# CHAPTER 4

# Discrete Equivalents to Continuous Transfer Functions: The Digital Filter

# 4.1 INTRODUCTION

One of the exciting fields of application of digital systems<sup>1</sup> is in signal processing and digital filtering. A filter is a device designed to pass desirable elements and hold back or reject undesirable ones; in signal processing it is common to represent signals as a sum of sinusoids and to define the "desirable elements" as those signals whose frequency components are in a specified band. Thus a radio receiver filter passes the band of frequencies transmitted by the station we want to hear and rejects all others. We would call such a filter a *bandpass filter*. In electrocardiography it often happens that power-line frequency signals are strong and unwanted, so we design a filter to pass signals between 1 and 500 Hz but to eliminate those at 60 Hz. The magnitude of the transfer function for this purpose may look like Fig. 4.1 on a log-frequency scale, where the amplitude response between 59.5 and 60.5 Hz might reach  $10^{-3}$ . Here we have a band-reject filter with a 60-dB rejection ratio in a 1-Hz band centered at 60 Hz.

In long-distance telephony some filters play a conceptually different role. There the issue is that transmission media—wires or microwaves—introduce distortion in the amplitude and phase of a sinusoid which must be removed. Filters to accomplish this correction are called *equalizers*. And in control we must control systems whose dynamics require modification in order that the complete system have satisfactory dynamic response. We call the devices that make these changes *compensators*.

<sup>&</sup>lt;sup>1</sup>Including microprocessors and special-purpose, very large-scale integration (VLSI) digital chips for signal processing, called DSP chips.



Figure 4.1 Magnitude of a low-frequency bandpass filter with a narrow rejection band.

Whatever the name-filter, equalizer, or compensator-many fields have found use for devices having specified characteristics of amplitude and phase transmission, and the trend is to perform these functions by digital means. The design of electronic filters is a well-established subject that includes not only very sophisticated techniques but also well-tested computer programs [Van Valkenburg(1982)]. Much of the effort in digital filter design has been directed toward the design of digital filters that have the same characteristics (as nearly as possible) as those of a satisfactory continuous design. For digital control systems we have much the same motivation: Because continuouscontrol designs are well established, we should like to know how to take advantage of a good continuous design and cause a digital computer to produce a discrete equivalent to the continuous compensator. This method of design is called *emulation*. Although much of our presentation in this book is oriented toward direct digital design and away from emulation of continuous designs with digital equivalents, it is important to understand the techniques of discrete equivalents for purposes of comparison and because it is widely used by practicing engineers.

Thus we are led to the specific problem of this chapter: Given a transfer function, H(s), what discrete transfer function will have approximately the same characteristics? We will present here three approaches to this task:

Method 1: numerical integration.

Method 2: pole-zero mapping.

Method 3: hold equivalence.

# 4.2 DESIGN OF DISCRETE EQUIVALENTS BY NUMERICAL INTEGRATION

The topic of numerical integration of differential equations is quite complex, and only the most elementary techniques are presented here. For example, we only consider formulas of low complexity and fixed step-size. The fundamental concept is to represent the given filter transfer function H(s) as a differential equation and to derive a difference equation whose solution is an approximation to that of the differential equation. For example, the system

$$\frac{U(s)}{E(s)} = H(s) = \frac{a}{s+a} \tag{4.1}$$

is equivalent to the differential equation

$$\dot{u} + au = ae. \tag{4.2}$$

Now, if we write (4.2) in integral form, we have a development much like that of (2.5) in Chapter 2, except that the integral is more complex here:

$$u(t) = \int_0^t [-au(\tau) + ae(\tau)] d\tau,$$
  

$$u(kT) = \int_0^{kT-T} [-au + ae] d\tau + \int_{kT-T}^{kT} [-au + ae] d\tau$$
  

$$= u(kT - T) + \left\{ \begin{array}{l} \operatorname{area of} -au + ae \\ \operatorname{over} kT - T \le \tau < kT \end{array} \right\}.$$
(4.3)

We can now develop many rules based on our selection of the approximation of the incremental area term. The first approximation leads to the forward rectangular rule<sup>2</sup> wherein we approximate the area by the rectangle looking forward from kT - T and take the amplitude of the rectangle to be the value of the integrand at kT - T. The width of the rectangle is T. The result is an equation in the first approximation,  $u_1$ :

$$u_1(kT) = u_1(kT - T) + T[-au_1(kT - T) + ae(kT - T)]$$
  
= (1 - aT)u\_1(kT - T) + aTe(kT - T). (4.4a)

<sup>2</sup>Also known as Euler's rule.

