, given that in closed loop the behavior frequently chosen as the desired behavior is that of a second order having a damping factor  $\zeta$  between 0.7 and 1, the following rule can be used (approximation of Equations 2.2.5 and 2.2.6):

$$0.25 \leq \omega_0 T_s \leq 1.5 \quad ; \quad 0.7 \leq \zeta \leq 1 \quad (2.2.7)$$

## 2.3 Discrete-time Models

### 2.3.1 Time Domain

Figure 2.8 illustrates the response of a continuous-time system to a step input, a response that can be simulated by a first order system (an integrator with a feedback gain indicated in the figure).

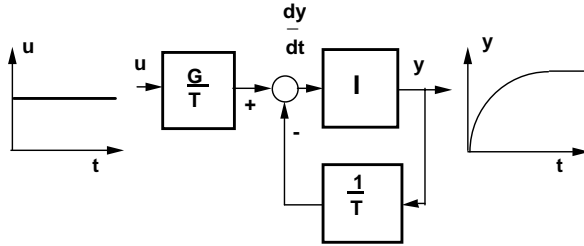


Figure 2.8. Continuous-time model

The corresponding model is described by the differential equation

$$\frac{dy}{dt} = -\frac{1}{T} y(t) + \frac{G}{T} u(t) \tag{2.3.1}$$

or by the transfer function

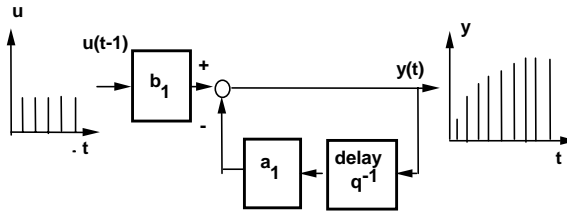
$$H(s) = \frac{G}{1 + sT} \tag{2.3.2}$$

where  $T$  is the time constant of the system and  $G$  is the gain.

If the input  $u(t)$  and the output  $y(t)$  are sampled with a specified sampling period, the representations of  $u(t)$  and  $y(t)$  are obtained as number sequences in which  $t$  (or  $k$ ) is now the *normalized discrete-time* (real time divided by the sampling period,  $t = t/T_s$ ). The relation between the input sequence  $\{u(t)\}$  and the output sequence  $\{y(t)\}$  can be simulated by the scheme given in Figure 2.9 by using a delay (backward shift) operator (symbolized by  $q^{-1}$ :  $y(t-1) = q^{-1} y(t)$ ), instead of an integrator.

This relation is described in the time domain by the algorithm (known as recursive equation or difference equation)

$$y(t) = -a_1 y(t-1) + b_1 u(t-1) \tag{2.3.3}$$



**Figure 2.9.** Discrete-time model

Let us now examine in greater detail the discrete-time model given by Equation 2.3.3 for a zero initial condition ( $y(0) = 0$ ) and a discrete-time unit step input:

$$u(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

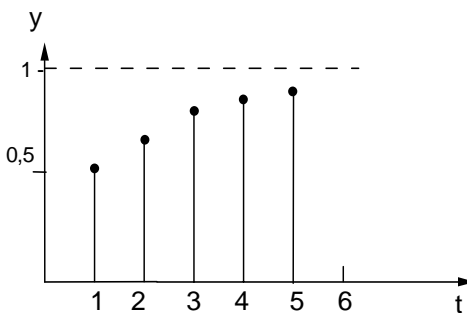
The response is directly computed by recursively using Equation 2.3.3 from  $t = 0$  (in the case of discrete-time models there are no problem with the integration of the differential equations like in continuous time). We shall examine two cases.

*Case 1.*  $a_1 = -0.5$  ;  $b_1 = 0.5$

The output values for different instants are given in Table 2.2 and the corresponding sequence is represented in Figure 2.10.

**Table 2.2.** Step response of a first-order discrete-time model ( $a_1 = -0.5, b_1 = 0.5$ )

T	0	1	2	3	4	5
y(t)	0	0.5	0.75	0.875	0.937	0.969



**Figure 2.10.** Step response of a first-order discrete-time model ( $a_1 = -0.5, b_1 = 0.5$ )

It is observed that the response obtained resembles the step response of a continuous-time first order system which has been sampled. An equivalent time

constant for the continuous-time system can even be determined (rise time from 0 to 90 %:  $t_R = 2.2 T$ ). From Table 2.2, one then obtains

$$\frac{3T_s}{2.2} < T < \frac{4T_s}{2.2}$$

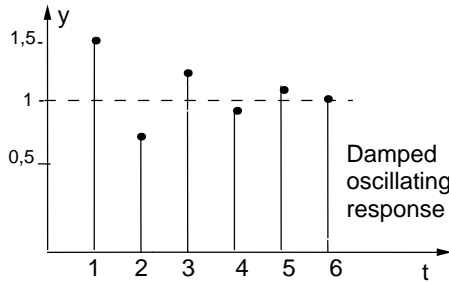
Case 2.  $a_1 = 0.5$  ;  $b_1 = 1.5$

Output values for different instants are given in Table 2.3 and the corresponding sequence is represented in Figure 2.11.

**Table 2.3.** Step response of a first-order discrete-time model ( $a_1=0.5$ ;  $b_1=1.5$ )

T	0	1	2	3	4	5
y(t)	0	1.5	0.75	1.125	0.937	1.062

An oscillatory damped response is observed with a period equal to two sampling periods. This type of phenomenon cannot result from the discretization of a continuous-time first order system, since this latter is always a-periodic. It may thus be concluded that the first order discrete-time model corresponds to the discretization of a first order continuous-time system only if  $a_1$  is negative<sup>1</sup>.



**Figure 2.11.** Step response of a first-order discrete-time model ( $a_1=0.5$ ;  $b_1=1.5$ )

We now go back to the method used to describe discrete-time models. The delay operator  $q^{-1}$  is used to obtain a more compact writing of the recursive (difference) equations which describe discrete-time models in the time domain (it has the same function as the operator  $p = d/dt$  for continuous-time systems). The following relations hold:

---

<sup>1</sup> For a positive  $a_1$ , this corresponds to the discretization of a 2<sup>nd</sup> order system, with a damped resonant frequency equal to  $0.5f_s$  (see Section 2.3.2).

$$\begin{aligned} q^{-1}y(t) &= y(t-1) \\ q^{-d}y(t) &= y(t-d) \end{aligned} \quad (2.3.4)$$

By using the operator  $q^{-1}$ , Equation 2.3.3 is rewritten as

$$(1 + a_1 q^{-1})y(t) = b_1 q^{-1}u(t) \quad (2.3.5)$$

Discrete-time models may also be obtained by the discretization of the differential equations describing continuous-time models. This operation is used for the simulation of continuous-time models on a digital computer.

Let us consider Equation 2.3.1 and approximate the derivative by

$$\frac{dy}{dt} = \frac{y(t+T_s) - y(t)}{T_s} \quad (2.3.6)$$

Equation 2.3.1 will be rewritten as

$$\frac{y(t+T_s) - y(t)}{T_s} + \frac{1}{T}y(t) = \frac{G}{T}u(t) \quad (2.3.7)$$

By multiplying both sides of Equation 2.3.7 by  $T_s$ , and with the introduction of the normalized time  $t (= t/T_s)$ , it follows that

$$y(t+1) + \left(\frac{T_s}{T} - 1\right)y(t) = \frac{G}{T}T_s u(t) \quad (2.3.8)$$

which can be further rewritten as:

$$(1 + a_1 q^{-1})y(t+1) = b_1 u(t) \quad (2.3.9)$$

where

$$a_1 = \frac{T_s}{T} - 1 \quad (< 0) ; \quad b_1 = \frac{G}{T}T_s$$

Shifting Equation 2.3.9 by one step, Equation 2.3.3 is obtained.

We point out that, in order to represent a first-order continuous model with Equation 2.3.9, the condition  $a_1 < 0$  must be verified. As a consequence, the sampling period  $T_s$  must be smaller than time constant  $T$  ( $T_s < T$ ). This result corresponds to the upper bound in Equation 2.2.4, introduced for sampling period selection of a first-order system as a function of the desired closed loop bandwidth.

If Equation 2.3.6 is the approximation of the “derivative”, the digital integrator equation can be directly deduced. Thus, if normalized time is used, Equation 2.3.6 is written as

$$\frac{d}{dt} y = py \approx y(t) - y(t-1) = (1 - q^{-1})y(t) \tag{2.3.10}$$

where  $(1 - q^{-1})$  is now equivalent to  $p$ . As the integration is the opposite of the differentiation, one obtains:

$$s(t) = \int y dt = \frac{1}{p}y \approx \frac{1}{1 - q^{-1}}y(t) \tag{2.3.11}$$

Multiplying both sides of Equation 2.3.11 by  $(1 - q^{-1})$ , it follows that

$$s(t) (1 - q^{-1}) = y(t) \tag{2.3.12}$$

which we can rewrite as

$$s(t) = s(t-1) + 1 \cdot y(t) \tag{2.3.13}$$

corresponding to the approximation of the integration operation by means of the rectangular rule, as illustrated in Figure 2.12 (if continuous-time is used, Equation 2.3.13 is written as  $s(t) = s(t-T_s) + T_s \cdot y(t)$ ).

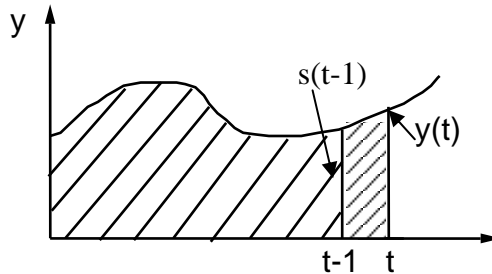


Figure 2.12. Numerical integration

### 2.3.2 Frequency Domain

The study of continuous-time models in the frequency domain has been carried out considering a periodic input of the complex exponential type

$$e^{j\omega t} = \cos \omega t + j \sin \omega t$$

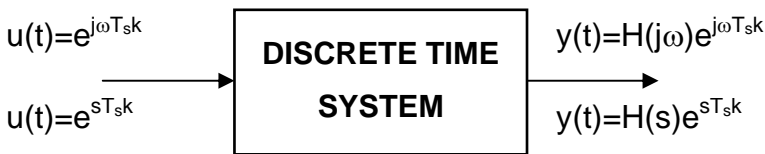
or  $e^{st}$  with  $s = \sigma + j\omega$ .

For the study of discrete-time models in the frequency domain we shall consider complex (sampled) exponentials, i.e. sequences resulting from complex continuous-time exponentials evaluated at the sampling instants  $t = k T_s$ .

These sequences will thus be written as

$$e^{j\omega T_s k} ; e^{s T_s k} ; k=1,2,3\dots$$

Since the discrete-time models being considered are linear, if a signal of a certain frequency is applied to the input, a signal of the same frequency, but amplified or attenuated according to the frequency, will be found at the output. This is summarized in Figure 2.13. in which  $H(s)$  is the “transfer function” of the system that expresses the dependence of the gain and the phase-deviation on the complex frequency  $s$  ( $s = \sigma + j\omega$ ).



**Figure 2.13.** Frequency response of a discrete-time system

If the input of the system is in the form  $e^{s T_s k}$ , the output will be

$$y(t) = H(s)e^{s T_s k} \tag{2.3.14}$$

and respectively

$$y(t - 1) = H(s)e^{s T_s (k-1)} = e^{-s T_s} H(s)e^{s T_s k} = e^{-s T_s} y(t) \tag{2.3.15}$$

It is thus observed that shifting backward by one step is equivalent to multiplying by  $e^{-s T_s}$ .

Let now determine the transfer function related to the recursive Equation 2.3.3. In this case  $u(t) = e^{s T_s k}$  and the output will be in the form of Equation 2.3.14. By also using Equation 2.3.15 one obtains:

$$(1 + a_1 e^{-s T_s})H(s)e^{s T_s k} = b_1 e^{-s T_s} e^{s T_s k} \tag{2.3.16}$$

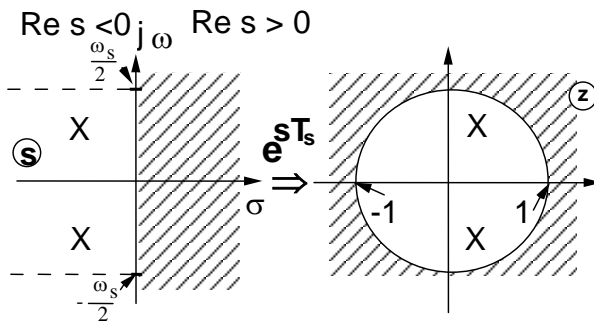
from which results

$$H(s) = \frac{b_1 e^{-sT_s}}{1 + a_1 e^{-sT_s}} \quad (2.3.17)$$

We consider now the following change of variable:

$$z = e^{sT_s} \quad (2.3.18)$$

which corresponds to the transformation of the left half-plane of the  $s$ -plane into the interior of the unit circle centered at the origin in the  $z$ - plane, as illustrated by Figure 2.14.



