7

DIGITAL SIMULATION OF CONTINUOUS-TIME SYSTEMS

7.1 PRELIMINARY CONSIDERATIONS: SAMPLING AND THE Z-TRANSFORM

Up to this point, the models and their associated signals have been assumed to be characterized in continuous time: there is an infinitesimal difference in time between one time point and the next time point. Ordinary or partial differential equations have been the mathematical representations that we have employed to characterize the dynamics of models of physiological systems, and these operate in the continuous-time domain. Analytical solutions exist for a large class of continuous-time differential equations, but for the rest the only viable path is to employ a numerical method of solution. However, to do so requires a conversion of the problem from one in continuous time to the equivalent problem in discrete time. The way this is achieved in practice is to sample the continuous-time signal (commonly referred to as the *analog* signal) on a periodic basis. Continuous-time *systems* can also be converted to discrete-time *systems*, and in this chapter, we will demonstrate that this can be accomplished using different methods, each with different ramifications.

In Figure 7.1, we consider from a theoretical perspective what exactly occurs in the transformation of a continuous-time signal $x(t)$ into a discrete-time signal $x_D(n)$. The first part of the transformation involves the multiplication of $x(t)$ by a train $p(t)$ of unit impulses uniformly spaced T time units apart. Thus, $p(t)$ is

Physiological Control Systems: Analysis, Simulation, and Estimation, Second Edition. Michael C.K. Khoo.

Companion website: www.wiley.com/go/khoo/controlsystems2e

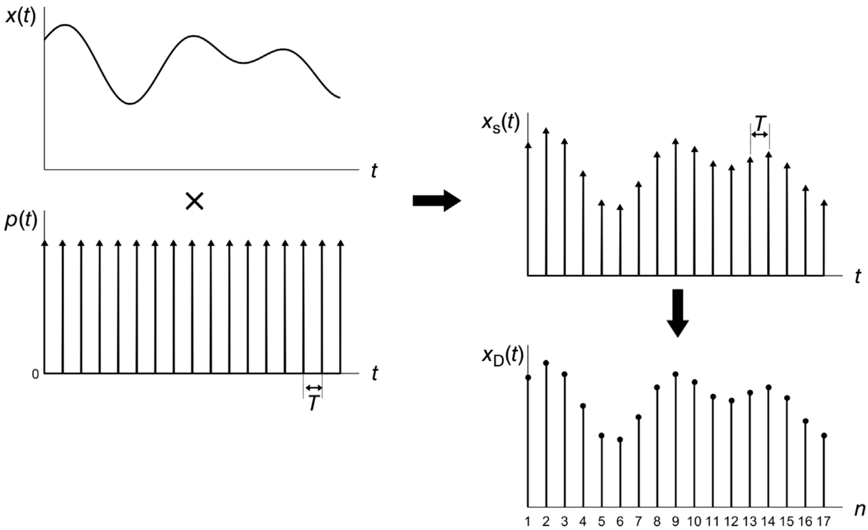

FIGURE 7.1 Schematic illustration of the process of sampling a continuous-time signal and converting it to discrete-time signal.

defined as

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \tag{7.1}$$

where $n = 0, \pm 1, \pm 2, \pm 3, \ldots, \pm\infty$, and $p(t) = 0$ when $t \neq nT$.

The resulting product is

$$x_s(t) = x(t) \cdot p(t) = x(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - nT) \tag{7.2}$$

Thus,

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(nT) \cdot \delta(t - nT) \tag{7.3}$$

Since

$$\delta(t - nT) = 0, \quad \text{when} \quad t \neq nT \tag{7.4}$$

at the time points at which $x(t)$ is sampled, we have

$$x_D(n) = x(nT) \tag{7.5}$$

Note that, as a discrete-time signal, $x_D(n)$ is not defined between consecutive values of n.

Now, take the two-sided Laplace transform of Equation 7.3:

$$X_s(s) = \int_{-\infty}^{\infty} e^{-st} x_s(t)dt = \int_{-\infty}^{\infty} e^{-st} \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT)dt \tag{7.6a}$$

$$= \sum_{n=-\infty}^{\infty} x(nT) \int_{-\infty}^{\infty} e^{-st}\delta(t - nT)\, dt \tag{7.6b}$$

$$= \sum_{n=-\infty}^{\infty} x(nT)e^{-snT} \tag{7.6c}$$

If we define the following equality

$$z = e^{sT} \tag{7.7}$$

and use Equation 7.5, the right-hand side of Equation 7.6c can be converted into a form that contains the variable z. We can then define the following function of z:

$$X_D(z) = \sum_{n=-\infty}^{\infty} x_D(n)z^{-n} \tag{7.8}$$

Equation 7.8 yields a mapping between the discrete-time signal $x(n)$ and the corresponding transformed quantity in the complex z-domain, $X_D(z)$. This "mapping" is called the *z-transform.*

The similarity in form between Equations 7.6c and 7.8 indicates that there is a one-to-one mapping between Laplace transform of $x_s(t)$ and the z-transform of $x_D(n)$. It can also be demonstrated that as T goes to zero, $X_D(z)$ converges to $X(s)$.

The utility of the z-transform for solving difference equations in discrete-time systems parallels that of the Laplace transform for solving differential equations in continuous time. A very simple result that is useful to keep in mind when employing the z-transform is the "delay theorem":

$$\sum_{n=-\infty}^{\infty} x_D(n - m)z^{-n} = z^{-m} \sum_{n=-\infty}^{\infty} x_D(n)z^{-n} = z^{-m}X_D(z) \tag{7.9}$$

We will employ this result frequently in the following sections when we convert difference equations into transfer functions.

7.2 METHODS FOR CONTINUOUS-TIME TO DISCRETE-TIME CONVERSION

In this section, we examine four ways of converting a continuous-time linear system to a discrete-time linear system. In order to keep our focus on the conceptual aspects of these four methods, we will base our considerations on a highly simplified linear system: the lung mechanics model displayed in Figure 4.1. We will assume further that the fluid inertance effects are negligible, and thus the inductance element L will be equal to zero. Let $P_A = y$ and $P_{ao} = x$. Then, from Equation 4.3, we have

$$\tau \frac{dy}{dt} + y = x \tag{7.10}$$

where $\tau = RC$. We showed in Equation 4.7 that the transfer function with $x(P_{ao})$ as input and $y(P_A)$ as output (for the open-loop configuration of the model) is given by

$$H(s) = \frac{Y(s)}{X(s)} = \frac{1}{\tau s + 1} \tag{7.11}$$

The corresponding impulse response for this model is

$$h(t) = \frac{1}{\tau} e^{-t/\tau} \tag{7.12}$$

7.2.1 Impulse Invariance

In the impulse invariance method, the impulse response is sampled at a uniform interval of T time units, and thus the resulting sampled discrete-time impulse response is

$$h_D(n) = \frac{1}{\tau} e^{-nT/\tau}, \quad n \geq 0 \tag{7.13a}$$

$$\text{and} \quad h_D(n) = 0, \quad n < 0 \tag{7.13b}$$

Note that Equation 7.13b holds because $h_D(n)$ is causal.