The transfer function corresponding to the forward rectangular rule in this case is

$$H_F(z) = \frac{aTz^{-1}}{1 - (1 - aT)z^{-1}}$$
  
=  $\frac{a}{(z - 1)/T + a}$  (forward rectangular rule). (4.4b)

A second rule follows from taking the amplitude of the approximating rectangle to be the value looking backward from kT toward kT-T, namely, -au(kT) + ae(kT). The equation for  $u_2$ , the second approximation,<sup>3</sup> is

$$u_2(kT) = u_2(kT - T) + T[-au_2(kT) + ae(kT)]$$
  
=  $\frac{u_2(kT - T)}{1 + aT} + \frac{aT}{1 + aT}e(kT).$  (4.5a)

Again we take the z-transform and compute the transfer function of the backward rule:

$$H_B(z) = \frac{aT}{1+aT} \frac{1}{1-z^{-1}/(1+aT)} = \frac{aTz}{z(1+aT)-1}$$
  
=  $\frac{a}{(z-1)/Tz+a}$  (backward rectangular rule). (4.5b)

Our final version of integration rules is the *trapezoid rule* found by taking the area approximated in (4.3) to be that of the trapezoid formed by the average of the previously selected rectangles. The approximating difference equation is

$$u_{3}(kT) = u_{3}(kT - T) + \frac{T}{2}[-au_{3}(kT - T) + ae(kT - T) - au_{3}(kT) + ae(kT)] = \frac{1 - (aT/2)}{1 + (aT/2)}u_{3}(kT - T) + \frac{aT/2}{1 + (aT/2)}[e_{3}(kT - T) + e_{3}(kT)].$$
(4.6a)

<sup>3</sup>It is worth noting that in order to solve for (4.5a) we had to eliminate u(kT) from the right-hand side where it entered from the integrand. Had (4.2) been nonlinear, the result would have been an implicit equation requiring an iterative solution. This topic is the subject of predictor-corrector rules, which are beyond our scope of interest. A discussion is found in most books on numerical analysis. See Golub and Van Loan(1983). The corresponding transfer function from the trapezoid rule is

$$H_T(z) = \frac{aT(z+1)}{(2+aT)z + aT - 2}$$
  
=  $\frac{a}{(2/T)[(z-1)/(z+1)] + a}$  (trapezoid rule). (4.6b)

Suppose we tabulate our results obtained thus far.

H(s)	Method	Transfer function	
$\frac{a}{s+a}$	Forward rule	$H_F = \frac{a}{(z-1)/T + a}$	
$\frac{a}{s+a}$	Backward rule	$H_B = \frac{a}{(z-1)/Tz + a}$	
$\frac{a}{s+a}$	Trapezoid rule	$H_T = \frac{a}{(2/T)\{(z-1)/(z+1)\} + a}$	
			(4.7)

From direct comparison of H(s) with the three approximations in this tabulation we can see that the effect of each of our methods is to present a discrete transfer function that can be obtained from the given Laplace transfer function H(s) by substitution of an approximation for the frequency variable as shown below:

Method	Approximation
Forward rule	$s \leftarrow \frac{z-1}{T}$
Backward rule	$s \leftarrow rac{z-1}{T  ilde{z}}$
Trapezoid rule	$s \leftarrow \frac{2}{T} \frac{z-1}{z+1}$

The trapezoid-rule substitution is also known, especially in digital and sampled-data control circles, as *Tustin's method* [Tustin (1947)] after the British engineer whose work on nonlinear circuits stimulated a great deal of interest in this approach. The transformation is also called the *bilinear transformation* from consideration of its mathematical form. The design method can be summarized by stating the rule: Given a continuous transfer function (filter), H(s), a discrete equivalent can be found by the substitution

$$H_T(z) = H(s)|_{s=(2/T)[(z-1)/(z+1)]}.$$
(4.9)