**Figure 2.14.** Effect of the transformation  $z = e^{sT_s}$

With the transformation given by Equation 2.3.18 the transfer function given in Equation 2.3.17 becomes

$$H(z^{-1}) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1}} \quad (2.3.19)$$

Note that the transfer function in  $z^{-1}$  can be directly obtained from the recursive Equation 2.3.3 by using the delay operator  $q^{-1}$  (see Equation 2.3.5), and afterwards by formally computing the ratio  $y(t)/u(t)$  and replacing  $q^{-1}$  with  $z^{-1}$ . This procedure can obviously be applied to all models described by linear difference equations with constant coefficients, regardless of their complexity. The same result can be also derived by means of the  $z$ -transform (see Appendix A, Section A.2)

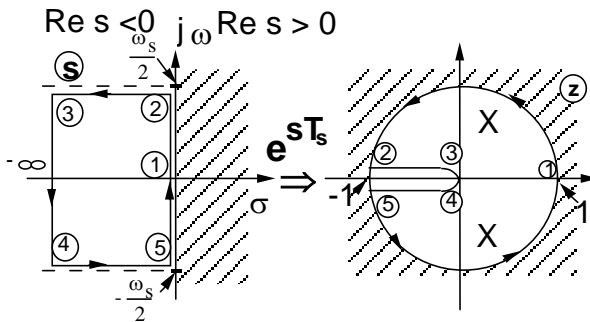
We also remark that the transfer functions of discrete-time models are often written in terms of  $q^{-1}$ . It is of course understood that the meaning of  $q^{-1}$  varies according to the context (delay operator or complex variable). When  $q^{-1}$  is considered as a delay operator, the expression  $H(q^{-1})$  is named “transfer operator”. It must be observed that the representation by transfer operators can also be used for models described by linear difference equations with time varying coefficients



as well. In contrast, the interpretation of  $q^{-1}$  as a complex variable ( $z^{-1}$ ) is only possible for linear difference equations with constant coefficients.

*Properties of the Transformation  $z = e^{sT_s}$*

The transformation of Equation 2.3.18 is not bijective because several points in the  $s$ - plane are transformed at the same point in the  $z$ -plane. Nevertheless, we are interested in the  $s$ -plane being delimited between the two horizontal lines crossing the points  $[0, +j\omega_s / 2]$  and  $[0, -j\omega_s / 2]$  where  $\omega_s = 2\pi f_s = 2\pi / T_s$ . This region is called “primary strip”.



**Figure 2.15.** Effects of the transformation  $z = e^{sT_s}$  on the points located in the “primary strip” in  $s$ -plane

The complementary bands are outside the frequency domain of interest if the conditions of the Shannon theorem (Section 2.2.1) have been satisfied.

Figure 2.15 gives a detailed image of the effects of the transformation  $z = e^{sT_s}$  for the points that are inside the “primary strip”.

Attention must be focused on an important aspect for continuous second-order systems in the form:

$$\frac{\omega_0^2}{\omega_0^2 + 2\zeta\omega_0s + s^2} \quad (\zeta < 1)$$

for which the resonant damped frequency is equal to half the sampling frequency:

$$\omega_0\sqrt{1 - \zeta^2} = \omega_s / 2$$

*The image of their conjugates poles*

$$s_{1,2} = -\zeta\omega_0 \pm j\frac{\omega_s}{2}$$

through the transformation  $z = e^{sT_s}$  corresponds to a single point placed on the real axis in the  $z$ -plane and with negative abscissa.

One gets:

$$z_{1,2} = e^{s_{1,2}T_s} = e^{-\zeta\omega_0T_s} \cdot e^{\pm j\frac{\omega_s}{2}T_s} = e^{-\zeta\omega_0T_s} \cdot e^{\pm j\pi} = -e^{-\zeta\omega_0T_s} = -e^{-\frac{\zeta}{\sqrt{1-\zeta^2}}\pi}$$

since:

$$\omega_0 = \frac{\omega_s}{2\sqrt{1-\zeta^2}}$$

This is the reason why discrete-time models in the form of Equation 2.3.3 such as

$$(I + a_1 q^{-1}) y(t) = b_1 q^{-1} u(t)$$

give oscillating step responses for  $a_1 > 0$  (damped if  $|a_1| < 1$ ) with period  $2T_s$  (see Section 2.3.1). These first-order discrete-time models have the same poles as the discrete-time models derived from second-order continuous-time systems having a damped resonant frequency equal to  $\omega_s/2$ .

### 2.3.3 General Forms of Linear Discrete-time Models

A linear discrete-time model is generally described as

$$y(t) = -\sum_{i=1}^{n_A} a_i y(t-i) + \sum_{i=1}^{n_B} b_i u(t-d-i) \quad (2.3.20)$$

in which  $d$  corresponds to a pure time delay which is an integer multiple of the sampling period.

Let us introduce the following notations:

$$1 + \sum_{i=1}^{n_A} a_i q^{-i} = A(q^{-1}) = 1 + q^{-1} A^*(q^{-1}) \quad (2.3.21)$$

$$A^*(q^{-1}) = a_1 + a_2 q^{-1} + \dots + a_{n_A} q^{-n_A+1} \quad (2.3.22)$$

$$\sum_{i=1}^{n_B} b_i q^{-i} = B(q^{-1}) = q^{-1} B^*(q^{-1}) \quad (2.3.23)$$

$$B^*(q^{-1}) = b_1 + b_2 q^{-1} + \dots + b_{n_B} q^{-n_B+1} \quad (2.3.24)$$

By using the delay operator  $q^{-1}$  in Equation 2.3.20 and taking into account the notations of Equations 2.3.21 to 2.3.24, the Equation 2.3.20 describing the discrete-time system is written as

$$A(q^{-1}) y(t) = q^{-d} B(q^{-1}) u(t) \quad (2.3.25)$$

or in the predictive form (by multiplying both sides by  $q^d$ )

$$A(q^{-1}) y(t+d) = B(q^{-1}) u(t) \quad (2.3.26)$$

Equation 2.3.25 can also be written in a compact form using the *pulse transfer operator*

$$y(t) = H(q^{-1}) u(t) \quad (2.3.27)$$

where the pulse transfer operator is given by

$$H(q^{-1}) = \frac{q^{-d} B(q^{-1})}{A(q^{-1})} \quad (2.3.28)$$

The *pulse transfer function* characterizing the system described by Equation 2.3.20 is obtained from the pulse transfer operator given in Equation 2.3.28 by replacing  $q^{-1}$  with  $z^{-1}$ <sup>2</sup>

$$H(z^{-1}) = \frac{z^{-d} B(z^{-1})}{A(z^{-1})} \quad (2.3.29)$$

### *Pulse Transfer Function Order*

To evaluate the order of a discrete time model represented by the pulse transfer function in the form of Equation 2.3.29, the representation in terms of positive power of  $z$  is needed. If  $d$  is the system pure time delay expressed as number of samples,  $n_A$  the degree of the polynomial  $A(z^{-1})$  and  $n_B$  the degree of the polynomial  $B(z^{-1})$ , one must multiply both numerator and denominator of  $H(z^{-1})$  by  $z^n$  in order to obtain a *proper*<sup>3</sup> pulse transfer function  $H(z)$  on the positive powers of  $z$ , where

<sup>2</sup> The pulse transfer operator  $H(q^{-1})$  can be used for a compact representation of the input-output relationship even in the case of  $A(q^{-1})$  and  $B(q^{-1})$  have time depending coefficients. The pulse transfer function  $H(z^{-1})$  is only defined for the case of  $A(q^{-1})$  and  $B(q^{-1})$  are with constant coefficients.

<sup>3</sup> This means that the denominator degree is greater than (or equal to) the numerator degree.

$$n = \max(n_A, n_B + d)$$

$n$  represents the *discrete-time system order* (the higher power of a term in  $z$  in the pulse transfer function denominator).

*Example 1:*

$$H(z^{-1}) = \frac{z^{-3}(b_1 z^{-1} + b_2 z^{-2})}{1 + a_1 z^{-1}}$$

$$n = \max(1, 5) = 5$$

$$H(z) = \frac{b_1 z + b_2}{z^5 + a_1 z^4}$$

*Example 2:*

$$H(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

$$n = \max(2, 2) = 2$$

$$H(z) = \frac{b_1 z + b_2}{z^2 + a_1 z + a_2}$$

One notes that the order  $n$  of an irreducible pulse transfer function also corresponds to the number of states for a minimal state space system representation associated to the transfer function (See Appendix C).

### 2.3.4 Stability of Discrete-time Systems

The stability of discrete-time systems can be studied either from the recursive (differences) equation describing the discrete-time system in the time domain, or from the interpretation of difference equations solutions as sums of discretized exponentials. We shall use examples to illustrate both these approaches.

Let us assume that the recursive equation is

$$y(t) = -a_1 y(t-1) ; \quad y(0) = y_0 \quad (2.3.30)$$

which is obtained from Equation 2.3.3 when the input  $u(t)$  is identically zero. The free response of the system is written as

$$y(1) = -a_1 y_0 ; \quad y(2) = (-a_1)^2 y_0 ; \quad y(t) = (-a_1)^t y_0 \quad (2.3.31)$$

The asymptotic stability of the system implies

$$\lim_{t \rightarrow \infty} y(t) = 0 \quad (2.3.32)$$

The condition of asymptotic stability thus results from Equation 2.3.31. It is necessary and sufficient that

$$|a_1| < 1 \quad (2.3.33)$$

On the other hand, it is known that the solution of the recursive (difference) equations is of the form (for a first-order system):

$$y(t) = K e^{sT_s t} = K z^t \quad (2.3.34)$$

By introducing this solution into Equation 2.3.30, and taking into account Equation 2.3.15, one obtains

$$(1 + a_1 e^{-sT_s}) K e^{sT_s t} = (1 + a_1 z^{-1}) K z^t = 0 \quad (2.3.35)$$

from which it follows that

$$z = e^{sT_s} = e^{(\sigma + j\omega)T_s} = e^{\sigma T_s} e^{j\omega T_s} = -a_1 \quad (2.3.36)$$

For this solution to be asymptotically stable, it is necessary that  $\sigma = \operatorname{Re} s < 0$  which implies that  $e^{\sigma T_s} < 1$  and respectively  $|z| < 1$  (or  $|a_1| < 1$ ).

However, the term  $(1 + a_1 z^{-1})$  is nothing more than the denominator of the pulse transfer function related to the system described by Equation 2.3.3 (see Equation 2.3.19).

The result obtained can be generalized. For a discrete-time system to be *asymptotically stable*, all the roots of the transfer function denominator must be inside the unit circle (see Figure 2.14):

$$1 + a_1 z^{-1} + \dots + a_n z^{-n} = 0 \Rightarrow |z| < 1 \quad (2.3.37)$$

In contrast, if one or several roots of the transfer function denominator are in the region defined by  $|z| > 1$  (outside the unit circle), this implies that  $\operatorname{Re} s > 0$  and thus the discrete-time system will be unstable.

As for the continuous-time case, some stability criteria are available (Jury criterion, Routh-Hurwitz criterion applied after the change of variable  $w = (z + 1)/(z - 1)$ ) for establishing the existence of unstable roots for a polynomial in the variable  $z$  with no explicit calculation of the roots (Åström and Wittenmark 1997).

A helpful tool to test  $z$ -polynomial stability is derived from a *necessary condition* for the stability of a  $z^{-1}$ -polynomial. This condition states: the evaluations of the polynomial  $A(z^{-1})$  given by Equation 2.3.37 in  $z = 1$ ,  $A(1)$  and in  $z = -1$  ( $A(-1)$ ) must be positive (the coefficient of  $A(z^{-1})$  corresponding to  $z^0$  is supposed to be positive).

*Example:*

$$A(z^{-1}) = 1 - 0.5 z^{-1} \quad (\text{stable system})$$

$$A(1) = 1 - 0.5 = 0.5 > 0 \quad ; \quad A(-1) = 1 + 0.5 = 1.5 > 0$$

$$A(z^{-1}) = 1 - 1.5 z^{-1}; \quad (\text{unstable system})$$

$$A(1) = -0.5 < 0 \quad ; \quad A(-1) = 2.5 > 0$$

### 2.3.5 Steady-state Gain

In the case of continuous-time systems, the steady-state gain is obtained by making  $s = 0$  (zero frequency) in the transfer function. In the discrete case,  $s = 0$  corresponds to

$$s = 0 \Rightarrow z = e^{sT} = 1 \quad (2.3.38)$$

and thus the steady-state gain  $G(0)$  is obtained by making  $z = 1$  in the pulse transfer function. Therefore for the first-order system one obtains:

$$G(0) = \left. \frac{b_1 z^{-1}}{1 + a_1 z^{-1}} \right|_{z=1} = \frac{b_1}{1 + a_1}$$

Generally speaking, the steady-state gain is given by the formula

$$G(0) = H(1) = H(z^{-1}) \Big|_{z=1} = \left. \frac{z^{-d} B(z^{-1})}{A(z^{-1})} \right|_{z=1} = \frac{\sum_{i=1}^{n_B} b_i}{1 + \sum_{i=1}^{n_A} a_i} \quad (2.3.39)$$

In other words, the steady-state gain is obtained as the ratio between the sum of the numerator coefficients and the sum of the denominator coefficients. This formula is quite different from the continuous-time systems, where the steady-state gain appears as a common factor of the numerator (if the denominator begins with 1).

The steady-state gain may also be obtained from the recursive equation describing the discrete-time models, the steady-state being characterized by  $u(t) = \text{const.}$  and  $y(t) = y(t-1) = y(t-2)\dots$

From Equation 2.3.3, it follows that

$$(1 + a_1) y(t) = b_1 u(t)$$

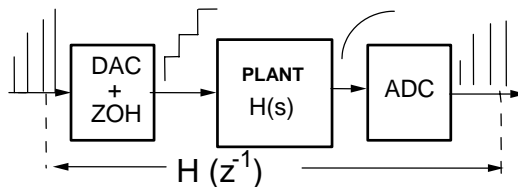
and respectively

$$y(t) = \frac{b_1}{1 + a_1} u(t) = G(0)u(t)$$

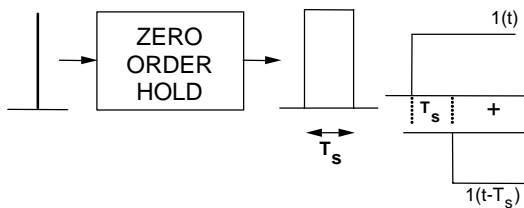
### 2.3.6 Models for Sampled-data Systems with Hold

Up to this point we have been concerned with sampled-data systems models corresponding to the discretization of inputs and outputs of a continuous-time system. However, in a computer controlled system, the control applied to the plant is not continuous. It is constant between the sampling instants (effect of the zero-order hold) and varies discontinuously at the sampling instants, as is illustrated in Figure 2.16.