Taking the z-transform of $h_D(n)$, we obtain

$$H_D(z) = \sum_{n=-\infty}^{n=\infty} h_D(n) z^{-n} \tag{7.14a}$$

But because of Equation 7.13b,

$$H_D(z) = \sum_{n=0}^{n=\infty} h_D(n) z^{-n} \tag{7.14b}$$

Substituting for $h_D(n)$ in Equation 7.14b, we get

$$H_D(z) = \frac{1}{\tau}\sum_{n=0}^{n=\infty} e^{-nT/\tau}\, z^{-n} = \frac{1}{\tau}\sum_{n=0}^{n=\infty}\left(e^{-T/\tau}z^{-1}\right)^n \tag{7.15}$$

However, note the equality:

$$\sum_{n=0}^{\infty} \propto^n = \frac{1}{1-\propto} \tag{7.16}$$

Using Equation 7.16 in Equation 7.15 yields

$$H_D(z) = \frac{1/\tau}{1-e^{-T/\tau}z^{-1}} \tag{7.17}$$

We can use Equation 7.17 to derive the equivalent difference equation relating the output $y(n)$ to the input $x(n)$ by recognizing that $H_D(z)$ is by definition $Y(z)/X(z)$. Then, substituting into Equation 7.17 and rearranging terms, we get

$$\left(1-e^{-T/\tau}z^{-1}\right)Y(z) = \frac{X(z)}{\tau} \tag{7.18}$$

Taking the inverse z-transform of both sides of Equation 7.18, we obtain

$$y(n) = e^{-T/\tau}y(n-1) + \frac{x(n)}{\tau} \tag{7.19}$$

Equation 7.19 gives the solution to the equivalent discrete-time equivalent of Equation 7.10 for any type of input. A hallmark of this solution is that it is recursive in nature: At each time point, y depends on its own past value, as well as the values of the input $x(n)$. Thus, note that if the input is a unit impulse, that is,

$$x(0) = 1 \quad \text{and} \quad x(n) = 0, \quad \text{for} \quad n > 0 \tag{7.20}$$

then Equation 7.19 yields the impulse response:

$$y(n) = h_D(n) = e^{-T/\tau}h_D(n-1) = e^{-2T/\tau}\, h_D(n-2) = \cdots = \frac{1}{\tau}e^{-nT/\tau} \tag{7.21}$$

which is the expression for the impulse response that we started with in Equation 7.13a.

7.2.2 Forward Difference

In this method, we make use of the following numerical approximation for dy/dt that becomes more and more exact as T tends toward zero:

$$\frac{dy}{dt} = \frac{y(n+1)-y(n)}{T} \tag{7.22}$$

Thus, substituting this into Equation 7.10 yields

$$\tau \frac{y(n+1) - y(n)}{T} + y(n) = x(n) \tag{7.23}$$

Rearranging terms in Equation 7.23, we obtain

$$y(n+1) = \left(1 - \frac{T}{\tau}\right) y(n) + \left(\frac{T}{\tau}\right) x(n) \tag{7.24a}$$

or equivalently,

$$y(n) = \left(1 - \frac{T}{\tau}\right) y(n-1) + \left(\frac{T}{\tau}\right) x(n-1) \tag{7.24b}$$

Equation 7.24b provides a somewhat different solution in discrete time to Equation 7.10 than we had found using the impulse invariance method (Equation 7.19).

Using Equation 7.20 in Equation 7.24b allows us to derive the discrete-time impulse response of this system:

$$h_D(n) = 0, \quad n = 0 \tag{7.25a}$$

$$h_D(n) = \left(1 - \frac{T}{\tau}\right)^{n-1} \frac{T}{\tau}, \quad n > 0 \tag{7.25b}$$

The corresponding discrete-time transfer function can be derived by taking the z-transform of Equation 7.24b:

$$Y(z) = \left(1 - \frac{T}{\tau}\right) z^{-1} Y(z) + \left(\frac{T}{\tau}\right) z^{-1} X(z) \tag{7.26a}$$

Rearranging terms, Equation 7.26a yields the z-transform for the discrete-time system in question:

$$H_D(z) \equiv \frac{Y(z)}{X(z)} = \frac{(T/\tau)\, z^{-1}}{1 - (1 - (T/\tau))z^{-1}} \tag{7.26b}$$

7.2.3 Backward Difference

Here, the following numerical approximation for dy/dt, based on the difference between the current time point and the previous time point, is used:

$$\frac{dy}{dt} = \frac{y(n) - y(n-1)}{T} \tag{7.27}$$

Substituting into Equation 7.10, we have

$$\tau\frac{y(n)-y(n-1)}{T}+y(n)=x(n) \tag{7.28}$$

Rearranging terms, we obtain

$$y(n)=\frac{1}{1+(T/\tau)}y(n-1)+\frac{T/\tau}{1+(T/\tau)}x(n) \tag{7.29}$$

As in the previous cases, we obtain the impulse response by setting $x(n)$ to be equal to the unit impulse (Equation 7.20):

$$h_{\mathrm{D}}(n)=\left(1+\frac{T}{\tau}\right)^{-n}\left(\frac{T/\tau}{1+(T/\tau)}\right),\quad n\geq 0 \tag{7.30}$$

The transfer function for the backward difference system can be derived from Equation 7.29 by taking the z-transform of Equation 7.29:

$$Y(z)=\frac{1}{1+(T/\tau)}z^{-1}Y(z)+\frac{T/\tau}{1+(T/\tau)}X(z) \tag{7.31a}$$

Rearranging terms in Equation 7.31a yields the transfer function:

$$H_{\mathrm{D}}(z)\equiv\frac{Y(z)}{X(z)}=\left(\frac{T/\tau}{1+(T/\tau)}\right)\left(1-\frac{1}{1+(T/\tau)}z^{-1}\right)^{-1} \tag{7.31b}$$

7.2.4 Bilinear Transformation

The bilinear transformation is best known through the following mapping between the s terms in the continuous-time transfer function $H(s)$ and the z terms in the discrete-time transfer function $H(z)$:

$$s=\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} \tag{7.32a}$$

The following expression characterizes the same mapping, but expressing z as a function of s:

$$z=\frac{1+(T/2)s}{1-(T/2)s} \tag{7.32b}$$

One disadvantage of the bilinear transformation is that we need to have the expression representing the transfer function $H(s)$ before the conversion to discrete time can be performed.

For a more intuitive interpretation of this transformation, consider the inverse of Equation 7.32a:

$$\frac{1}{s} = \frac{T}{2}\frac{1+z^{-1}}{1-z^{-1}} \tag{7.32c}$$

Denoting the right-hand side of Equation 7.32c as $H(z)$, and remembering that it is by definition equal to $Y(z)/X(z)$:

$$\frac{Y(z)}{X(z)} = \frac{T}{2}\frac{1+z^{-1}}{1-z^{-1}} \tag{7.33a}$$

Rearranging terms in Equation 7.33a, we obtain

$$Y(z) - z^{-1}Y(z) = \frac{T}{2}\left(X(z) + z^{-1}X(z)\right) \tag{7.33b}$$

Taking the inverse z-transform of Equation 7.33b yields

$$y(n) = y(n-1) + \frac{T}{2}(x(n) + x(n-1)) \tag{7.34}$$

Equation 7.34 provides a useful, practical interpretation of Equation 7.32c. What it says is that the bilinear transformation converts the integration operation in continuous time (represented by $1/s$) into the equivalent operation of numerical integration in discrete time by employing the "trapezoidal rule."