(4.8)

Each of the approximations given in (4.8) can be viewed as a map from the *s*-plane to the *z*-plane. A further understanding of the maps can be obtained by considering them graphically. For example, because the  $(s = j\omega)$ axis is the boundary between poles of stable systems and poles of unstable systems, it would be interesting to know how the  $j\omega$ -axis is mapped by the three rules and where the left (stable) half of the *s*-plane appears in the *z*-plane. For this purpose we must solve the relations in (4.8) for *z* in terms of *s*. We find

i) 
$$z = 1 + Ts$$
, forward rectangular rule,  
ii)  $z = \frac{1}{1 - Ts}$ , backward rectangular rule,  
iii)  $z = \frac{1 + Ts/2}{1 - Ts/2}$ , bilinear rule. (4.10)

If we let  $s = j\omega$  in these equations, we obtain the boundaries of the regions in the z-plane which originate from the stable portion of the s-plane. The shaded areas sketched in the z-plane in Fig. 4.2 are these stable regions for each case. To show that rule (ii) results in a circle,  $\frac{1}{2}$  is added to and subtracted from the right-hand side to yield

$$z = \frac{1}{2} + \left\{ \frac{1}{1 - Ts} - \frac{1}{2} \right\}$$
$$= \frac{1}{2} - \frac{1}{2} \frac{1 + Ts}{1 - Ts}.$$
(4.11)

Now it is easy to see that with  $s = j\omega$ , the magnitude of  $z - \frac{1}{2}$  is constant,

$$|z-\frac{1}{2}|=\frac{1}{2},$$

and the curve is thus a circle as drawn in Fig. 4.2(b). Because the unit circle is the stability boundary in the z-plane, it is apparent from Fig. 4.2 that the forward rectangular rule could cause a stable continuous filter to be mapped into an unstable digital filter.

It is especially interesting to notice that the bilinear rule maps the stable region of the s-plane exactly into the stable region of the z-plane although the entire  $j\omega$ -axis of the s-plane is stuffed into the  $2\pi$ -length of the unit circle! Obviously a great deal of distortion takes place in the mapping in spite of the congruence of the stability regions. As our final rule deriving from numerical



Figure 4.2 Maps of the left-half s-plane to the z-plane by the integration rules of (4.8). Stable s-plane poles map into the shaded regions in the z-plane. The unit circle is shown for reference. (a) Forward rectangular rule. (b) Backward rectangular rule. (c) Trapezoid or bilinear rule.

integration ideas, we discuss a formula that extends Tustin's rule one step in an attempt to correct for the inevitable distortion of real frequencies mapped by the rule. We begin with our elementary transfer function (4.1)and consider the bilinear rule approximation

$$H_T(z) = rac{a}{(2/T)[(z-1)/(z+1)] + a}.$$

The original H(s) had a pole at s = -a, and for real frequencies,  $s = j\omega$ , the magnitude of  $H(j\omega)$  is given by

$$|H(j\omega)|^{2} = \frac{a^{2}}{\omega^{2} + a^{2}}$$
$$= \frac{1}{\omega^{2}/a^{2} + 1}.$$
(4.12)

Thus our reference filter has a half-power point,  $|H|^2 = \frac{1}{2}$ , at  $\omega = a$ . It will be interesting to know where  $H_T(z)$  has a half-power point.

As we saw in Chapter 2, signals with poles on the imaginary axis in the s-plane (sinusoids) map into signals on the unit circle of the z-plane. A sinusoid of frequency  $\omega_1$  corresponds to  $z_1 = e^{j\omega_1 T}$ , and the response of  $H_T(z)$  to a sinusoid of frequency  $\omega_1$  is  $H_T(z_1)$ . We consider now (4.6b) for  $H_T(z_1)$  and manipulate it into a more convenient form for our present

#### 4.3 ZERO-POLE MAPPING EQUIVALENTS 147

These results can be tabulated for convenient reference. Suppose we have a continuous system described by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}e(t),$$
$$u(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}e(t).$$