It is important to be able to relate the model of the discretized system, which gives the relation between the control sequence (produced by the digital controller) and the output sequence (obtained after the analog-to-digital converter), to the transfer function  $H(s)$  of the continuous-time system. The zero-order hold, whose operation is reviewed in Figure 2.17 introduces a transfer function in cascade with  $H(s)$ .



**Figure 2.16.** Control system using an analog-to-digital converter followed by a zero-order hold



**Figure 2.17.** Operation of the zero-order hold

The hold converts a Dirac pulse given by the digital-to-analog converter at the sampling instant into a rectangular pulse of duration  $T_s$ , which can be interpreted as the difference between a step and the same step shifted by  $T_s$ . As the step is the integral of the Dirac pulse, it follows that the zero-order hold transfer function is

$$H_{ZOH}(s) = \frac{1 - e^{-sT_s}}{s} \quad (2.3.40)$$

Equation 2.3.40 allows one to consider the zero-order hold as a filter having a frequency response given by

$$H_{ZOH}(j\omega) = \frac{1 - e^{-j\omega T_s}}{j\omega} = T_s \frac{\sin(\omega T_s / 2)}{\omega T_s / 2} e^{-j\omega \frac{T_s}{2}}$$

From the study of this response in the frequency region  $0 \leq f \leq f_s / 2$  ( $0 \leq \omega \leq \omega_s / 2$ ), one can conclude:

1. The ZOH gain at the zero frequency is equal to:  $G_{ZOH}(0) = T_s$ .
2. The ZOH introduces an attenuation at high frequencies. For  $f = f_s / 2$  one

$$\text{gets } G(f_s / 2) = \frac{2}{\pi} T_s = 0.637 T_s \text{ (-3.92 dB)}.$$

3. The ZOH introduces a phase lag which grows with the frequency. This phase lag is between 0 (for  $f = 0$ ) and  $-\pi/2$  (for  $f = f_s / 2$ ) and should be added to the phase lag due to  $H(s)$ .

The global continuous-time transfer function will be

$$H'(s) = \frac{1 - e^{-sT_s}}{s} H(s) \quad (2.3.41)$$

to which a pulse transfer function is associated.

Tables which give the discrete-time equivalent of systems with a zero-order hold are available. Some typical situations are summarized in Table 2.4.

The computation of ZOH sampled models for transfer functions of different orders can be done by means of the functions: *cont2disc.sci* (Scilab) or *cont2disc.m* (MATLAB<sup>®</sup>). The corresponding sampled model (with Z.O.H) for a second-order system characterized by  $\omega_0$  and  $\zeta$  can be obtained with the functions *ft2pol.sci* (Scilab) or *ft2pol.m* (MATLAB<sup>®</sup>)<sup>4</sup>.

---

<sup>4</sup> To be downloaded from the book website.



### 2.3.7 Analysis of First-order Systems with Time Delay

The continuous-time model is characterized by the transfer function

$$H(s) = \frac{G e^{-s\tau}}{1 + T_s s} \quad (2.3.42)$$

where  $G$  is the gain,  $T$  is the time constant and  $\tau$  is the pure time delay. If  $T_s$  is the sampling period, then  $\tau$  is expressed as

$$\tau = d T_s + L ; \quad 0 < L < T_s \quad (2.3.43)$$

where  $L$  is the fractional time delay and  $d$  is the integer number of sampling periods included in the delay and corresponding to a sampled delay of  $d$ -periods. From Table 2.4, one derives the transfer function of the corresponding sampled model (when a zero-order hold is used)

$$H(z^{-1}) = \frac{z^{-d} (b_1 z^{-1} + b_2 z^{-2})}{1 + a_1 z^{-1}} = \frac{z^{-d-1} (b_1 + b_2 z^{-1})}{1 + a_1 z^{-1}} \quad (2.3.44)$$

with

$$a_1 = -e^{-\frac{T_s}{T}} ; \quad b_1 = G(1 - e^{-\frac{L-T_s}{T}}) ; \quad b_2 = Ge^{-\frac{T_s}{T}} (e^{\frac{L}{T}} - 1)$$

The effect of the fractional time delay can be seen in the appearance of the coefficient  $b_2$  in the transfer function. For  $L = 0$ , one gets  $b_2 = 0$ . On the other hand, if  $L = T_s$ , it follows that  $b_1 = 0$ , which correspond to an additional delay of one sampling period. For  $L < 0.5T_s$  one has  $b_2 < b_1$ , and for  $L > 0.5T_s$  one has  $b_2 > b_1$ . For  $L = 0.5T_s$ ,  $b_2 \approx b_1$ . Therefore, a fractional delay introduces a zero in the pulse transfer function. For  $L > 0.5 T_s$  the relation  $|b_2| > |b_1|$  holds and the zero is outside the unit circle (unstable zero)<sup>5</sup>.

The pole-zero configuration in the  $z$  plane for the first-order system with ZOH is represented in Figure 2.18. The term  $z^{-d-1}$  introduces  $d+1$  poles at the origin  $[H(z) = (b_1 z + b_2) / z^{d+1} (z + a_1)]$ .

---

<sup>5</sup> The presence of unstable zeros has no influence on the system stability, but it imposes constraints on the use of controller design techniques based on the cancellation of model zeros by controller poles.

**Table 2.4.** Pulse transfer functions for continuous-time systems with zero-order hold

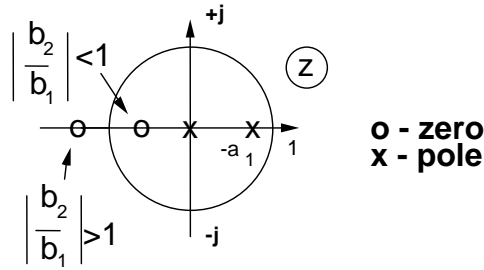
$H(s)$	$H(z^{-1})$
$\frac{1}{s}$	$\frac{T_s z^{-1}}{1 - z^{-1}}$
$\frac{G}{1 + sT}$	$\frac{b_1 z^{-1}}{1 + a_1 z^{-1}} ; b_1 = G(1 - e^{-T_s/T}) ; a_1 = -e^{-T_s/T}$
$\frac{Ge^{-sL}}{1 + sT} ; L < T_s$	$\frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1}} ; b_1 = G(1 - e^{(L-T_s)/T}) ;$ $b_2 = Ge^{-T_s/T} (e^{L/T} - 1) ; a_1 = -e^{-T_s/T}$
$\frac{\omega_0^2}{\omega_0^2 + 2\zeta\omega_0 s + s^2}$  $\omega = \omega_0 \sqrt{1 - \zeta^2}$ $\zeta < 1$	$\frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$  $b_1 = 1 - \alpha \left( \beta + \frac{\zeta\omega_0}{\omega} \delta \right) ; b_2 = \alpha^2 + \alpha \left( \frac{\zeta\omega_0}{\omega} \delta - \beta \right)$  $a_1 = -2\alpha\beta ; a_2 = \alpha^2$  $\alpha = e^{-\zeta\omega_0 T_s} ; \beta = \cos(\omega T_s) ; \delta = \sin(\omega T_s)$

Figure 2.19 represents the step responses for a system characterized by a pulse transfer function

$$H(z^{-1}) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1}} \quad (2.3.45)$$

with  $\frac{b_1}{1 + a_1} = 1$  (steady-state gain = 1) for different values of the parameter  $a_1$ :

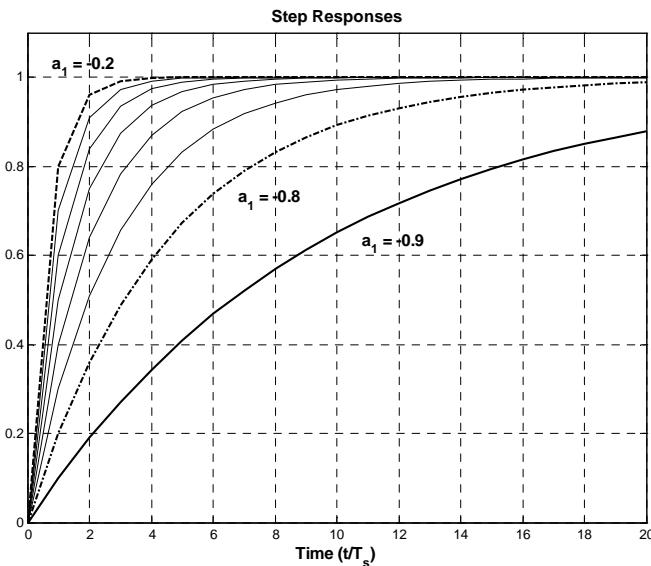
$$a_1 = -0.2 ; -0.3 ; -0.4 ; -0.5 ; -0.6 ; -0.7 ; -0.8 ; -0.9$$



**Figure 2.18.** Pole-zero configuration of the sampled-data system described by Equation 2.3.44 ( first order system with ZOH)

On the basis of these responses, it is easy to derive the time constant of the corresponding continuous-time system, expressed in terms of the sampling period (the time constant is equal to the time required to reach 63% of the final value).

The presence of a time delay equal to an integer multiple of the sampling period only causes a time shift in the responses given in Figure 2.19.



**Figure 2.19.** Step responses of the discrete-time system  $b_1 z^{-1}/(1+a_1 z^{-1})$  for different values of  $a_1$  and  $[b_1/(1+ a_1)] = 1$

The presence of a fractional time delay has as a main consequence a modification at the beginning of the step response, if compared to the case with no fractional time delay.

*Exercise.* Assuming that the sampled-data system model is

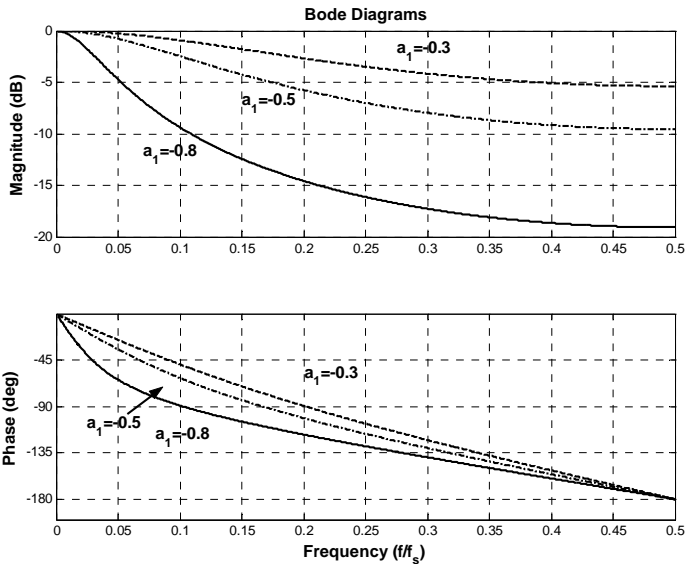
$$y(t) = -0.6 y(t-1) + 0.2 u(t-1) + 0.2 u(t-2)$$

What is the corresponding continuous-time model?

It is interesting to analyze the relation between the location of the pole ( $z = -a_1$ ) and the rising time of the system. Figure 2.19 indicates that the response of the system becomes slower as the pole of the system moves toward the point  $[1, j0]$ , and it becomes faster as the pole of the system approaches the origin ( $z = 0$ ). These considerations can be applied to systems with several poles.

In the case of systems with more than one pole, the term “dominant pole(s)” is introduced to characterize the pole (or the poles) that is (are) the closest to the point  $[1, j0]$ , i.e. which is the slowest pole(s).

Figure 2.20 shows the frequency responses (magnitude and phase) of the first-order discrete system given by 2.3.45 for  $a_1 = -0.8; -0.5; -0.3$ . It can be observed that the bandwidth increases when the system pole is approaching the origin (faster pole). We can also remark that the phase lag at the frequency  $0.5f_s$  is  $-180^\circ$  due to the presence of the ZOH (see Section 2.3.6).



**Figure 2.20.** Frequency responses (magnitude and phase) of the discrete-time model  $b_1 z^{-1} / (1 + a_1 z^{-1})$  for different values of  $a_1$  and  $b_1$

### 2.3.8 Analysis of Second-order Systems

The pulse transfer function corresponding to the discretization with a zero-order hold of a normalized second-order continuous-time system, characterized by a natural frequency  $\omega_0$  and a damping  $\zeta$ , is given by

$$H(z^{-1}) = \frac{z^{-d}(b_1 z^{-1} + b_2 z^{-2})}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad (2.3.46)$$

where  $d$  represents the integer number of sampling periods contained in the delay. The values of  $a_1$ ,  $a_2$ ,  $b_1$ ,  $b_2$  as a function of  $\omega_0$  and  $\zeta$  for a pure time delay  $\tau = dT_s$  are given in Table 2.4.

It is interesting to express the poles of the discretized system as a function of  $\omega_0$ ,  $\zeta$  and the sampling period  $T_s$  (or the sampling frequency  $f_s$ ).