We now turn back to deriving the discrete-time equivalent of the first-order continuous-time transfer function given in Equation 7.11. Starting with Equation 7.11 and applying the transformation defined in Equation 7.32a, we obtain

$$H_{\mathrm{D}}(z) = \frac{1}{1 + (2\tau/T)\left[(1-z^{-1})/(1+z^{-1})\right]} \tag{7.35}$$

From Equation 7.35, it is easy to show that the corresponding finite difference equation is

$$y(n) = \frac{1 - (T/2\tau)}{1 + (T/2\tau)}y(n-1) + \frac{T/2\tau}{1 + (T/2\tau)}(x(n) + x(n-1)) \tag{7.36}$$

The corresponding discrete-time impulse response can be derived from Equation 7.36 by setting $x(n)$ to be equal to the unit impulse (Equation 7.20):

$$h_{\mathrm{D}}(n) = \frac{T/2\tau}{1 + (T/2\tau)}, \quad n = 0 \tag{7.37a}$$

$$h_{\mathrm{D}}(n) = \frac{T/\tau}{(1 + (T/2\tau))^2}, \quad n = 1 \tag{7.37b}$$

$$h_D(n) = \left(\frac{1-(T/2\tau)}{1+(T/2\tau)}\right)^{n-1}\left(\frac{T/\tau}{1+(T/2\tau)^2}\right), \quad n > 1 \tag{7.37c}$$

7.3 SAMPLING

In the previous section, we showed that a given continuous-time system can be converted into more than one equivalent discrete-time systems, depending on the method employed to perform the analog-to-digital transformation. Another important parameter in this process is the rate at which the sampling of the continuous-time signal is carried out. It should be quite intuitive that, with a very slow sampling rate, one could miss much of the dynamics of a particular signal. In the example discussed above, different values of the ratio T/τ could lead to discrete-time equivalents with very different system dynamics. In addition, there is another fundamental phenomenon that poses its own challenges, if certain constraints are not kept – and this is the problem of *aliasing* that arises from employing sampling rates that are too low, relative to the dynamics of the continuous-time system.

Consider a continuous-time impulse response $h(t)$. The frequency response of this system is given by the Fourier transform of $h(t)$:

$$H(\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t}dt \tag{7.38}$$

Conversely, the inverse Fourier transform of $H(\omega)$ yields the impulse response:

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)e^{j\omega t}d\omega \tag{7.39}$$

If $h(t)$ is sampled at uniform intervals of T, the values of $h(t)$ at those points would be

$$h(nT) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)e^{jn\omega T}d\omega \tag{7.40}$$

Equation 7.40 can be reformulated in a somewhat different way for use later. We make the following change in variables:

$$\phi = \omega T \tag{7.41a}$$

From Equation 7.41a, we obtain

$$d\omega = \frac{1}{T}d\phi \tag{7.41b}$$

Then, Equation 7.40 can be rewritten as

$$
\begin{aligned}
h(nT) &= \frac{1}{2\pi T}\int_{-\infty}^{\infty} H(\phi)e^{jn\phi}\,d\phi \\
&= \frac{1}{2\pi T}\sum_{m=-\infty}^{\infty}\int_{(2m-1)\pi}^{(2m+1)\pi} H(\phi)e^{jn\phi}\,d\phi
\end{aligned}
\tag{7.42a}
$$

Define

$$\Omega \equiv \phi - 2\pi m \tag{7.43}$$

Substituting into Equation 7.42a yields

$$
\begin{aligned}
h(nT) &= \frac{1}{2\pi T}\sum_{m=-\infty}^{\infty}\int_{-\pi}^{\pi} H(\Omega + 2\pi m)e^{jn\phi}\,d\Omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{m=-\infty}^{\infty}\frac{H(\Omega + 2\pi m)}{T}e^{jn\phi}\,d\Omega
\end{aligned}
\tag{7.42b}
$$

To better appreciate how aliasing affects the sampling process, we first return to the definition of the z-transform in Equation 7.8. Recall that the mapping between the s-domain and the z-domain is given by Equation 7.7. If we are interested in determining how the frequency response of a continuous-time system translates to the frequency response of its equivalent discrete-time system, what we would do is to evaluate the z-transform of the discrete-time system along the contour of the unit circle in the z-domain, that is,

$$z = e^{-j\Omega n} \tag{7.44}$$

Thus, the z-transform of the discrete-time impulse response becomes

$$H_{\mathrm{d}}(\Omega) = \sum_{n=-\infty}^{\infty} h_{\mathrm{D}}(nT)e^{-j\Omega n} \tag{7.45}$$

Note that Equation 7.45 says that the frequency response of the discrete-time system is given by the *discrete-time Fourier transform* of its impulse response. Conversely, we can represent the discrete-time impulse response as the inverse of the Fourier transform of its frequency response:

$$h_{\mathrm{d}}(nT) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_{\mathrm{D}}(\Omega)e^{j\Omega n}\,d\Omega \tag{7.46}$$

Since $h(nT)$ in Equation 7.42b is equal to $h_D(nT)$ in Equation 7.46, we can equate the term within the integral on the right-hand-side of Equation 7.42b to the corresponding term in Equation 7.46:

$$H_d(\Omega) = \sum_{m=-\infty}^{\infty} \frac{H(\Omega + 2\pi m)}{T} \tag{7.47}$$

From Equations 7.41a and 7.43, we see that

$$\omega = \frac{\Omega + 2\pi m}{T} \tag{7.48}$$

Substituting Equation 7.48 into Equation 7.47, we obtain the following result:

$$H_d(\Omega) = \sum_{m=-\infty}^{\infty} \frac{H(\omega - (2\pi m/T))}{T} \tag{7.49}$$

Equation 7.49 is highly significant in that it shows the fundamental relationship between the frequency response of a continuous-time system and the corresponding frequency response of its discrete-time equivalent derived by sampling the impulse response of the continuous-time system. This is best understood by presenting the concept in graphical form, as displayed in Figure 7.2. Figure 7.2a shows the frequency response (magnitude) plot of the continuous-time system, with the abscissa representing angular frequency ω. Since the (absolute) sampling frequency is $1/T$, where T is the sampling interval, the *angular* sampling frequency ω_s is related to T in the following way:

$$\omega_s = \frac{2\pi}{T} \tag{7.50}$$

Figure 7.2b displays the frequency response (magnitude) of the corresponding discrete-time system. Equation 7.49 shows that this frequency response is a reduced-amplitude version of the original frequency response. In addition, the frequency response of the continuous-time system is duplicated an infinite number of times and centered around multiples of $2\pi m$, where $m = 0, \pm 1, \pm 2$, and so on. Figure 7.2 highlights another important detail that appears in Equation 7.49: That the magnitude of the frequency response of the continuous-time system $H(\omega)$ (Figure 7.2a) is scaled by the factor T in the frequency response of the corresponding discrete-time system $H_D(\Omega)$ (Figure 7.2b). The scaled duplicates of the original frequency response are known as *aliases*. The relationship between $H(\omega)$ and $H_D(\Omega)$ is best understood if we recall from Equation 7.7 that the sampling process is equivalent to mapping the imaginary axis ($j\omega$) of the s-plane into the unit circle of the z-plane, that is,

$$e^{j\Omega} = e^{j\omega T} \tag{7.51}$$

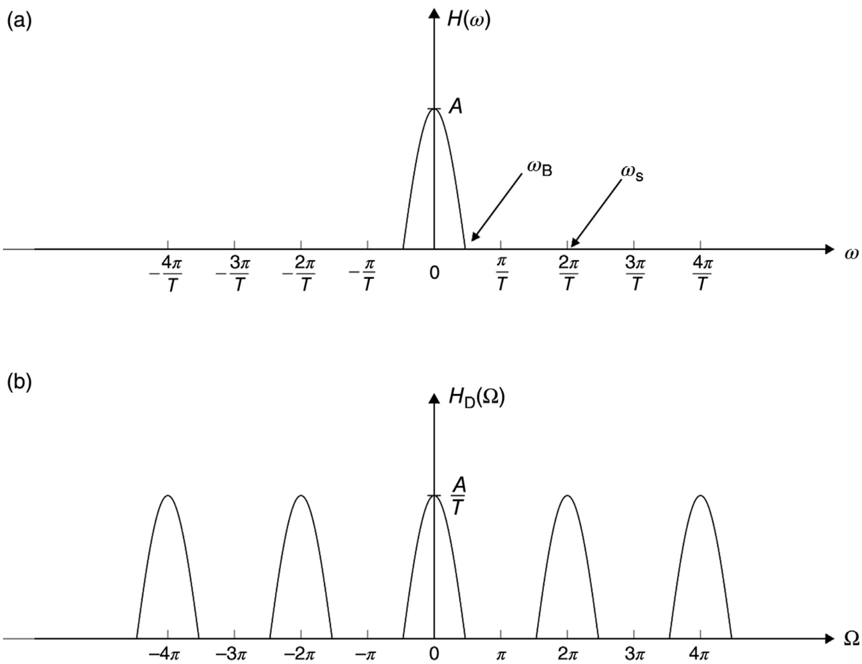

FIGURE 7.2 (a) Frequency response of a continuous-time system. (b) Frequency response of the discrete-time equivalent of the above continuous-time system derived using impulse invariance. ω_B = highest frequency associated with dynamics of the continuous-time system; ω_s = sampling frequency.