Then a discrete equivalent at sampling period T will be described by

$$\mathbf{w}(k+1) = \mathbf{\Phi}\mathbf{w}(k) + \mathbf{\Gamma}e(k),$$
  
 $u(k) = \mathbf{H}\mathbf{w}(k) + \mathbf{J}e(k),$ 

where  $\Phi$ ,  $\Gamma$ , H, and J are given as follows:<sup>7</sup>

	Forward	Backward	Bilinear
Φ	$(\mathbf{I} + \mathbf{A}T)$	$(I - AT)^{-1}$	$(I + \frac{AT}{2})(I - \frac{AT}{2})^{-1}$
Г	$\mathbf{B}T$	$(\mathbf{I} - \mathbf{A}T)^{-1}\mathbf{B}T$	$(\mathbf{I} - \frac{\mathbf{A}T}{2})^{-1}\mathbf{B}\sqrt{T}$
н	С	$C(I - AT)^{-1}$	$\sqrt{T}C(I-\frac{AT}{2})^{-1}$
J	D	$\mathbf{D} + \mathbf{C}(\mathbf{I} - \mathbf{A}T)^{-1}\mathbf{B}T$	$D + C(I - \frac{AT}{2})^{-1}BT/2$

# 4.3 ZERO-POLE MAPPING EQUIVALENTS

A very simple but effective method of obtaining a discrete equivalent to a continuous transfer function is to be found by extrapolation of the relation derived in Chapter 2 between the s- and z-planes. If we take the z-transform of samples of a continuous signal e(t), then the poles of the discrete transform E(z) are related to the poles of E(s) according to  $z = e^{sT}$ . We must go through the z-transform process to locate the zeros of E(z), however. The idea of the zero-pole mapping technique is that the map  $z = e^{sT}$  could reasonably be applied to the zeros also. The technique consists of a set of heuristic rules for locating the zeros and poles and setting the gain of a z-transform that will describe a discrete, equivalent transfer function that approximates the given H(s). The rules are as follows:

1. All poles of H(s) are mapped according to  $z = e^{sT}$ . If H(s) has a pole at s = -a, then  $H_{zp}(z)$  has a pole at  $z = e^{-aT}$ . If H(s) has a pole at -a + jb then  $H_{zp}(z)$  has a pole at  $re^{j\theta}$ , where  $r = e^{-aT}$  and  $\theta = bT$ .

<sup>&</sup>lt;sup>7</sup>These formulas are easily implemented in CAD packages; for example, see EQUIVNT in Table E.1 in Appendix E.

- 2. All finite zeros are also mapped by  $z = e^{sT}$ . If H(s) has a zero at  $s = -then H_{zp}(z)$  has a zero at  $z = e^{-aT}$ , and so on.
- 3. The zeros of H(s) at  $s = \infty$  are mapped in  $H_{zp}(z)$  to the point z =
  - a) One zero of H(s) at  $s = \infty$  is mapped into  $z = \infty$ . That is,  $H_{zp}$  is left with the number of zeros one less than the number of poles the finite plane. The series expansion of H(z) in powers of  $z^{-1}$  whave no constant term, and thus the corresponding pulse response h(k), has a one-unit delay. This choice means that the computer h one sample period to do the computation that corresponds to the transfer function because there is no direct transmission term.
- 4. The gain of the digital filter is selected to match the gain of H(s) at band center or a similar critical point. In most control applications, the critical frequency is s = 0, and hence we typically select the gain so the select the gain set of the gain set of the select the select the gain set of the select the sele

$$H(s)|_{s=0} = H_{zp}(z)|_{z=1}.$$

The rationale behind rule 3 is that the map of real frequencies fr  $j\omega = 0$  to increasing  $\omega$  is onto the unit circle at  $z = e^{j0} = 1$  until z  $e^{j\pi} = -1$ . Thus the point z = -1 represents, in a real way, the high frequency possible in the discrete transfer function, so it is appropriate t if H(s) is zero at the highest (continuous) frequency,  $|H_{zp}(z)|$  should be z at z = -1, the highest frequency that can be processed by the digital filt

**Example 4.2:** Application of these rules to H(s) = a/(s+a) gi

$$H_{zp}(z) = \frac{(z+1)(1-e^{-aT})}{2(z-e^{-aT})},$$
(4.)

or, using rule 3(a), we get

$$H_{zp}(z) = \frac{1 - e^{-aT}}{z - e^{-aT}}.$$
(4.)