From Table 2.4 the following relations are easily found (for  $\zeta < 1$ ):

$$\begin{aligned} a_1 &= -2e^{-\zeta\omega_0 T_s} \cos \omega_0 \sqrt{1-\zeta^2} T_s = -e^{-\zeta\omega_0 T_s} (e^{j\omega_0 \sqrt{1-\zeta^2} T_s} + e^{-j\omega_0 \sqrt{1-\zeta^2} T_s}) \\ a_2 &= -e^{-2\zeta\omega_0 T_s} \end{aligned}$$

The poles of the pulse transfer function (roots of the denominator) are found by solving the equation

$$z^2 + a_1 z + a_2 = (z - z_1)(z - z_2) = 0$$

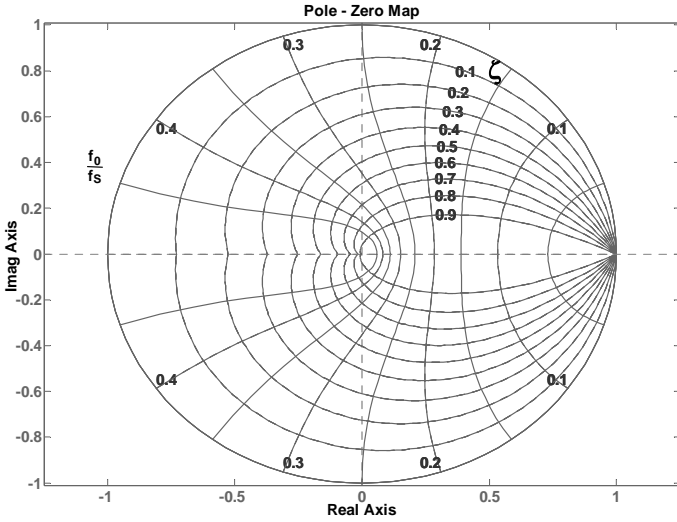
From the expressions of  $a_1$  and  $a_2$  the solutions are directly derived:

$$z_{1,2} = e^{-\zeta\omega_0 T_s \pm j\omega_0 \sqrt{1-\zeta^2} T_s}$$

Note that the poles of the discretized system correspond to the poles of the continuous-time system  $s_{1,2} = \zeta\omega_0 \pm j\omega_0 \sqrt{1-\zeta^2}$  by applying the transformation  $z = e^{sT_s}$ .

For  $\zeta < 1$ , the poles of the discretized system are complex conjugate and, consequently, symmetric with respect to the real axis. They are characterized by a module and a phase given by

$$\begin{aligned} |z_{1,2}| &= e^{-\zeta\omega_0 T_s} = e^{-2\pi\zeta \frac{f_0}{f_s}} = e^{-2\pi\zeta \frac{\omega_0}{\omega_s}} \\ \angle z_{1,2} &= \pm \sqrt{1-\zeta^2} \omega_0 T_s = \pm 2\pi \sqrt{1-\zeta^2} \frac{f_0}{f_s} = \pm 2\pi \sqrt{1-\zeta^2} \frac{\omega_0}{\omega_s} \end{aligned}$$



**Figure 2.21.** The curves  $\zeta = \text{constant}$  and  $\omega_0 T_s / 2\pi = f_0 / f_s = \text{constant}$  in the  $z$ -plane for a second-order discrete-time system

Note that the poles location depends upon  $\zeta$  and  $\omega_0 T_s$  (or  $\omega_0 / \omega_s = f_0 / f_s$ ).

That is:

$$z = f(\zeta, \omega_0 T_s / 2\pi) = f(\zeta, f_0 / f_s)$$

and in the  $z$ -plane the following curves can be drawn:

$$z = f(\omega_0 T_s / 2\pi) = f(f_0 / f_s) \qquad \text{for } \zeta = \text{constant}$$

and

$$z = f(\zeta) \qquad \text{for } \omega_0 T_s / 2\pi = f(f_0 / f_s) = \text{constant}$$

We must remember (see Figure 1.9) that in the  $s$ -plane (continuous system) the curves  $\zeta = \text{constant}$  are straight lines forming an angle  $\theta = \cos^{-1} \zeta$  with the real axis and the curves  $\omega_0 = \text{constant}$  are circles with radius  $\omega_0$  (these two sets of curves are orthogonal). In the  $z$ -plane the curves  $z = f(\omega_0 T_s)$  for  $\zeta = \text{constant}$  are logarithmic spirals that are orthogonal in each point to the curves  $z: f(\zeta)$  for  $\omega_0 T_s = \text{constant}$ .

Figure 2.21 shows the set of curves  $z = f(\zeta)$  for  $\omega_0 T_s/2\pi = \text{constant}$  and  $z = f(\omega_0 T_s/2\pi)$  for  $\zeta = \text{constant}$  corresponding to different values of  $\zeta$  and  $\omega_0 T_s/2\pi$  (respectively  $f_0/f_s$ ).

We should also remember (see Section 2.3.2) that for  $f_0/f_s = \omega_0/\omega_s = 1/(2\sqrt{1-\zeta^2})$ , the corresponding poles in the  $z$ - plane are confounded ( $\angle z_{1,2} = \pm\pi$ ), and they are located on the segment of the real axis  $(-1, 0)$  having an abscissa coordinate equal to  $-e^{-\frac{\pi\zeta}{\sqrt{1-\zeta^2}}}$ .

The stability domain of the second-order discrete-time system in the plane of the parameters  $a_1 - a_2$  is a triangle (see Figure 2.22). For values of  $a_1, a_2$  placed inside of the triangle, the roots of the denominator of the pulse transfer function are inside the unit circle.

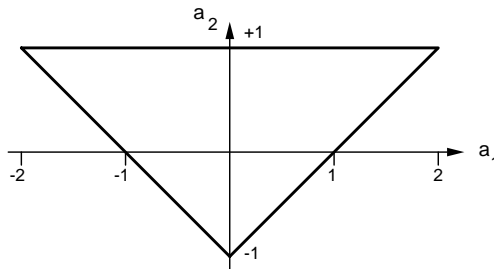


Figure 2.22. Stability domain for the second-order discrete-time system

## 2.4 Closed Loop Discrete-time Systems

### 2.4.1 Closed Loop System Transfer Function

Figure 2.23 gives the diagram of a closed loop discrete-time system. The transfer function on the feedforward channel can result from the cascade of a digital controller and of the group DAC+ ZOH + continuous-time system + ADC (discretized system).

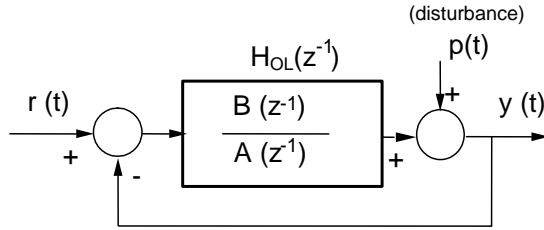


Figure 2.23. Closed loop discrete-time system

Let

$$H_{OL}(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} \tag{2.4.1}$$

be the feed-forward channel transfer function with

$$B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_B} z^{-n_B} \tag{2.4.2}$$

where the coefficients  $b_1, b_2 \dots b_d$  may be zero if there is a time delay of  $d$  sampling periods.

In the same way as for continuous-time systems, the closed loop transfer function connecting the reference signal  $r(t)$  to the output  $y(t)$  is written as

$$H_{CL}(z^{-1}) = \frac{H_{OL}(z^{-1})}{1 + H_{OL}(z^{-1})} = \frac{B(z^{-1})}{A(z^{-1}) + B(z^{-1})} \tag{2.4.3}$$

The denominator of the closed loop transfer function, whose roots correspond to the closed loop system poles, is also called *characteristic polynomial of the closed loop*.

### 2.4.2 Steady-state Error

The steady-state is obtained for  $r(t) = \text{constant}$  by making  $z = 1$ , corresponding to the zero frequency ( $z = e^{sTs} = 1$  for  $s = 0$ ).

It follows from Equation 2.4.3 that

$$y = H_{CL}(1)r = \frac{\sum_{i=1}^{n_B} b_i}{A(1) + \sum_{i=1}^{n_B} b_i} r \tag{2.4.4}$$



where  $H_{CL}(1)$  is the steady-state gain (static gain) of the closed loop system. In order to obtain a zero steady-state error between the reference signal  $r$  and the output  $y$ , it is necessary that

$$H_{CL}(1) = 1 \quad (2.4.5)$$

From Equation 2.4.4 the following conditions are derived:

$$\sum_{i=1}^{n_B} b_i \neq 0 \text{ and } A(1) = 0 \quad (2.4.6)$$

In order to obtain  $A(1) = 0$ ,  $A(z^{-1})$  must have the following structure:

$$A(z^{-1}) = (1 - z^{-1}) \cdot A'(z^{-1}) \quad (2.4.7)$$

where

$$A'(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_{n_A} z^{-n_A} \quad (2.4.8)$$

and thus the global transfer function of the feedforward channel must be of the type

$$H_{OL}(z^{-1}) = \frac{1}{1 - z^{-1}} \cdot \frac{B(z^{-1})}{A'(z^{-1})} \quad (2.4.9)$$

It is thus observed that *the feedforward channel must contain a digital integrator* in order to obtain a *zero steady-state error* in closed loop. This situation is similar to the continuous case (see Section 1.2.3) and *internal model principle* is also applicable to discrete-time systems.

### 2.4.3 Rejection of Disturbances

In the presence of a disturbance  $p(t)$  acting on the controlled output (see Figure 2.23), the objective is to reduce its effect as much as possible, at least in some frequency regions.

In particular, the constant disturbance effect (a step), often called “load disturbance”, is expected to be zero in steady-state ( $t \rightarrow \infty$ ,  $z \rightarrow 1$ ).

The pulse transfer function, which links the disturbance to the output, is

$$S_{yp}(z^{-1}) = \frac{1}{1 + H_{OL}(z^{-1})} = \frac{A(z^{-1})}{A(z^{-1}) + B(z^{-1})}$$

As for the continuous-time case,  $S_{yp}(z^{-1})$  is called “output sensitivity function”. The steady-state is obtained for  $z = 1$ . It follows that

$$y = S_{yp}(1) p = \frac{A(1)}{A(1) + B(1)} p$$

(where  $p$  is the stationary value of the disturbance).

In order to achieve a perfect steady-state disturbance rejection, it is necessary that  $S_{yp}(1) = 0$  and thus  $A(1) = 0$ . This implies that  $A(z^{-1})$  must have the form given in Equation 2.4.7, corresponding to the *integrator* insertion in the feedforward channel.

Similarly, to the continuous case, a perfect steady-state disturbance rejection implies that the feedforward channel must contain *the internal model* of the disturbance (the transfer function that produces  $p(t)$  from a Dirac pulse).

As in the continuous-time case, it should be avoided that an amplification of the disturbance effect occurs in certain frequency regions. This is the reason why  $|S_{yp}(e^{-j\omega})|$  must be lower than a specified value for all frequencies  $f = \omega/2\pi \leq f_s/2$ .

A typical value used as upper bound is

$$|S_{yp}(e^{-j\omega})| \leq 2 \quad 0 \leq \omega \leq \pi f_s$$

Furthermore, if it is known that a disturbance has its energy concentrated in a particular frequency region,  $|S_{yp}(e^{j\omega})|$  may be constrained to introduce a desired attenuation in this frequency region.

## 2.5 Basic Principles of Modern Methods for Design of Digital Controllers

### 2.5.1 Structure of Digital Controllers

Figure 2.24 gives the diagram of a PI type analog controller. The controller contains two channels (a proportional channel and an integral channel) that process the error between the reference signal and the output.

In the case of sampled-data systems the controller is digital, and the only operations it can carry out are additions, multiplications, storage and shift. All the digital control algorithms have the same structure. Only “the memory” of the controller is different, that is the number of coefficients.

Figure 2.25 illustrates the computation structure of the control  $u(t)$  applied to the plant at the instant  $t$  by the digital controller. This control is a weighted average of the measured output at instants  $t, t-1, \dots, t-n_A \dots$ , of the previous control values at instants  $t-1, t-1, \dots, t-n_B \dots$  and of the reference signal at instants  $t, t-1, \dots$ , the weights being the coefficients of the controller.



$$\int xdt = \frac{1}{p} x \approx \left[ \frac{T_s}{1-q^{-1}} \right] x(t) \tag{2.5.3}$$

and the equation of the PI controller becomes

$$u(t) = \frac{K(1-q^{-1}) + \frac{KT_s}{T_i}}{1-q^{-1}} [r(t) - y(t)] \tag{2.5.4}$$

Multiplying both sides of Equation 2.5.4 by  $(1 - q^{-1})$ , the equation of the digital PI controller is written as

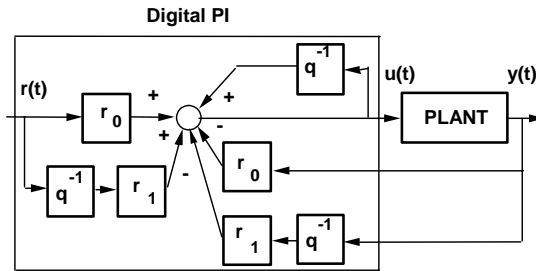
$$S(q^{-1}) u(t) = T(q^{-1}) r(t) - R(q^{-1}) y(t) \tag{2.5.5}$$

where

$$S(q^{-1}) = 1 - q^{-1} = 1 + s_1 q^{-1} \quad (s_1 = -1) \tag{2.5.6}$$

$$R(q^{-1}) = T(q^{-1}) = K (1 + T_s/T_i) - K q^{-1} = r_0 + r_1 q^{-1} \tag{2.5.7}$$

which leads to the diagram represented in Figure 2.26.



**Figure 2.26.** Digital PI controller

Taking into account the expression of  $S(q^{-1})$ , the control signal  $u(t)$  is computed on the basis of Equation 2.5.5, by means of the formula

$$\begin{aligned} u(t) &= u(t-1) - R(q^{-1}) y(t) + T(q^{-1}) r(t) \\ &= u(t-1) - r_0 y(t) - r_1 y(t-1) + r_0 r(t) + r_1 r(t-1) \end{aligned} \tag{2.5.8}$$

which corresponds to the diagram given in Figure 2.26.

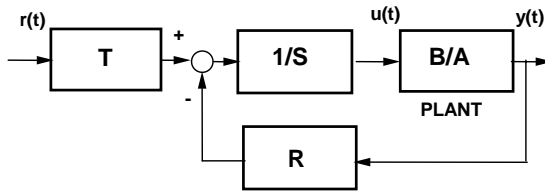
## 2.5.2 Digital Controller Canonical Structure

Dividing by  $S(q^{-1})$  both sides of Equation 2.5.5, one obtains

$$u(t) = -\frac{R(q^{-1})}{S(q^{-1})}y(t) + \frac{T(q^{-1})}{S(q^{-1})}r(t) \quad (2.5.9)$$

from which we derive the digital controller canonical structure presented in Figure 2.27 (three branched RST structure).