As illustrated in Figure 7.3, the segment of the $j\omega$ axis from $\omega = -\pi/T$ to $\omega = \pi/T$ gets mapped to the unit circle of the z-plane from $\Omega = -\pi$ to $\Omega = \pi$ in the anticlockwise direction. What about the segment of $j\omega$ from $\omega = \pi/T$ to $\omega = 3\pi/T$ in the s-plane? One can surmise from Figures 7.2 and 7.3 that this next segment is mapped into the z-plane as another anticlockwise wrap around the unit circle, from $\Omega = -\pi$ to $\Omega = \pi$. Similarly, each "strip" of length $2\pi/T$ of the $j\omega$ axis gets wrapped around the unit circle in the z-plane.

Based on the considerations illustrated in Figures 7.2 and 7.3, it is not too difficult to understand why $H_D(\Omega)$ contains multiple aliases of $H(\omega)$. However, it is important to note that each alias of $H_D(\Omega)$ takes the exact form of $H(\omega)$ under certain constraints. In Figure 7.2, $H(\omega)$ is shown to have a bandwidth (highest frequency) of ω_B, and ω_B is less than π/T. Since the (angular) sampling frequency ω_s is equal to $2\pi/T$,

$$\omega_B < \frac{\omega_s}{2} \tag{7.52}$$

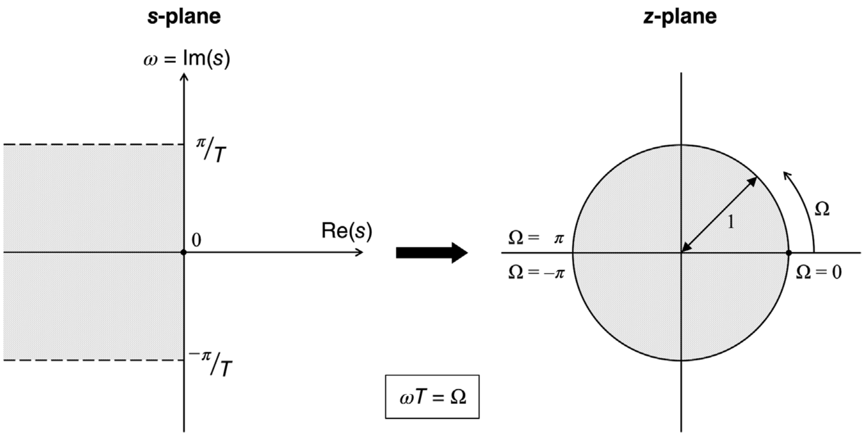

FIGURE 7.3 Schematic illustration of how the s-plane (associated with continuous-time system) maps to the z-plane (associated with discrete-time system), according to the relationship $z = e^{sT}$.

Figure 7.4 shows an example when the condition specified in Equation 7.52 does not hold. Here, because $\omega_B > \omega_s/2$, the ends of the frequency responses of the main transfer function and its aliases run into one another, causing distortion in the regions where there is overlap. When this happens, parts of the system response with frequencies higher than $\omega_s/2$ appear as components in frequencies lower than $\omega_s/2$. This phenomenon is known as *aliasing*. $\omega_s/2$ is also known as the *Nyquist frequency* or *folding frequency*. Equation 7.52 represents the concise version of the Nyquist–Shannon sampling theorem, namely, that a continuous-time signal can be fully reconstructed from its discrete-time equivalent only when the sampling frequency is greater than or equal to twice the highest frequency component of original signal. A corollary of this theorem is that if one samples a periodic phenomenon at the primary frequency of the process, then the aliasing effect will make the dynamic phenomenon appear static. This is the principle by which the stroboscopic effect works.

7.4 DIGITAL SIMULATION: STABILITY AND PERFORMANCE CONSIDERATIONS

In this section, we compare how well the various methods of converting continuous-time systems to discrete-time systems work when they are implemented in MATLAB. We use the example of the first-order system discussed in Section 7.3. Assuming the associated time constant τ to be equal to 1 s, we will consider how the dynamics of the equivalent discrete-time systems derived from the four methods of discrete-to-continuous-time conversion compare with the dynamics of the original continuous-time system. The MATLAB program

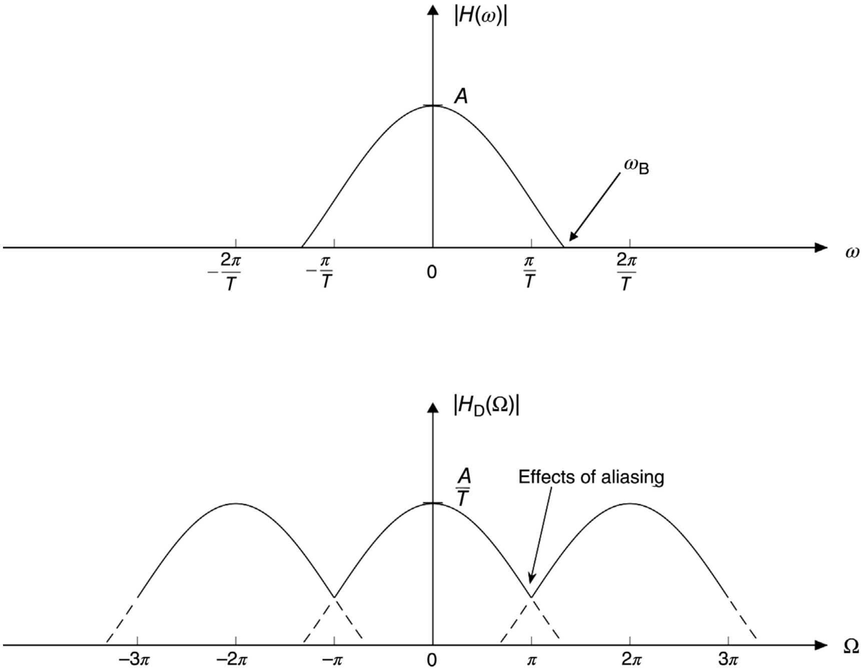

FIGURE 7.4 Aliasing occurs when the bandwidth or the highest frequency of the continuous-time system is greater than half the sampling frequency, that is, $\omega_B > \omega_s/2$. This is equivalent to $\omega_s < 2\omega_B$, thus violating the sampling theorem.