A state-space algorithm to generate the zero-pole equivalent is also reily constructed with the utilities of a control design package.<sup>8</sup> The freque

<sup>8</sup>See EQUIVNT in Table E.1.

response of the zero-pole equivalent was found for the third-order case of Example 4.1 and is shown in Fig. 4.8 along with other equivalents for purposes of comparison.

# 4.4 HOLD EQUIVALENTS

For this technique, we construct the situation sketched in Fig. 4.4. The purpose of the samplers in Fig. 4.4(b) is to require that  $H_{h0}(z)$  has a discrete signal at its input and produces a discrete signal at its output, and thus  $H_{h0}(z)$  can be realized as a discrete transfer function. The philosophy of the design is the following. We are asked to design a discrete system that, with an input consisting of samples of e(t), has an output that approximates the output of the continuous filter H(s) whose input is the continuous e(t). We generate the discrete equivalent by first approximating e(t) from the samples e(k) and then putting this  $\hat{e}(t)$  through the given H(s). First consider the possibilities for approximation. These are techniques for taking a sequence of samples and extrapolating or holding them to produce a continuous signal.<sup>9</sup> Suppose we have the e(t) as sketched in Fig. 4.5. This figure also shows a sketch of a piecewise constant approximation to e(t) obtained by the operation of holding  $\hat{e}(t)$  constant at e(k) over the interval from kT to (k+1)T. This operation is the zero-order hold (or ZOH) we've discussed before. If we use a first-order polynomial for extrapolation, we have a first-order hold (or FOH), and so on for second-, and *n*th-order holds.

### 4.4.1 Zero-Order Hold Equivalent

If the approximating hold is the zero-order hold, then we have for our approximation exactly the same situation that in Chapter 2 was analyzed as a sampled-data system.<sup>10</sup> Therefore, the zero-order-hold equivalent to H(s) is given by (2.39), or

$$H_{h0}(z) = (1 - z^{-1})\mathcal{Z}\left\{\frac{H(s)}{s}\right\}.$$
 (4.31)

<sup>&</sup>lt;sup>9</sup>Some books on digital-signal processing suggest using no hold at all, using the equivalent  $H(z) = \mathbb{Z}{H(s)}$ . This choice is called the z-transform equivalent.

<sup>&</sup>lt;sup>10</sup>Recall that we noticed in Chapter 3 that the signal  $\hat{e}$  is, on the average, delayed from e by T/2 sec. The size of this delay is one measure of the quality of the approximation and can be used as a guide to the selection of T.





An example will fix ideas.

Example 4.3: Suppose we again take the first-order filter

$$H(s)=\frac{a}{s+a}.$$

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

Then

and

$$\mathcal{Z}\left\{\frac{H(s)}{s}\right\} = \mathcal{Z}\left\{\frac{1}{s}\right\} - \mathcal{Z}\left\{\frac{1}{s+a}\right\},\qquad(4)$$





### 4.4 HOLD EQUIVALENTS 151





and, by definition of the operation given in (4.32),

$$\mathcal{Z}\left\{\frac{H(s)}{s}\right\} = \sum_{0}^{\infty} z^{-k} - \sum_{0}^{\infty} z^{-k} e^{-akT}$$
$$= \frac{1}{1 - z^{-1}} - \frac{1}{1 - e^{-aT}z^{-1}}$$
$$= \frac{(1 - e^{-aT}z^{-1}) - (1 - z^{-1})}{(1 - z^{-1})(1 - e^{-aT}z^{-1})}.$$
(4.33)

Finally, substituting (4.33) in (4.31), we get the zero-order-hold equivalent of H(s), namely,

$$H_{h0}(z) = \frac{(1 - e^{-aT})}{z - e^{-aT}}.$$
(4.34)

We note that for the trivial example given, the zero-order-hold equivalent of (4.34) is identical to the matched zero-pole equivalent given by (4.30). However, this is not generally true as is evident in the comparison with other equivalents for the third-order example (4.1) in Fig. 4.8.

## 4.4.2 Triangle Hold Equivalent

An interesting hold equivalent can be constructed by imagining that we have a noncausal hold impulse response, as sketched in Fig. 4.6. The result is called the triangle-hold equivalent. The effect of this hold filter is to extrapolate the samples so as to connect sample to sample in a straight line. Although the continuous system is noncausal, the discrete equivalent is not.