In general,  $T(q^{-1})$  in Figure 2.27 is different from  $R(q^{-1})$ .



**Figure 2.27.** Digital controller canonical structure

Consider

$$H(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} \quad (2.5.10)$$

as the pulse transfer function of the cascade DAC + ZOH + continuous-time system + ADC, then the transfer function of the open loop system is written as

$$H_{OL}(z^{-1}) = \frac{B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})} \quad (2.5.11)$$

and the closed loop transfer function between the reference signal  $r(t)$  and the output  $y(t)$ , using a digital controller canonical structure, has the expression

$$H_{CL}(z^{-1}) = \frac{B(z^{-1})T(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} = \frac{B(z^{-1})T(z^{-1})}{P(z^{-1})} \quad (2.5.12)$$

where

$$P(z^{-1}) = A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1}) = 1 + p_1z^{-1} + p_2z^{-2} + \dots \quad (2.5.13)$$

is the denominator of the closed loop transfer function that defines the closed loop system poles. Note that  $T(q^{-1})$  introduces one more degree of freedom, which

allows one to establish a distinction between tracking and regulation performances specifications.

We also remark that  $r(t)$  is often replaced by a “desired trajectory”  $y^*(t)$ , obtained either by filtering the reference signal  $r(t)$  with the so-called shaping filter or tracking reference model, or saving in the memory of the digital computer the sequence of the desired trajectory values.

The digital controller represented in Figure 2.27 is also defined as “RST digital controller”. It is a *two degrees of freedom* controller, which allows one to impose different specifications in terms of desired dynamics for the tracking and regulation problems.

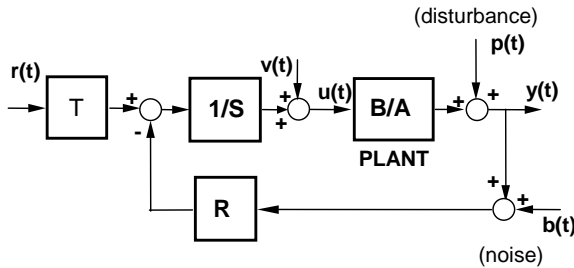
The goal of the digital controller design is to find the polynomials  $R$ ,  $S$ , and  $T$  in order to obtain the closed loop transfer functions, with respect to the reference and disturbance signals, satisfying the desired performances.

This explains why the desired closed loop performances will be expressed, (if not, they will be converted) in terms of desired closed loop poles, and eventually in terms of desired zeros (in this way the closed loop transfer function will be completely imposed).

In the presence of disturbances (see Figure 2.28) there are other four important transfer functions to consider, relating the disturbance to the output and the input of the plant.

The transfer function between the disturbance  $p(t)$  and the output  $y(t)$  (*output sensitivity function*) is given by

$$S_{yp}(z^{-1}) = \frac{A(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} \tag{2.5.14}$$



**Figure 2.28.** Digital control system in presence of disturbances and noise

This function allows the characterization of the system performances from the point of view of disturbances rejection. In addition, certain components of  $S(z^{-1})$  can be pre-specified in order to obtain satisfactory disturbance rejection properties.

Thus, if a perfect disturbance rejection is required at a specified frequency,  $S(z^{-1})$  must include a zero corresponding to this frequency. In particular, if a perfect load disturbance rejection in steady-state (i.e. zero frequency) is desired,  $S_{yp}(z^{-1})$  must include a term  $(I - z^{-1})$  in the numerator, which leads to a value of the gain

equal to zero for  $z = 1$ . This is coherent with the result given in Section 2.4.3., because a zero of  $S_{yp}(z^{-1})$  corresponds to a pole of the open loop system.

The transfer function between the disturbance  $p(t)$  and the input of the plant  $u(t)$  (*input sensitivity function*) is given by

$$S_{up}(z^{-1}) = \frac{-A(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} \quad (2.5.15)$$

The analysis of this function allows one to evaluate the influence of a disturbance upon the plant input, and to specify a factor of the polynomial  $R(z^{-1})$  if the controller must not react to disturbances concentrated in a particular frequency region.

When noise is added to the measured output (see Figure 2.28), important information can be retrieved by the transfer function that relates the noise  $b(t)$  to the plant output  $y(t)$  (*noise-output sensitivity function*).

$$S_{yb}(z^{-1}) = \frac{-B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} \quad (2.5.16)$$

As the noise energy is often concentrated at high frequency, attention should be paid in order to obtain a low gain of the transfer function  $|S_{yb}(e^{-j\omega})|$  in this frequency region.

For  $T=R$ , the sensitivity function between  $r$  and  $y$  (also called *complementary sensitivity function*) is defined as

$$S_{yr}(z^{-1}) = \frac{B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} = -S_{yb}(z^{-1}) \quad (2.5.17)$$

Note that

$$S_{yp}(z^{-1}) - S_{yb}(z^{-1}) = S_{yp}(z^{-1}) + S_{yr}(z^{-1}) = 1$$

which implies an interdependence between these sensitivity functions.

Notice that  $S_{ub}(z^{-1})$ , the transfer function between the noise and the plant input, is equal to  $S_{up}(z^{-1})$ .

Another important transfer function describes the influence on the output of a disturbance  $v(t)$  on the plant input. This sensitivity function (*input disturbance-output sensitivity function*) is given by

$$S_{yv}(z^{-1}) = \frac{B(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} \quad (2.5.18)$$

The importance of this sensitivity function is that it enhances the possible simplification of unstable plant poles by the zeros of  $R(z^{-1})$ .

In order to clarify this point, let us consider the assumption  $R(z^{-1})=A(z^{-1})$  (plant poles compensation by controller zeros) and suppose that the plant to be controlled is unstable ( $A(z^{-1})$  has roots outside the unit circle). In this case

$$\begin{aligned} S_{yp}(z^{-1}) &= \frac{A(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})A(z^{-1})} = \frac{S(z^{-1})}{S(z^{-1}) + B(z^{-1})} \\ S_{up}(z^{-1}) &= -\frac{A(z^{-1})A(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})A(z^{-1})} = -\frac{A(z^{-1})}{S(z^{-1}) + B(z^{-1})} \\ S_{yb}(z^{-1}) &= -\frac{B(z^{-1})A(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})A(z^{-1})} = -\frac{B(z^{-1})}{S(z^{-1}) + B(z^{-1})} \\ S_{yv}(z^{-1}) &= \frac{B(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})A(z^{-1})} = \frac{B(z^{-1})S(z^{-1})}{A(z^{-1})[S(z^{-1}) + B(z^{-1})]} \end{aligned}$$

Note that  $S_{yp}$ ,  $S_{up}$ ,  $S_{yb}$  are stable transfer functions if  $S(z^{-1})$  is chosen in order to have  $S(z^{-1})+B(z^{-1})$  stable, that is

$$S(z^{-1}) + B(z^{-1}) = 0 \Rightarrow |z| < 1$$

while the sensitivity function  $S_{yv}(z^{-1})$  is unstable.

This remark yields to the following general statement:

*The feedback system presented in Figure 2.28 is asymptotically stable if and only if all the four sensitivity functions  $S_{yp}$ ,  $S_{up}$ ,  $S_{yb}$  (or  $S_{yr}$ ) and  $S_{yv}$  (describing the relations between disturbances on one hand and plant input or output on the other hand) are asymptotically stable.*

The set of five transfer functions  $H_{OL}(z^{-1})$ ,  $S_{yp}(z^{-1})$ ,  $S_{up}(z^{-1})$ ,  $S_{yb}(z^{-1})$  (or  $S_{yr}(z^{-1})$ ) and  $S_{yv}(z^{-1})$  also play an important role in the closed loop system robustness analysis.

### 2.5.3 Control System with PI Digital Controller

In this section the design of digital PI controllers will be illustrated. The transfer (function) operator of the discretized plant with zero-order hold is given by

$$H(q^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{B(q^{-1})}{A(q^{-1})} = \frac{b_1 q^{-1}}{1 + a_1 q^{-1}} \quad (2.5.19)$$

*For the sake of notation uniformity, we shall often use, in the case of constant coefficients,  $q^{-1}$  notation both for the delay operator and the complex variable  $z^{-1}$ .*



The  $z^{-1}$  notation will be specially employed when an interpretation in the frequency domain is needed (in this case  $z = e^{j\omega T_s}$ ).

The digital PI controller is characterized by the polynomials (see Equations 2.5.6 and 2.5.7):

$$R(q^{-1}) = T(q^{-1}) = r_0 + r_1 q^{-1} \quad (2.5.20)$$

$$S(q^{-1}) = 1 - q^{-1} \quad (2.5.21)$$

The closed loop system transfer function (with respect to the reference  $r(t)$ ) in the general form is given by Equation 2.5.12.

The characteristic polynomial  $P(q^{-1})$ , whose roots are the desired closed loop system poles, essentially defines the performances. As a general rule, it is chosen as a second-order polynomial corresponding to the discretization of a second-order continuous-time system with a specified natural frequency  $\omega_0$  and damping  $\zeta$  ( $\omega_0$  and  $\zeta$ , for example, and can be obtained on the basis of the diagrams given in Figures 1.10 or 1.11) starting from specifications in the time domain. The coefficients corresponding to the polynomial  $P(q^{-1})$  are obtained either by conversion tables mentioned in Table 2.4, or by Scilab and MATLAB<sup>®</sup> functions given in Section 2.3. In this case, sampling period  $T_s$ , natural frequency  $\omega_0$  and damping  $\zeta$  must be specified.

We recall that the relation between  $\omega_0$  and  $T_s$  must be respected (see Section 2.2.2, Equation 2.2.7):

$$0.25 \leq \omega_0 T_s \leq 1.5 ; 0.7 \leq \zeta \leq 1 \quad (2.5.22)$$

For a plant having an equivalent discrete-time transfer operator (function) given by Equation 2.5.19, and the use of a digital PI controller, the closed loop system poles are given by Equation 2.5.13, and they are

$$(1 + a_1 q^{-1})(1 - q^{-1}) + b_1 q^{-1}(r_0 + r_1 q^{-1}) = 1 + p_1 q^{-1} + p_2 q^{-2} \quad (2.5.23)$$

By rearranging the terms in Equation 2.5.23 in ascending  $q^{-1}$  powers, we get

$$1 + (a_1 - 1 + r_0 b_1) q^{-1} + (b_1 r_1 - a_1) q^{-2} = 1 + p_1 q^{-1} + p_2 q^{-2} \quad (2.5.24)$$

For the polynomial Equation 2.5.24 to be verified, it is necessary that the coefficients of the same  $q^{-1}$  powers must be equal on both sides. Thus the following system is obtained:

$$\begin{cases} a_1 - 1 + r_0 b_1 & = & p_1 \\ b_1 r_1 - a_1 & = & p_2 \end{cases} \quad (2.5.25)$$

which gives for  $r_0$  and  $r_1$  the results

$$r_1 = \frac{p_2 + a_1}{b_1} \quad ; \quad r_0 = \frac{p_1 - a_1 + 1}{b_1} \quad (2.5.26)$$

One can see that the parameters of the controller depend upon the performance specifications (the desired closed loop poles) and the plant model parameters.

By using Equation 2.5.7, one can obtain the parameters of the continuous-time PI controller:

$$K = -r_1 \quad ; \quad T_i = -\frac{T_s r_1}{r_1 + r_0}$$

## 2.6 Analysis of the Closed Loop Sampled-Data Systems in the Frequency Domain

### 2.6.1 Closed Loop Systems Stability

In the case of continuous-time systems, it was shown in Chapter 1, Section 1.2.5, how to use the open loop transfer function representation in the complex plane (the Nyquist plot) in order to analyze the closed loop system stability and the robustness with respect to the parameters variations (or uncertainties on the parameters value). The same approach can be applied to the case of sampled-data systems. The Nyquist plot for sampled-data systems can be drawn using the functions *Nyquist-ol.sci* (Scilab) and *Nyquist-ol.m* (MATLAB®)<sup>6</sup>.

Figure 2.29 shows the Nyquist plot of an open loop sampled-data system including a plant (represented by the corresponding transfer function  $H(z^{-1}) = B(z^{-1}) / A(z^{-1})$ ) and a RST controller.

In this case, the open loop transfer function is given by

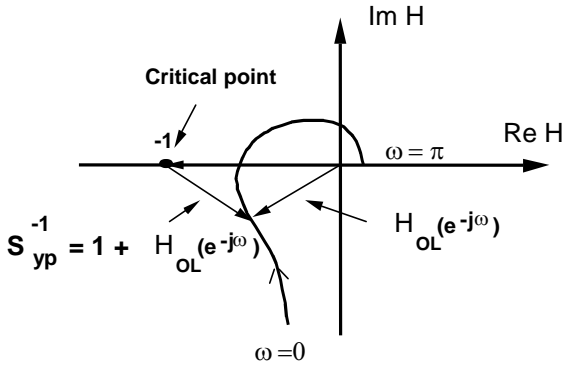
$$H_{OL}(e^{-j\omega}) = \frac{B(e^{-j\omega})R(e^{-j\omega})}{A(e^{-j\omega})S(e^{-j\omega})} \quad (2.6.1)$$

The vector linking the plane origin to a point belonging to the Nyquist plot of the transfer function represents  $H_{OL}(e^{-j\omega})$  for a specified normalized radian frequency  $\omega = \omega T_s = 2 \pi f / f_s$ . The considered range of variation of the radian natural

---

<sup>6</sup> To be downloaded from the book website.

frequency  $\omega$  is between  $0$  and  $\pi$  (corresponding to an unnormalized frequency variation between  $0$  and  $0.5 f_s$ ).