`CT2DTsys_impresp.m` implements the four methods of conversion discussed in the last section:

```
%% Continuous time
h = 1/tau * exp(-tc/tau);

%% Discrete time
x = zeros(size(t));
x(t==0) = 1; %input = unit impulse x(n=0) = 1
n0 = find(t==0); %n=0

%Method 1: impulse invariance
yii = nan(size(t));
yii(t<0) = 0; %y(n<0)
for nn=n0:length(t)
    yii(nn) = exp(-T/tau) * yii(nn-1) + x(nn)/tau; %y(n>0)
end

%Method 2: forward difference
yfd = nan(size(t));
yfd(t<0) = 0; %y(n<0)
```

```
for nn=n0:length(t)
    yfd(nn) = (1 - T/tau)*yfd(nn-1) + (T/tau)*x(nn-1);
end
yfd = yfd/T;   % scaling factor

%Method 3: backward difference
ybd = nan(size(t));
ybd(t<0) = 0; %y(n<0)
Ky = 1/(1 + T/tau); %scaling factor of y
Kx = (T/tau)/(1 + T/tau); %scaling factor of x
for nn=n0:length(t)
    ybd(nn) = Ky*ybd(nn-1) + Kx*x(nn);
end
ybd = ybd/T;   % scaling factor

%Method 4: bilinear transformation
ybt = nan(size(t));
ybt(t<0) = 0; %y(n<0)
Ky = (1 - T/(2*tau)) / (1 + T/(2*tau)); %scaling factor of y
Kx = (T/(2*tau)) / (1 + T/(2*tau)); %scaling factor of x
for nn=n0:length(t)
    ybt(nn) = Ky*ybt(nn-1) + Kx*(x(nn) + x(nn-1));
end
ybt = ybt/T;  % scaling factor
```

As displayed in Figure 7.5a and b, the thick black curve represents the impulse response of the continuous-time system, that is, $h(t)$ in Equation. 7.12. The impulse invariance method consists of simply sampling $h(t)$ at uniform time intervals of T; the mathematical representation is given in Equation 7.13a. Figure 7.5 displays the sampled points as closed black circles that lie on the trajectory of $h(t)$ at two different sampling intervals: $T = 0.1$ s (Figure 7.5a) and $T = 1$ s (Figure 7.5b). This is clearly the reason why this method is known as "impulse invariance." With the forward difference method, the impulse response of the discrete-time equivalent is described by Equations 7.25a and 7.25b. As shown in Figure 7.5 (upright triangles), the peak of the impulse response is delayed by one point. In the case for $T = 0.1$, since the impulse response begins with magnitude zero at time zero, the discrepancy between the discrete-time and continuous-time impulse responses is largest before $t = 0.5$ s, but both responses converge subsequently. However, when $T = 1$, the impulse response of the discrete-time system generated using forward difference oscillates between values of -2 and 2. Thus, clearly, with relatively large T (with respect to τ), the stable continuous-time system gets converted into an unstable discrete-time equivalent. On the other hand, with the backward difference method (inverted triangles, Figure 7.5), the impulse response of the discrete-time system converges toward the trajectory of $h(t)$, regardless of whether T is 0.1 or 1 s, following an initial period of discrepancy. Similarly, the impulse response of the discrete-time equivalent obtained by using the bilinear transformation (open squares, Figure 7.5) converges toward $h(t)$ after the second

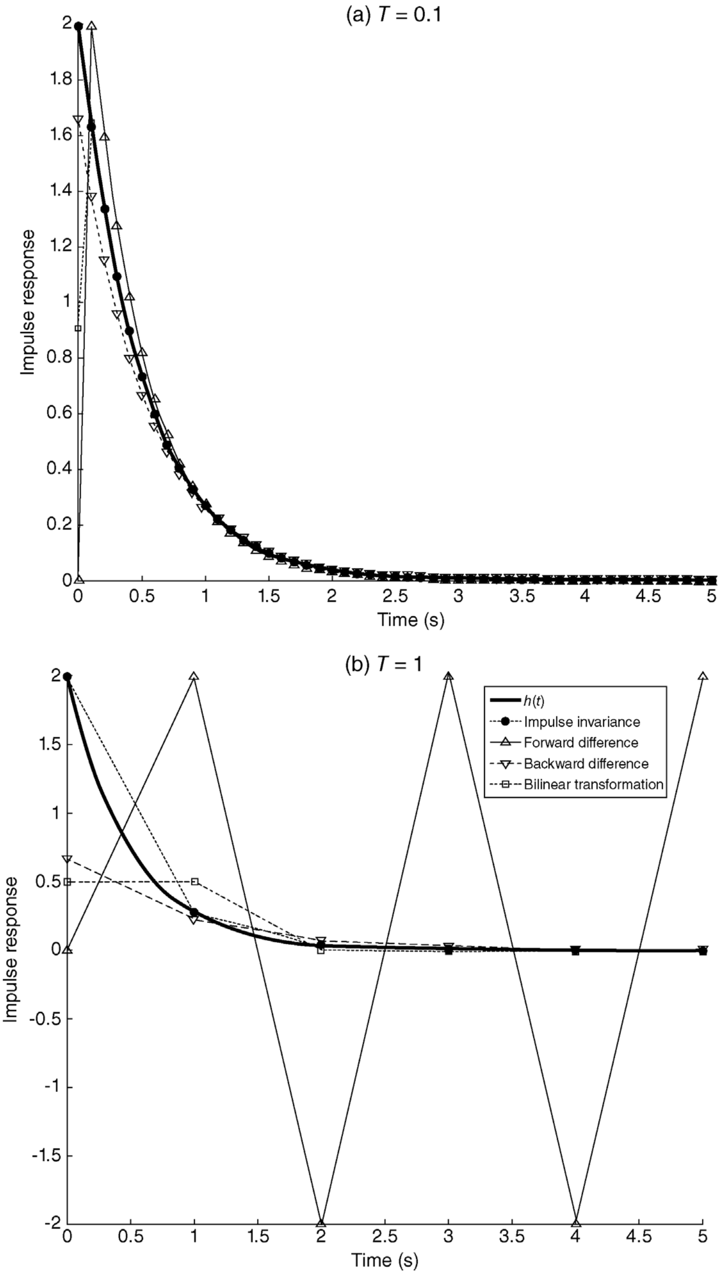

FIGURE 7.5 Impulse responses of discrete-time systems derived from a first-order low-pass continuous-time system using various methods of CT–DT conversion: impulse invariance (closed circles), forward difference (upright triangles), backward difference (inverted triangles), and bilinear transformation (squares), using time step (T) of (A) 0.1 s and (B) 1 s.

point. Note that, in the code for `CT2DTsys_impresp.m` (displayed above), a scaling factor of $1/T$ is applied to the solutions corresponding to the forward difference, backward difference, and bilinear transformation systems in order to make the magnitudes of the discrete-time impulse responses comparable to that of the continuous-time system. This scaling factor stems from the fundamental difference between continuous-time systems and discrete-time systems. In continuous time, the "impulse" takes the form of an infinitely high and infinitely thin "spike," but the total area under the spike is one. In discrete time, the "impulse" simply takes on the value of 1 at time zero. While the discrete-time system generated using impulse invariance has an impulse response function whose values fall on the impulse response of the continuous-time system, its step response needs to be scaled appropriately to match the step response of the continuous-time system.