Thus the triangle equivalent of  $[\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{J}]$  with sample period T is given by

$$\begin{split} \mathbf{A} &= \mathbf{\Phi}, \\ \mathbf{B} &= \mathbf{\Gamma}_1 + \mathbf{\Phi}\mathbf{\Gamma}_2 - \mathbf{\Gamma}_2, \\ \mathbf{C} &= \mathbf{H}, \\ \mathbf{D} &= \mathbf{J} + \mathbf{H}\mathbf{\Gamma}_2, \end{split} \tag{4.43}$$

where  $\Phi$ ,  $\Gamma_1$ , and  $\Gamma_2$  are defined by (4.40).<sup>11</sup>

In Fig. 4.8 the responses of the zero-pole, the zero-order hold, and the triangle-hold equivalents are compared, again for the third-order lowpass filter. We notice in particular that the triangle hold has excellent phase responses, even with sampling period of T = 2, which corresponds to a sampling frequency to passband frequency ratio of only  $\omega_s/\omega_p = \pi$ .

# 4.5 SUMMARY

In this chapter we have presented several techniques for the construction of discrete equivalents to continuous transfer functions so that known design methods for continuous systems—controls and filters—can be used as a basis for the design of discrete systems. The methods presented were:

- 1. Numerical integration
  - a) Forward rectangular rule
  - b) Backward rectangular rule
  - c) Trapezoid or Tustin's rule
  - d) Bilinear transformation with prewarping
- 2. Zero-pole mapping
- 3. Hold equivalence

All methods, except the forward rectangular rule, guarantee a stable, discrete system from a stable, continuous prototype. The bilinear transformation with prewarping affords exact control over the transmission at a selected critical frequency, which must be less than 1/2T. Zero-pole mapping is the simplest method to apply computationally if the desired filter is given

<sup>&</sup>lt;sup>11</sup>See TRIANG in Table E.1.

in terms of its zeros and poles; but with a reasonable computer-aided-design tool, the designer can select that method that best meets the requirements of the design.

# PROBLEMS AND EXERCISES

**4.1** Sketch the zone in the z-plane where poles corresponding to the left half of the *s*-plane will be mapped by the zero-pole mapping technique and the zero-order-hold technique.

- **4.2**Show that <math>(4.15) is true.
- 4.3 a) The following transfer function is a lead network designed to add about  $60^{\circ}$  phase lead at  $\omega_1 = 3$  rad:

$$H(s)=\frac{s+1}{0.1s+1}.$$

For each of the following design methods compute and plot in the z-plane the pole and zero locations and compute the amount of phase lead given by the equivalent network at  $z_1 = e^{j\omega_1 T}$  if T = 0.25 sec and the design is via:

- i) Forward rectangular rule
- ii) Backward rectangular rule
- iii) Bilinear rule
- iv) Bilinear with prewarping (Use  $\omega_1$  as the warping frequency.)
- v) Zero-pole mapping
- vi) Zero-order-hold equivalent
- vii) Triangular-hold equivalent
- b) Plot over the frequency range  $\omega_l = 0.1 \rightarrow \omega_h = 100$  the amplitude and phase Bode plots for each of the above equivalents.
- 4.4 a) The following transfer function is a lag network designed to introduce gain attenuation of a factor of 10 (20 dB) at  $\omega_1 = 3$ :

$$H(s) = \frac{10s+1}{100s+1}.$$

For each of the following design methods, compute and plot on the z-plane the zero-pole patterns of the resulting discrete equivalents and give the gain attenuation at  $z_1 = e^{j\omega_1 T}$ . Let T = 0.25 sec.

- i) Forward rectangular rule
- ii) Backward rectangular rule
- iii) Bilinear rule

- iv) Bilinear with prewarping (Use  $\omega_1 = 3$  radians as the warping frequency.)
- v) Zero-pole mapping
- vi) Zero-order-hold equivalent
- vii) Triangle-hold equivalent
- b) For each case computed, plot the Bode amplitude and phase curves over the range  $\omega_l = 0.01 \rightarrow \omega_h = 10$  rad.