**Figure 2.29.** Nyquist plot for a sampled-data system transfer function and the critical point

In this diagram the point  $[-1, j0]$  is the “critical point”. As Figure 2.29 clearly shows, the vector linking the point  $[-1, j0]$  to the Nyquist plot of  $H_{OL}(e^{-j\omega})$  has the expression

$$S_{yp}^{-1}(z^{-1}) = 1 + H_{OL}(z^{-1}) = \frac{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})} \quad (2.6.2)$$

This vector represents the inverse of the output sensitivity function  $S_{yp}(z^{-1})$  (see Equation 2.5.14) and the zeros of  $S_{yp}^{-1}(z^{-1})$  correspond to the closed loop system poles (see Equation 2.5.13). In order to have an asymptotically stable closed loop system, it is necessary that all the zeros of  $S_{yp}^{-1}(z^{-1})$  (that are the poles of  $S_{yp}(z^{-1})$ ) be inside the unit circle ( $|z| < 1$ ). The necessary and sufficient conditions for the asymptotic stability of the closed loop system are given by the Nyquist criterion.

For systems having stable poles in open loop (in this case  $A(z^{-1}) \neq 0$  and  $S(z^{-1}) \neq 0 \rightarrow |z| \leq 1$ ) the Nyquist stability criterion states (stable open loop system): *The Nyquist plot of  $H_{OL}(z^{-1})$  traversed in the sense of growing frequencies (from  $\omega = 0$  to  $\omega = \pi$ ), leaves the critical point  $[-1, j0]$  on the left.*

As a general rule, for the given nominal plant model  $B(z^{-1})/A(z^{-1})$ , polynomials  $R(z^{-1})$  and  $S(z^{-1})$  are computed in order to have

$$A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1}) = P(z^{-1}) \quad (2.6.3)$$

where  $P(z^{-1})$  is a polynomial with asymptotically stable roots. As a consequence, for the nominal values of  $A(z^{-1})$  and  $B(z^{-1})$ , since the closed loop system is stable, the open loop transfer function:

$$H_{OL}(z^{-1}) = \frac{B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})}$$

does not encircle the critical point (if  $A(z^{-1})$  and  $S(z^{-1})$  have their roots inside the unit circle).

In the case of an *unstable open loop system*, either if  $A(z^{-1})$  has some pole outside the unit circle (unstable plant), or if the computed controller is unstable in open loop ( $S(z^{-1})$  has some pole outside the unit circle), the stability criterion is: *The Nyquist plot of  $H_{OL}(z^{-1})$  traversed in the sense of growing frequencies (from  $\omega = 0$  to  $\omega = \pi$ ), leaves the critical point  $[-1, j0]$  on the left and the number of counter clockwise encirclements of the critical point should be equal to the number of unstable poles of the open loop system<sup>7</sup>.*

Note that the Nyquist locus between  $0.5 f_s$  and  $f_s$  is the symmetric of the Nyquist locus between  $0$  and  $0.5 f_s$  with respect to the real axis.

The general Nyquist criterion formula that gives the number of encirclements around the critical point is

$$N = P_{CL}^i - P_{OL}^i$$

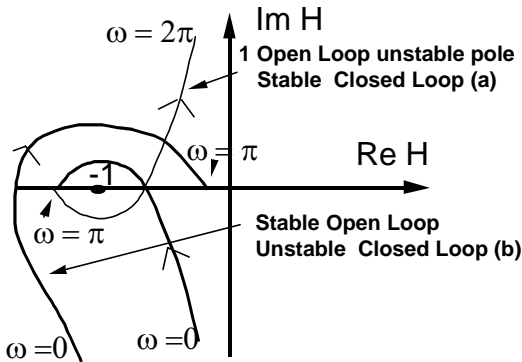
where  $P_{CL}^i$  is the number of closed loop unstable poles and  $P_{OL}^i$  is the number of open loop unstable poles. Positive values of  $N$  correspond to clockwise encirclements around the critical point. In order that the closed loop system be asymptotically stable it is necessary that  $N = -P_{OL}^i$ . Figure 2.30 shows two interesting Nyquist loci.

If the plant is stable in open loop and the controller is computed on the basis of Equation 2.6.3 to obtain a desired stable closed loop polynomial  $P(z^{-1})$  (this means that the nominal closed loop system is stable too), then, if a Nyquist plot of the form of Figure 2.30a is obtained, one concludes that the controller is unstable in open loop. This situation must be generally avoided<sup>8</sup>, and this can be achieved by reducing the desired closed loop dynamic performances (by modifying  $P(z^{-1})$ ).

---

<sup>7</sup> The criterion holds even if an unstable pole-zero cancellation occurs. The number of encirclements should be equal to the number of unstable poles without taking into account the possible cancellations.

<sup>8</sup> Note that there exist some « pathological » transfer functions  $B(z^{-1})/A(z^{-1})$  with unstable poles and/or zeros that can be only stabilized by controllers that are unstable in open loop.



**Figure 2.30.** Nyquist plots: **a)** unstable system in open loop but stable in closed loop; **b)** stable system in open loop but unstable in closed loop

## 2.6.2 Closed Loop System Robustness

When designing a control system, one has to take into account the plant model uncertainties (uncertainties of the parameter values or of the frequency characteristics, variations of the parameters, etc.). It is therefore extremely important to assess if the stability of the closed loop is guaranteed in the presence of the plant model uncertainties. The closed loop will be termed “robust” if the stability is guaranteed for a given set of model uncertainties.

The *robustness* of the closed loop is related to the minimal distance between the Nyquist plot for the nominal plant model and the “critical point” as well as to the frequency characteristics of the modulus of the sensitivity functions.

The following elements help to evaluate how far is the critical point  $[-1, j0]$  (see Figure 2.31):

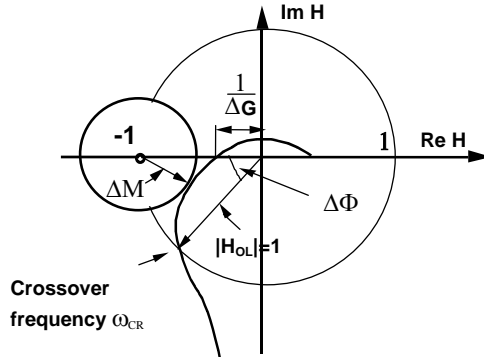
- Gain margin;
- Phase margin;
- Delay margin;
- Modulus margin.

### Gain Margin

The gain margin ( $\Delta G$ ) equals the inverse of  $H_{OL}(e^{-j\omega})$  gain for the frequency corresponding to a phase shift  $\angle\phi = -180^\circ$  (see Figure 2.31).

The gain margin is often expressed in dB. In other words, the gain margin gives the maximum admissible increase of the open loop gain for the frequency corresponding to  $\angle\phi(\omega) = -180^\circ$ .

$$\Delta G = \frac{1}{|H_{OL}(e^{-j\omega_{180}})|} \quad \text{for } \angle\phi(\omega_{180}) = -180^\circ$$



**Figure 2.31.** Gain, phase and modulus margins

Typical values for a good gain margin are

$$\Delta G \geq 2 \quad (6 \text{ dB}) \quad [\text{min: } 1.6 \text{ (4 dB)}]$$

If the Nyquist plot crosses the real axis at several frequencies  $\omega_{i\pi}$  characterized by a phase lag

$$\angle \phi(\omega_{i\pi}) = -i 180^\circ; \quad i = 1, 3, 5 \dots$$

and the corresponding gains of the open loop system are denoted by  $|H_{OL}(e^{-j\omega_{i\pi}})|$ , then the *gain margin* is defined by<sup>9</sup>

$$\Delta G = \min_i \frac{1}{|H_{OL}(e^{-j\omega_{i\pi}})|}$$

### Phase Margin

The phase margin ( $\Delta\phi$ ) is the additional phase that we must add at the crossover frequency, for which the gain of the open-loop system equals 1, in order to obtain a total phase shift  $\angle\phi = -180^\circ$  (see Figure 2.31).

$$\Delta\phi = 180^\circ - \angle\phi(\omega_{cr}) \quad \text{for} \quad |H_{OL}(e^{-j\omega_{cr}})| = 1$$

in which  $\omega_{cr}$  is called *crossover frequency* and it corresponds to the frequency for which the Nyquist plot crosses the unit circle (see Figure 2.31).

<sup>9</sup> Note that if the Nyquist plot crosses the real axis for values less than  $-1$  and leaves the critical point to the left, there is a minimal value of the gain margin under which the system becomes unstable.

Typical values for a good phase margin are

$$30^\circ \leq \Delta\phi \leq 60^\circ$$

If the Nyquist plot crosses the unit circle at several frequencies  $\omega_{cr}^i$  characterized by the corresponding phase margins:

$$\Delta\phi_i = 180^\circ - \angle\phi(\omega_{cr}^i)$$

then the system *phase margin* is defined as

$$\Delta\phi = \min_i \Delta\phi_i$$

### *Delay Margin*

A time-delay introduces a phase shift proportional to the frequency  $\omega$ . For a certain frequency  $\omega_0$ , the phase shift introduced by a time-delay  $\tau$  is

$$\angle\phi(\omega_0) = -\omega_0\tau$$

We can therefore convert the phase margin in a “time-delay margin”, i.e. to compute the maximum admissible increase of the delay of the open-loop system without making the closed-loop system unstable. It then follows that:

$$\Delta\tau = \frac{\Delta\phi}{\omega_{cr}}$$

If the Nyquist plot intersects the unit circle at several frequencies  $\omega_{cr}^i$ , characterized by the corresponding phase margins  $\Delta\phi_i$ , the *delay margin* is defined as

$$\Delta\tau = \min_i \frac{\Delta\phi_i}{\omega_{cr}^i}$$

Note that a *good phase margin does not guarantee a good delay margin* (if the frequency  $\omega_{cr}$  is high, then the delay margin is low even if the phase margin is important).

The typical value of the delay margin is  $\Delta\tau \geq T_S$  [min:  $0.75T_S$ ]

### *Modulus Margin*

This concerns a more global measure of the distance between the critical point  $[-1, j0]$  and the plot of  $H_{OL}(z^{-1})$ . The *modulus margin* ( $\Delta M$ ) is defined as the radius

of the circle centered in  $[-1, j0]$  and tangent to the plot of  $H_{OL}(z^{-1})$  (see Figure 2.31).

From the definition of Equation 2.6.2 of the vector connecting the critical point  $[-1, j0]$  to the plot of  $H_{OL}(e^{-j\omega})$  it follows immediately that

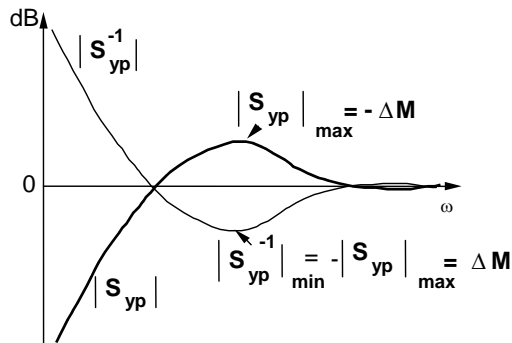
$$\Delta M = \left| 1 + H_{OL}(z^{-1}) \right|_{\min} = \left| S_{yp}^{-1}(z^{-1}) \right|_{\min} = \left( \left| S_{yp}(z^{-1}) \right|_{\max} \right)^{-1} =$$

$$\left( \left| \frac{A(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})} \right|_{\max} \right)^{-1} \quad \text{for } z^{-1} = e^{-j2\pi f}$$
(2.6.4)

In other words, the *modulus margin*  $\Delta M$  is equal to the inverse of the maximum value of the sensitivity function  $S_{yp}(z^{-1})$  magnitude. By plotting  $S_{yp}(z^{-1})$  magnitude in dB scale, the following condition is immediately derived:

$$\left| S_{yp}(e^{-j\omega}) \right|_{\max} dB = \Delta M^{-1} dB = -\Delta M dB$$
(2.6.5)

Figure 2.32 shows the relation between the sensitivity function  $S_{yp}$  and the modulus margin.



**Figure 2.32.** Relation between the output sensitivity function and the modulus margin

Therefore, the reduction (or the minimization) of  $|S_{yp}(j\omega)|_{\max}$  will lead to an increase (or maximization) of the modulus margin<sup>10</sup>.

Typical values for a good modulus margin are

$$\Delta M \geq 0.5 \text{ (-6 dB)} \quad [\text{min: } 0.4 \text{ (-8 dB)}]$$

<sup>10</sup>  $|S_{yp}(j\omega)|_{\max}$  corresponds to the  $H_{\infty}$  norm of the output sensitivity function.



Note that  $\Delta M \geq 0.5$  implies a gain margin  $\Delta G \geq 2$  (6 dB) and a phase margin  $\Delta\phi > 29^\circ$ . As a general rule, a good modulus margin guarantees satisfactory values for the gain and phase margins<sup>11</sup>.

To summarize, typical values for the stability margins in a robust design are:

- gain margin:  $\Delta G \geq 2$  (6dB) [min.: 1.6 (4dB)]
- phase margin:  $30^\circ \leq \Delta\phi \leq 60^\circ$
- delay margin:  $\Delta\tau = \frac{\Delta\phi}{\omega_{cr}} \geq T_s$  [min.: 0.75  $T_s$ ]
- modulus margin:  $\Delta M \geq 0.5$  (-6dB), [min.: 0.4 (-8dB)]

If the plant model is known with a very good precision for a certain region of operation, the imposed robustness margins can eventually be less restrictive.

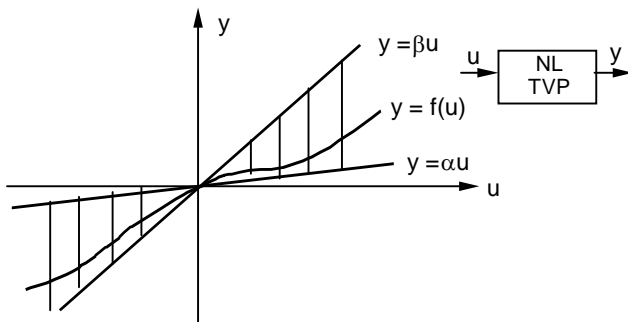
The *modulus margin* is very important because:

- It defines the maximum admissible value for the modulus of the *output sensitivity function* and therefore the low limits of the performance in disturbance rejection;
- It defines the tolerance with respect to nonlinear or time varying elements that may belong to the system (the circle criterion - see below).