The relationship between τ, which reflects the dynamics of the continuous-time system, and the sampling interval T used in developing the discrete-time equivalent is simply another manifestation of the relationship between the system bandwidth ω_B and the sampling frequency ω_s, as we had discussed in Section 7.3. When the sampling frequency is less than twice the bandwidth of the continuous-time system in question, aliasing occurs. Equivalently, when the ratio of T to τ becomes too large, aliasing introduces "distortion" into the dynamics of the discrete-time system vis-à-vis the original continuous-time system. This is the reason why the impulse responses of the discrete-time systems become progressively more different from that of the parent continuous-time system as T increases. On the other hand, the unstable behavior of the discrete-time system generated via the forward difference method with large T is derived from a different source. Recall, from Chapter 6, that the poles of a stable continuous-time system are always located on the left-hand side of the s-plane, that is, the real parts of the poles must be negative. Now, consider Figure 7.3 that shows how the s-plane gets mapped into the z-plane. Notice that the left-hand side of the s-plane maps into the area within the unit circle in the z-plane (shaded regions in Figure 7.3). Thus, discrete-time equivalent of a continuous-time system will be stable as long as the poles of the discrete-time system fall within the unit circle. Now, consider Equation. 7.26b, the transfer function corresponding to the discrete-time system derived using the forward difference method. This transfer function can be rewritten as

$$H_D(z) = \frac{T/\tau}{z - (1 - (T/\tau\,))} \tag{7.26c}$$

Thus, the pole at $s = -1/\tau$ in the continuous-time system gets mapped into a pole at $z = 1 - T/\tau$ in the discrete-time equivalent derived using forward difference method. Note that, when T ranges between zero and 2τ, the pole of the discrete-time system falls within the unit circle (i.e., $-1 < z < 1$). But when $T > 2\tau$, this pole will lie outside the unit circle – This is when the discrete-time equivalent of the stable continuous-time system becomes unstable. Similar considerations can be applied to the backward difference and bilinear transformation methods – But in these cases, the corresponding discrete-time system always remains stable.

7.5 PHYSIOLOGICAL APPLICATION: THE INTEGRAL PULSE FREQUENCY MODULATION MODEL

A major motivation for converting a model containing continuous-time systems and signals into a representation in which these systems and signals are now expressed in discrete time is that this allows for more convenient estimation of the model parameters, especially when the measurements employed for estimation are collected on a sample-by-sample basis. This is essentially what happens anyway since analog signals have to be digitized before being acquired on any computer. However, when sampling frequency is very high relative to the dynamics of the system under study, we can still employ continuous-time models (e.g., in the form of differential equations) but use a wide plethora of numerical integration techniques to solve these equations. But there are many instances in which the physiological variables under study occur naturally on a sample-by-sample basis. The obvious examples are cardiac variables, such as heart period and stroke volume, both of which can be quantified on a beat-to-beat basis. Respiratory variables can be expressed on a per-breath basis. Physiological oscillations are so ubiquitous that it is not unusual to quantify the underlying time base in units of “cycles.” Since arterial blood pressure fluctuates between systolic and diastolic levels within each cardiac cycle, one can define new descriptors such as the cycle-averaged blood pressure, systolic pressure, and diastolic pressure on a beat-by-beat basis.

Neural signals are another excellent example of the kind of model where continuous-time inputs can yield outputs that may be approximated as discrete “spikes.” In this case, the underlying “drive” may be continuous, but the output is in the form of a train of neural impulses. Generally, when the “drive” is high, the neural system would depolarize more rapidly and generate an action potential more quickly – As such, a high drive would produce a high rate of neuronal firing. The integral pulse frequency modulation (IPFM) model, introduced by Bayly (1968), has been employed in many theoretical studies of neuronal dynamics. Figure 7.6 displays a schematic diagram that highlights how the IPFM model works. The following equations specify the operations of each of the modules in

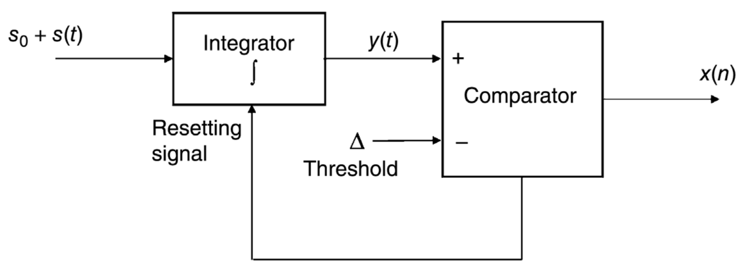

FIGURE 7.6 Schematic diagram of the integral pulse frequency modulation (IPFM) model. (Adapted from Chiu and Kao (2001).)

the IPFM model:

$$y(t) = \int_{t_n}^{t} [(s_0 + s(t))dt] \qquad (7.53)$$

where $t_n \le t \le t_{n+1}$, and

$$y(t_{n+1}) - y(t_n) = \Delta \qquad (7.54)$$

Note, in Equation 7.53, that s_0 represents the intrinsic drive, while $s(t)$ represents the modulated component of the drive. The square brackets [. . .] operate by disallowing any negative values to occur; if the argument becomes negative, the square brackets will function as a thresholding operation, setting everything to zero if the argument within the square brackets goes negative. Δ represents a threshold that determines the intrinsic frequency of the generated pulses when $s(t)$ is equal to zero. The output of the integral in Equation 7.53 is constantly compared to the selected threshold Δ, and once the difference between y at time t_{n+1} and y at time t_n equals Δ, a spike is generated at the output of the comparator module. At that same instant, a signal is sent to the integrator to reset and start integrating the input again.

A SIMULINK implementation of the IPFM model (`IPFM.slx`) is displayed in Figure 7.7. Figure 7.7a shows the IPFM model as a subsystem that receives the neural drive input and outputs the corresponding response in the form of a spike train. Figure 7.7b shows the internal workings of the IPFM. The "neural drive" takes the form of a continuous-time signal with mean value s_0 and fluctuating component $s(t)$. It has units of impulses (or cycles) per second – Hence, it represents the instantaneous neural firing frequency. This continuous-time signal is first integrated and compared with the threshold Δ. When the integral has risen to the point at which it attains the value of the threshold, the model generates a "spike." This "process" may be thought of as being analogous to the depolarization of the nerve cell membrane prior to the point at which an action potential is generated. The SIMULINK implementation shown here assumes a threshold value of 1, and we can consider this example as a model of how the totality of autonomic input to the heart generates the surge of electrical activity that triggers ventricular contraction (observable via the electrocardiogram as the "R-wave"). The "neural drive" in this case would be the instantaneous heart rate. In the SIMULINK implementation, the instantaneous heart rate (in cycles per second) is integrated continuously until the integral attains the value of 1. At this point, the "hit crossing" block generates a unit impulse (spike), the integrator is reset to zero, and integration of the input (instantaneous heart rate) resumes, starting from zero. The total duration over which each cycle of integration takes place equals the heart period for that beat. In the SIMULINK implementation, the tracking of the "R-to-R interval" is taken care of through the use of the second integrator, which integrates a constant