#### *Tolerance with Respect to Nonlinear Elements*

In control systems we frequently have components with static nonlinear or time-varying characteristics (often in the actuators).

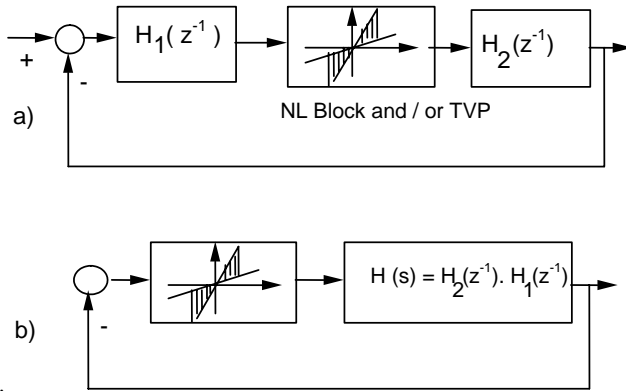
The characteristics of these components, without being accurately known, generally lie inside the conic region defined by a minimum linear gain ( $\alpha$ ) and a maximum linear gain ( $\beta$ ) – see Figure 2.33.



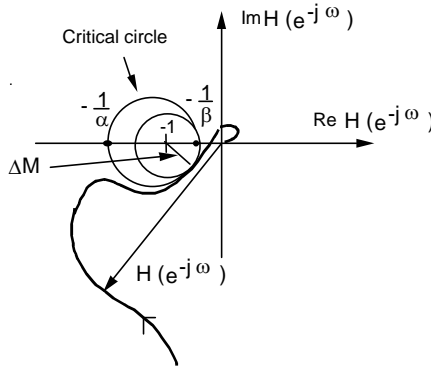
**Figure 2.33.** Nonlinear or time-varying characteristics, contained in the conic domain ( $\alpha$ ,  $\beta$ )

The closed-loop system looks, for example, like those in Figure 2.34a.

<sup>11</sup> The converse is not true. Systems having satisfactory gain and phase margins may have a very low modulus margin.



**Figure 2.34.** Closed-loop systems containing a nonlinear block (NL) and / or time-varying parameters (TVP): **a)** block diagram **b)** equivalent representation



**Figure 2.35.** Circle stability criterion and modulus margin for discrete time systems

From the stability analysis point of view, we may use an equivalent representation of such systems, given in Figure 2.34b, where  $H_{OL}(z^{-1}) = H_1(z^{-1})H_2(z^{-1})$ .

For this kind of system we have a generalization of the Nyquist criterion, known as “the circle criterion” (Popov-Zames).

*Circle (Stability) Criterion*

*The feedback system represented in Figure 2.34b is asymptotically stable for the set of nonlinear and/or time-varying characteristics lying in the conic domain  $[\alpha, \beta]$  (with  $\alpha, \beta > 0$ ) if the plot of  $H_{OL}(z^{-1})$ , traversed in the sense of growing frequencies, leaves on the left, without crossing it, the circle centered on the real axis and passes through the points  $[-1/\beta, j0]$  and  $[-1/\alpha, j0]$ .*

The modulus margin  $\Delta M$  defines a circle of radius  $\Delta M$  centered in  $[-1, j0]$  that is outside the Nyquist plot of the open loop transfer function.

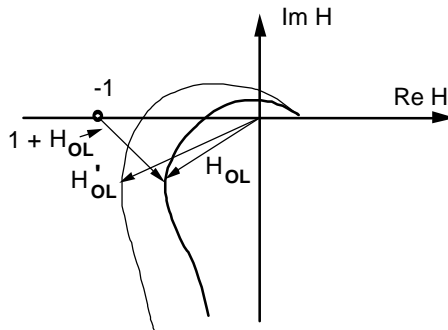
Thus, the closed loop system can tolerate non-linear blocks or time-variable parameters described by input-output characteristics lying in a conic sector delimited by a minimum linear gain  $(1/(1+\Delta M))$  and a maximum linear gain  $(1/(1-\Delta M))$  (see Figure 2.35).

*Tolerances to Plant Transfer Function Uncertainties and/or Parameters Variations.*

Figure 2.36 shows the effect of the plant nominal model uncertainties and parameters variations on the Nyquist plot of the open loop transfer function. As a general rule, the Nyquist plot of the plant nominal model lies inside a “tube” corresponding to the accepted tolerances of the parameters variations (or the uncertainties) of the plant model transfer function.

In order to ensure the stability of the closed loop system for an open loop transfer function  $H'_{OL}(z^{-1})$  that is different from the nominal one  $H_{OL}(z^{-1})$  (but having the same number of unstable poles as  $H_{OL}(z^{-1})$ ), it is necessary that the Nyquist plot of the open loop transfer function  $H'_{OL}(z^{-1})$  leaves the critical point  $[-1, j0]$  on the left when traversed in the sense of growing frequencies from  $0$  to  $0.5 f_s$ . This condition is satisfied if the difference between the real open loop transfer function  $H'_{OL}(z^{-1})$  and the nominal one  $H_{OL}(z^{-1})$  is smaller than the distance between the Nyquist plot of the open loop nominal transfer function and the critical point for all frequencies (see Figure 2.36). This robust stability condition is expressed by the inequality

$$\begin{aligned} \left| H'_{OL}(z^{-1}) - H_{OL}(z^{-1}) \right| < \left| 1 + H_{OL}(z^{-1}) \right| = \left| S_{yp}^{-1}(z^{-1}) \right| = \\ \left| \frac{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})} \right| = \left| \frac{P(z^{-1})}{A(z^{-1})S(z^{-1})} \right| ; \quad z^{-1} = e^{-j\omega} \end{aligned} \quad (2.6.6)$$



**Figure 2.36.** Nyquist plot for the nominal open loop transfer function and the real open loop transfer function in presence of uncertainties and parameters variations ( $H_{OL}$  and  $H'_{OL}$  are stable)

where  $S(z^{-1})$  and  $R(z^{-1})$  are computed on the basis of Equation 2.6.3 for the nominal values of  $A(z^{-1})$  and  $B(z^{-1})$ .

In other words, the magnitude of  $S_{yp}^{-1}(e^{j\omega})$  function (evaluated in dB units), obtained by symmetry from  $S_{yp}(e^{j\omega})$  (see Figure 2.32), gives, at each frequency, a sufficient condition for the accepted difference (computed as the Euclidian distance) between the real open loop transfer function and the nominal open loop transfer function, in order to guarantee the stability of the closed loop.

This tolerance is higher at low frequencies (see Figure 2.32) where the gain of the open loop system is high (especially when an integrator is included), and it has a minimum value at the frequency (or frequencies) where  $S_{yp}^{-1}(e^{j\omega})$  reaches its minimum ( $= \Delta M$ ), that is the frequency where  $S_{yp}(e^{j\omega})$  has the maximum value.

It is necessary to ensure that at these frequencies the plant model variations are compatible with the obtained modulus margin. If this is not the case, the solution is to provide a more accurate model, or to modify the specifications in order to maintain the closed loop stability.

Equation 2.6.6 expresses a robustness condition in terms of open loop transfer function variations (controller + plant). It is interesting to express this robustness condition in terms of the plant model variations only. Note that Equation 2.6.6 can be further expressed as

$$\begin{aligned} \left| \frac{B'(z^{-1})R(z^{-1})}{A'(z^{-1})S(z^{-1})} - \frac{B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})} \right| &= \left| \frac{R(z^{-1})}{S(z^{-1})} \right| \cdot \left| \frac{B'(z^{-1})}{A'(z^{-1})} - \frac{B(z^{-1})}{A(z^{-1})} \right| \\ &< \left| \frac{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}{A(z^{-1})S(z^{-1})} \right| \end{aligned} \quad (2.6.7)$$

where  $B(z^{-1})/A(z^{-1})$  corresponds to the nominal plant transfer function.

Multiplying by  $|S(z^{-1})/R(z^{-1})|$  both sides of Equation 2.6.7 one gets the condition

$$\begin{aligned} \left| \frac{B'(z^{-1})}{A'(z^{-1})} - \frac{B(z^{-1})}{A(z^{-1})} \right| &< \left| \frac{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}{A(z^{-1})R(z^{-1})} \right| \\ &= \left| \frac{P(z^{-1})}{A(z^{-1})R(z^{-1})} \right| = |S_{up}^{-1}(z^{-1})| \end{aligned} \quad (2.6.8)$$

By plotting the inverse of the input sensitivity function magnitude, sufficient conditions for tolerated (additive) variations (or uncertainties) of the plant transfer function are obtained. The inverse of the magnitude of the input sensitivity function is symmetric to the input sensitivity function magnitude in dB units with respect to the axis at 0 dB (see Figure 2.37).

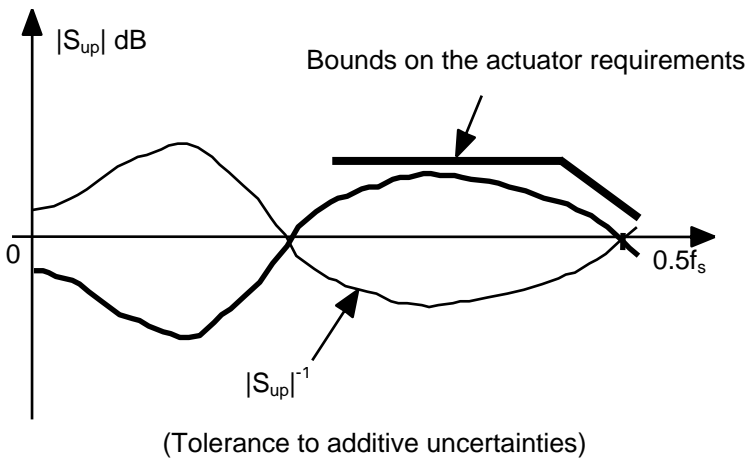
As plant model uncertainties at high frequencies are often present, one must verify that the maximum of  $|S_{up}^{-1}(e^{j\omega})|$  at high frequencies is small. On the other

hand, the input sensitivity function  $S_{up}$  is an effective image of the actuator stress in the frequency domain when disturbances act on the system. The physical characteristics of the actuator often impose a bound on actuator stress at high frequencies, and an upper bound of the maximum of  $|S_{up}(e^{j\omega})|$  at these frequencies should be imposed.

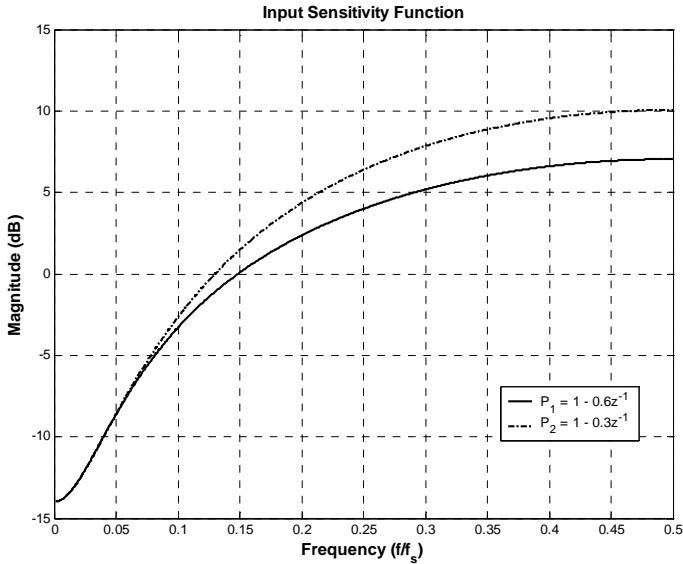
Notice that (from Equation 2.6.8) the admitted tolerances (neglecting the term  $1/R(z^{-1})$ ) depend to a large extent upon the relation between the open loop system poles (defined by  $A(z^{-1})$ ) and the desired closed loop poles (defined by  $P(z^{-1})$ ).

In order to understand this phenomenon in greater detail, Figure 2.38 shows the  $S_{up}(z^{-1})$  magnitude functions for a plant model characterized by  $A(z^{-1})=1-0.8z^{-1}$ ;  $B(z^{-1})=z^{-1}$  and for two different desired closed loop system characteristic polynomials:  $P_1(z^{-1})=1-0.6z^{-1}$  and  $P_2(z^{-1})=1-0.3z^{-1}$  (the controller includes an integrator). Note that  $P_2(z^{-1})$  corresponds to a closed loop system faster than the one specified by  $P_1(z^{-1})$ , and both closed loop systems are faster than the plant (open loop system).

The  $|S_{up}(z^{-1})|$  maximum for  $P_2(z^{-1})$  is greater than for  $P_1(z^{-1})$ , and then the inverse of  $|S_{up}(z^{-1})|$  will be smaller. As a consequence, the accepted tolerances for the frequency response variations (especially at high frequencies) are smaller in the case of  $P_2(z^{-1})$  with respect to the case of desired closed loop performances imposed by  $P_1(z^{-1})$ .



**Figure 2.37.** The input sensitivity function and its inverse



**Figure 2.38.** Input sensitivity function for the plant model  $q^{-1} / (1 - 0.8 q^{-1})$  as a function of the desired closed loop performances

Equation 2.6.8 gives a sufficient condition for the accepted (additive) tolerance in terms of real plant transfer function parameters variations (or uncertainties) with respect to the nominal plant transfer function.