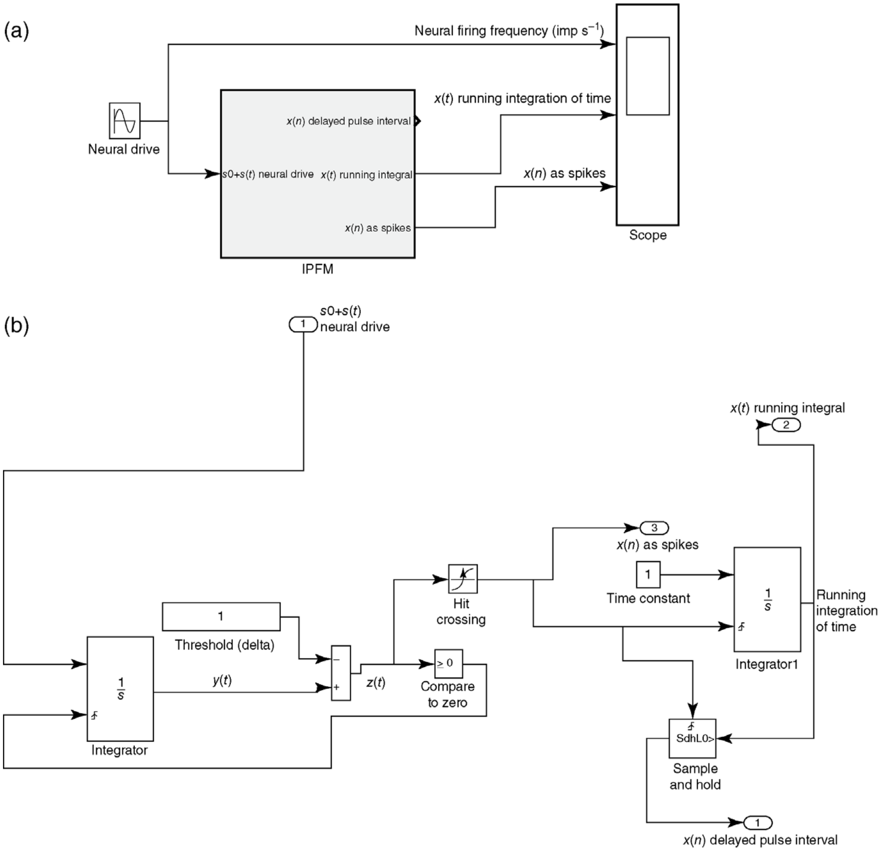

FIGURE 7.7 SIMULINK implementation of IPFM model. (a) Overall model showing IPFM subsystem with input (neural drive or neural firing frequency) and outputs. (b) SIMULINK structure of IPFM mechanism (see text for explanation).

input value of 1 until it is reset to zero by the next "spike" issued by the "hit crossing" block:

$$x(n) = \int_{t_n}^{t_{n+1}} 1\, dt \tag{7.55}$$

Figure 7.8 displays the results of running `IPFM.slx` with a constant level of cardiac autonomic input equivalent to a heart rate of 0.5 beat s^{-1} (or 30 beats min^{-1}) for the first 30 s and a different constant level of 1 beat s^{-1} (or 60 beats min^{-1}) for the following 30 s (part (a)). Figure 7.8b shows, for each beat, the running time count (output of the second integrator) that occurs in parallel with the integration of the cardiac autonomic input signal (accomplished through the first integrator in

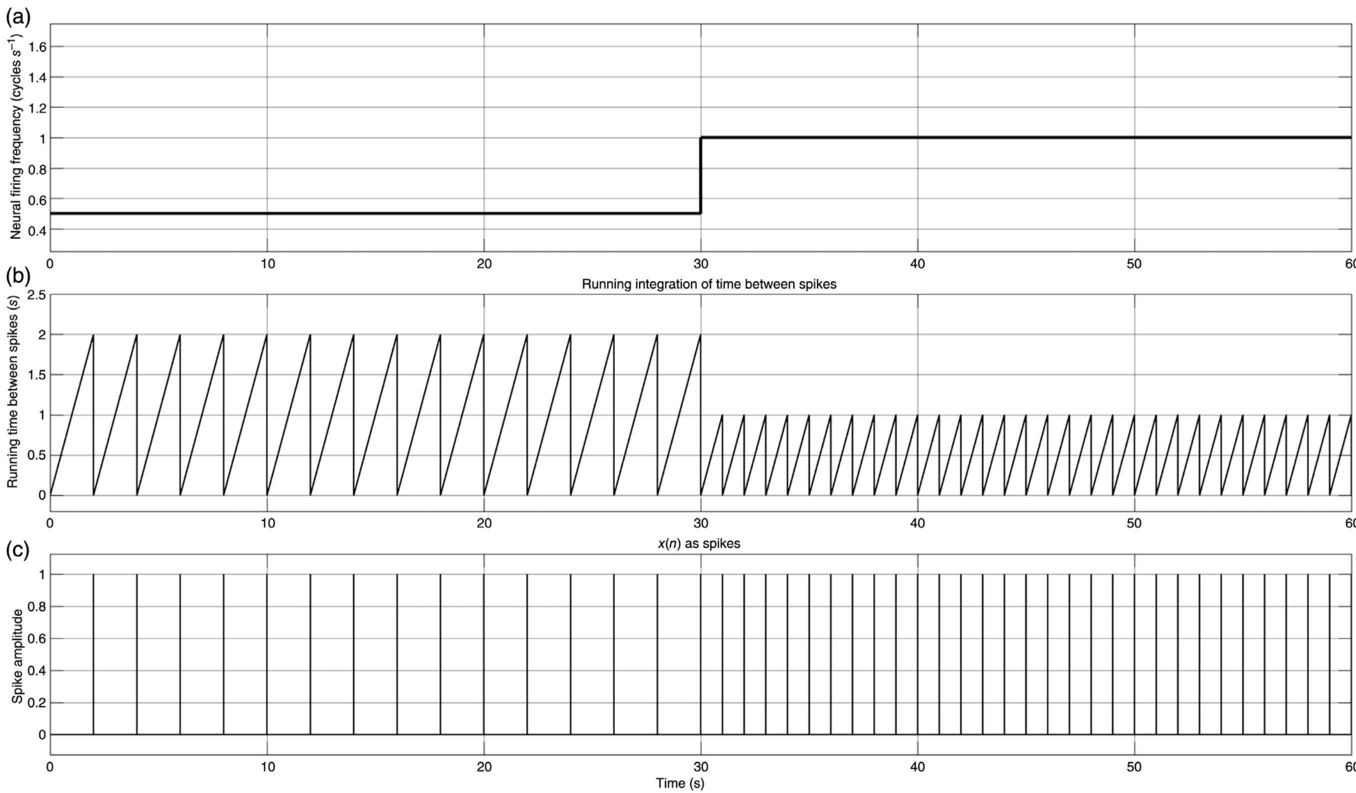

FIGURE 7.8 Simulation of IPFM model with fixed neural drives.