Moreover, we may be interested in the evaluation of the accepted relative tolerance with respect to the nominal plant transfer function magnitude. It follows from Equation 2.6.8 that

$$\frac{\left| \frac{B'(z^{-1})}{A'(z^{-1})} - \frac{B(z^{-1})}{A(z^{-1})} \right|}{\left| \frac{B(z^{-1})}{A(z^{-1})} \right|} < \left| \frac{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}{B(z^{-1})R(z^{-1})} \right| \tag{2.6.9}$$

$$= \left| \frac{P(z^{-1})}{B(z^{-1})R(z^{-1})} \right| = \left| S_{yb}^{-1}(z^{-1}) \right| = \left| S_{yr}^{-1}(z^{-1}) \right|$$

where  $S_{yb}$  is the “noise-output sensitivity function” and  $S_{yr}$  is the *complementary sensitivity function*.

The noise-output sensitivity function  $S_{yb}$  allows the definition of a frequency “template” to ensure that the “delay margin” constraint is fulfilled. Let consider the case of a delay margin  $\Delta\tau = 1 \cdot T_s$ .

It follows that

$$H(z^{-1}) = \frac{z^{-d}B(z^{-1})}{A(z^{-1})} \quad ; \quad H'(z^{-1}) = \frac{z^{-d-1}B(z^{-1})}{A(z^{-1})} \quad (2.6.10)$$

and consequently

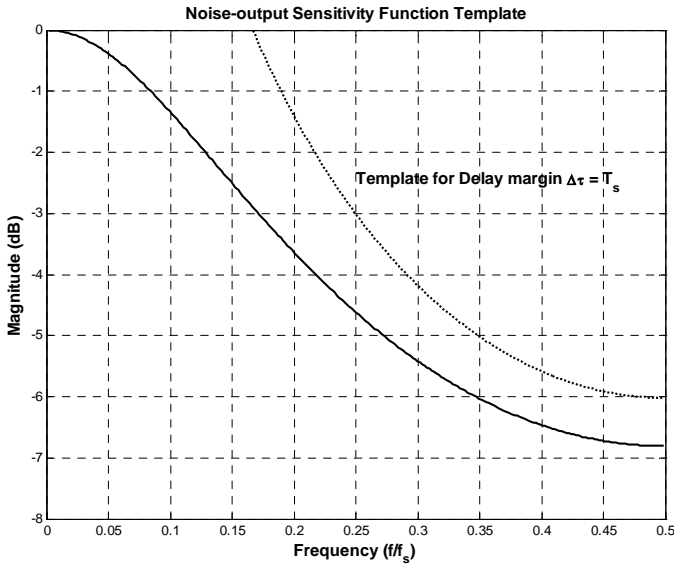
$$\frac{H'(z^{-1}) - H(z^{-1})}{H(z^{-1})} = z^{-1} - 1 \quad (2.6.11)$$

Equation 2.6.9 becomes

$$\left| S_{yb}(z^{-1}) \right| < \frac{1}{\left| z^{-1} - 1 \right|} \quad ; \quad z = e^{j\omega} \quad 0 \leq \omega \leq \pi \quad (2.6.12)$$

or in dB units

$$\left| S_{yb}^{-1}(z^{-1}) \right|_{dB} < -20 \log \left| 1 - z^{-1} \right| \quad ; \quad z = e^{j\omega} \quad 0 \leq \omega \leq \pi \quad (2.6.13)$$



**Figure 2.39.** Frequency template for the noise-output sensitivity function and  $\Delta\tau = T_s$

This expression defines a frequency robustness template for the sensitivity function  $S_{yb}$ . This template corresponds to the frequency response of a digital integrator and is represented in Figure 2.39.

As the modulus margin introduces a frequency template on the output sensitivity function ( $|S_{yp}(e^{-j\omega})| \leq -\Delta M ; 0 \leq \omega \leq \pi$ ), we are interested in finding what template is introduced by the delay margin on  $|S_{yp}|$ . From Equations 2.5.14 and 2.5.17 it results that

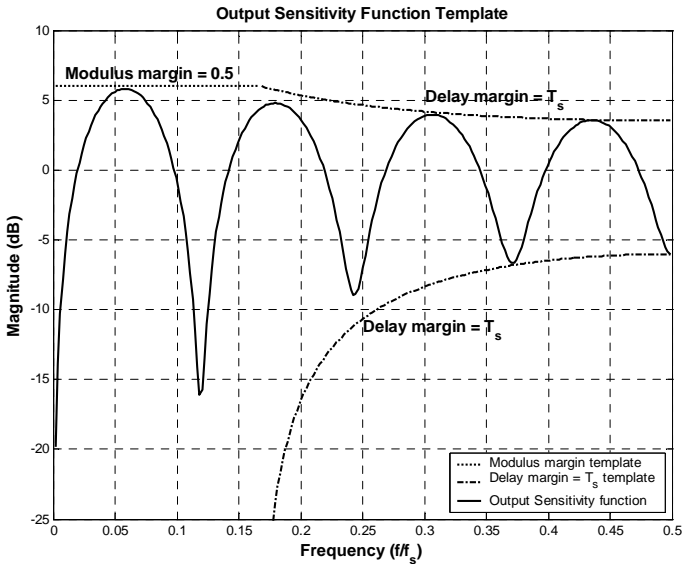
$$S_{yp}(z^{-1}) = 1 + S_{yb}(z^{-1}) \tag{2.6.14}$$

and by means of the triangle inequality it follows that

$$1 - |S_{yb}(z^{-1})| \leq |S_{yp}(z^{-1})| \leq 1 + |S_{yb}(z^{-1})| \tag{2.6.15}$$

Taking into account the frequency bound on  $S_{yb}$  given by Equation 2.6.12, the following condition is obtained :

$$1 - |1 - z^{-1}|^{-1} \leq |S_{yp}(z^{-1})| \leq 1 + |1 - z^{-1}|^{-1} \tag{2.6.16}$$



**Figure 2.40.** Frequency template on the output sensitivity function for  $\Delta M = 0.5$  (-6dB) and  $\Delta \tau = T_S$



that leads to the robustness template on  $|S_{yp}|$  represented in Figure 2.40.

Notice that, from the point corresponding to  $0.17 f_s$ ,  $|S_{yp}|$  must lie inside a region delimited by an upper and a lower bound and that, for frequencies below  $0.17 f_s$ , the frequency template for the modulus margin also assures the delay margin constraint to be respected.

It is important to note that the template on  $S_{yp}$  will not always guarantee the desired delay margin (it is an approximation). If the condition on  $|S_{yb}|$  is satisfied, then the condition on  $|S_{yp}|$  will also be satisfied. However, if the condition on  $|S_{yb}|$  is violated, this will not imply necessarily that the condition on  $|S_{yp}|$  will also be violated. In practice, the results obtained by using the template on the  $|S_{yp}|$  are very reliable.

The following remark is important: *the closed loop system robustness will be, in general, reduced when the closed loop system bandwidth is increased with respect to the open loop system bandwidth. Conversely, for a relevant reduction of the rise time for the closed loop system, with respect to the open loop system rise time, a good estimation of the plant model is required (especially in the frequency regions where  $|S_{yp}(z^{-1})|$  is high).*

As a consequence, robustness constraints can imply either a small reduction of the closed loop system rise time (with respect to the open loop system rise time), or a controller design which takes into account the bounds on the sensitivity functions.

An important challenge in control system design is the maximization of the controller robustness for given performances. This is obtained by minimizing the sensitivity functions maximum in the critical frequency regions.

## 2.7 Concluding Remarks

Recursive (differences) equations of the form

$$y(t) = -\sum_{i=1}^{n_A} a_i y(t-i) + \sum_{i=1}^{n_B} b_i u(t-d-i) \quad (2.7.1)$$

where  $u$  is the input,  $y$  is the output and  $d$  is the discrete-time delay, are used to describe discrete-time dynamic models.

The delay operator  $q^{-1}$  [ $q^{-1} y(t) = y(t-1)$ ] is a simple tool to handle recursive equations. If the operator  $q^{-1}$  is used, the recursive Equation 2.7.1 takes the form

$$A(q^{-1})y(t) = q^{-d} B(q^{-1})u(t)$$

where

$$A(q^{-1}) = 1 + \sum_{i=1}^{n_A} a_i q^{-i} \quad B(q^{-1}) = 1 + \sum_{i=1}^{n_B} b_i q^{-i}$$

The input-output relation for a discrete-time model is also conveniently described by the pulse transfer operator  $H(q^{-1})$ :

$$y(t) = H(q^{-1})u(t)$$

where

$$H(q^{-1}) = \frac{q^{-d} B(q^{-1})}{A(q^{-1})}$$

The pulse transfer function of a discrete-time linear system is expressed as function of the complex variable  $z = e^{sT_s}$  ( $T_s$  = sampling period). The pulse transfer function can be derived from the pulse transfer operator  $H(q^{-1})$  by replacing  $q^{-1}$  with  $z^{-1}$ .

The asymptotical stability of a discrete-time model is ensured if, and only if, all pulse transfer function poles (in  $z$ ) lie inside the unit circle.

The order of a pulse transfer function is

$$n = \max(n_A, n_B + d)$$

In computer controlled systems, the input signal applied to the plant is held constant between two sampling instants by means of a zero-order hold (ZOH). The zero-order hold is characterized by the following transfer function:

$$H_{ZOH}(s) = \frac{1 - e^{-sT_s}}{s}$$

Therefore, the continuous-time part of the system (between digital-to-analog converter and the analog-to-digital converter) is characterized by the continuous-time transfer function

$$H^*(s) = H_{ZOH}(s) \cdot H(s)$$

where  $H(s)$  is the plant transfer function.

In computer controlled systems, the input signal applied to the plant at time  $t$  is a weighted average of the plant output at times  $t, t-1, \dots, t-n_A+1$ , of the previous input signal values at instants  $t-1, t-2, \dots, t-n_B-d$ , and of the reference signal at

instants  $t, t-1, \dots$ , the weights being the coefficients of the controller. The corresponding control law (controller RST) is written as

$$S(q^{-1})u(t) = -R(q^{-1})y(t) + T(q^{-1})r(t) \quad (2.7.2)$$

where  $u(t)$  is the control (input) signal to the plant,  $y(t)$  is the plant output and  $r(t)$  is the reference.

The transfer function of the closed loop system (between the reference signal and the plant output) that includes the digital controller of Equation 2.7.2 is given by

$$H_{CL}(z^{-1}) = \frac{B(z^{-1})T(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}$$

where  $H(z^{-1}) = B(z^{-1})/A(z^{-1})$  is the pulse transfer function of the discretized plant (in this case  $B(z^{-1})$  may include possible delays).

The characteristic polynomial defining the closed loop system poles is given by

$$P(z^{-1}) = A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})$$

The disturbance rejection properties on the output result from the output sensitivity function frequency response

$$S_{yp}(z^{-1}) = \frac{A(z^{-1})S(z^{-1})}{A(z^{-1})S(z^{-1}) + B(z^{-1})R(z^{-1})}$$

Robust stability of the closed loop system, with respect to the plant transfer function uncertainties or parameters variations, is essentially characterized by the *modulus margin and the delay margin*.

The *modulus margin* and the *delay margin* introduce frequency constraints on the magnitude of the sensitivity functions. These constraints lead to the definition of frequency robustness templates that must be respected.

The robust stability (or performance) of the closed loop system robustness, with respect to the plant transfer function uncertainties or parameters variations, depends upon the choice of the desired closed loop system performances (bandwidth, rise time) with respect to the open loop system dynamics. A significant reduction of the closed loop system rise time (or a significant augmentation of the bandwidth of the closed loop system), compared to the open loop system rise time (or bandwidth), requires a good estimation of the plant model.

In order to ensure closed loop system robustness, when a good estimation of the plant model is not available, or when large system parameters variations occur, the closed loop system rise time acceleration, compared to the open loop system rise time, must be moderate. However, some methods exist for maximizing the

controller robustness with respect to plant model uncertainties (or parameters variations), for given nominal performance.

## 2.8 Notes and References

Some basic books on computer control system are:

- Kuo B. (1980) Digital Control Systems, Holt Saunders, Tokyo.  
 Åström K.J., Wittenmark B. (1997) Computer Controlled Systems - Theory and Design, 3rd edition, Prentice-Hall, Englewood Cliffs, N.J.  
 Ogata K. (1987) Discrete-Time Control Systems, Prentice Hall, N.J.  
 Franklin G.F., Powell J.D., Workman M.L. (1998) Digital Control of Dynamic Systems, 3rd edition, Addison Wesley, Reading, Mass.  
 Wirk G.S. (1991) Digital Computer Systems, MacMillan, London.  
 Phillips, C.L., Nagle, H.T. (1995) Digital Control Systems Analysis and Design, 3rd edition, Prentice Hall, N.J.

For an introduction to robust control theory see:

- Doyle J.C., Francis B.A., Tanenbaum A.R. (1992) Feedback Control Theory, Mac Millan, N.Y.  
 Kwakernaak H. (1993) Robust control and  $H_{\infty}$  optimization – a tutorial Automatica, vol.29, pp.255-273.  
 Morari M., Zafiriou E. (1989) Robust Process Control, Prentice Hall International, Englewood Cliffs, N.J.

The delay margin was introduced by:

- Anderson B.D.O, Moore J.B. (1971) Linear Optimal Control, Prentice Hall, Englewood Cliffs, N.J.

For a detailed discussion on the delay margin:

- Bourlès H., Irving E. (1991) La méthode LQG/LTR: une interprétation polynomiale, temps continu / temps discret, RAIRO-APII, Vol. 25, pp. 545-568.  
 Landau I.D. (1995) Robust digital control systems with time delay (the Smith predictor revisited), Int. J. of Control, Vol. 62, no. 2, pp. 325-347.

The *circle criterion* is presented in:

- Zames G. (1966) On the input-output stability of time-varying nonlinear feedback systems, IEEE-TAC, vol. AC-11, April, pp. 228-238, July pp. 445-476.  
 Narendra K.S., Taylor J.H. (1973) Frequency Domain Criteria for Absolute Stability, Academic Press, New York.

For a generalization of the concepts of robustness margins and robust stability introduced in this chapter, see Appendix D.