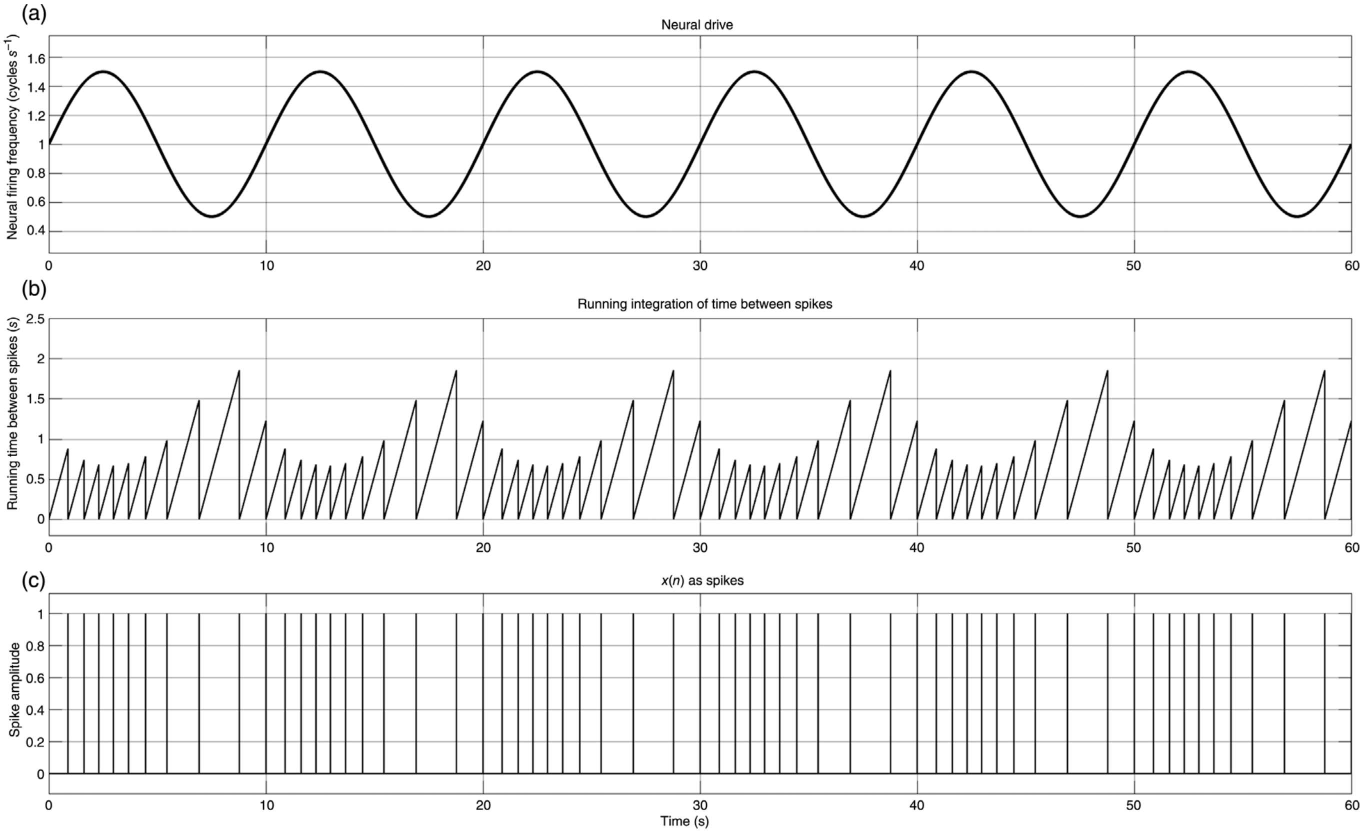

FIGURE 7.9 Simulation of IPFM model with sinusoidally modulated neural drive.

Figure 7.7). Recall that the first integrator resets to zero once the integral attains the value of 1. The second integrator resets to zero simultaneously, but in this case, the highest value of the integral achieved before it is reset yields the duration of time elapsed since the previous beat. In this example, this time interval equals 2 s in the first half of the simulation and 1 s in the second half. Figure 7.8c displays the main output of the IPFM model, that is, the "spikes" of unit amplitude that are generated with periodicities consistent with the input neural drive. In this first half of this simulation, the heart rate is 0.5 beat s^{-1}, equivalent to a heart period of 2 s, whereas in the second half, the heart rate of 1 beat s^{-1} yields a heart period of 1 s. Figure 7.9 shows another simulation, but this time, the autonomic input to the heart fluctuates sinusoidally with an amplitude of 0.5 beat s^{-1} around a mean level of 1 beat s^{-1}. This input represents an oscillatory drive that should make instantaneous heart rate vary between 0.5 and 1.5 beats s^{-1}. When it is sent through the IPFM, the output is a train of spikes (beats) that varies in interbeat interval between 0.6 and 1.8 s (Figure 7.9b and c).

PROBLEMS

P7.1. Consider a saline-filled catheter that has been inserted into the brachial artery of a patient so that the proximal tip of the catheter is exposed to blood flowing through the artery at pressure P_a. The distal tip of the catheter is connected to a pressure transducer. The transducer works by means of an internal thin diaphragm that deflects by an amount proportional to the difference between the pressure in the transducer chamber (P_m) and the ambient pressure (which we will consider to be equal to zero). This arrangement is displayed in Figure P7.1. Under static conditions, P_m should be exactly equal to P_a. However, this will not be true if P_a varies dynamically. How much P_m differs

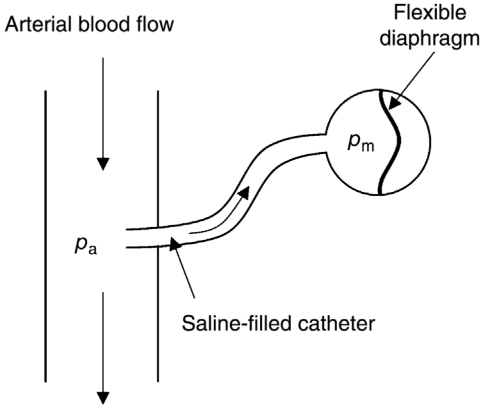

FIGURE P7.1 Schematic illustration of catheter–transducer system for measuring arterial pressure.

from P_a at any given time would depend on the response characteristics of the catheter–transducer system. If the mechanical properties of the transducer diaphragm and the dynamics of fluid motion in the catheter are known (based on prior testing), it is possible to employ a simple model to determine how the true arterial pressure signal is likely to be distorted dynamically by the measurement process.

(a) Derive the simplest linear lumped parameter model of the catheter–transducer arrangement that relates P_m to P_a. Include in the model the effects of (i) resistance R to fluid motion along the catheter, (ii) inertance L due to fluid acceleration along the catheter, and (iii) compliance C of the transducer diaphragm. We will consider the saline inside the catheter to be incompressible and the catheter wall to be nondistensible.

(b) Use the forward difference (Euler) method for converting the continuous-time model above into a discrete-time model. With the resulting difference equation, compute the responses in P_m of the discrete-time model to a unit step in P_a when the time step (sampling interval) $T = 0.1$ s and when $T = 2$ s. Assume the following values for the model parameters: $R = 0.05$, $L = 0.1$, and $C = 10$.

(c) Using SIMULINK, determine the response of the *continuous-time system* to a unit step, and display this alongside the two responses obtained in part (b).

(d) Determine an expression for the transfer function of the discrete-time system ($P_m(z)/P_a(z)$). By examining the locations of the poles of this system on the z-plane, explain why the stability properties of the two discrete-time representations ($T = 0.1$ versus $T = 2$) are different.

P7.2. In an experiment on humans, the ventilatory response to a single-breath challenge of CO_2 was measured, that is, during one breath, the inhaled CO_2 concentration was changed abruptly from 0 to 10% against a background mixture of air. Subsequently, the same subjects were exposed to the same CO_2 challenge, except that this was performed against a background mixture of hypoxic gas. In all subjects, the following model was found to provide an adequate fit to the data:

$$y(n) = ay(n-1) + bx(n-M)$$

where x and y represent changes from the mean levels of inhaled CO_2 concentration and ventilation, respectively. n represents the current breath, and M represents the delay (in number of breaths) between exposure to CO_2 and the change in ventilation that follows. Note that the measurements were made on a breath-by-breath basis, and therefore as a first approximation, they may be considered samples of an underlying continuous-time process that were acquired with a sampling interval equal to the subject's average breath period (T).

(a) Derive an expression for the transfer function (i.e., $H(s) = Y(s)/X(s)$) of the equivalent continuous-time model, assuming impulse invariance. Show clearly how a, b, M, and T are related to the parameters of $H(s)$.

(b) In one subject, suppose the following values were estimated from the data:

Normoxia: $a = 0.684$, $b = 0.059$, $M = 3$, $T = 3.3$ s
Hypoxia: $a = 0.624$, $b = 0.257$, $M = 2$, $T = 2.6$ s

How has hypoxia affected the steady-state gain and time constant of the underlying continuous-time model for the CO_2 ventilatory response for this particular subject?

P7.3. Consider the following discrete-time linear system with transfer function (z-domain) as given below:

$$H(z) = \frac{z}{z - 0.5}$$

(a) Derive the corresponding finite difference equation for this system that will enable you to determine how the output $y(n)$ would respond to the input $x(n)$ in the (discrete) time domain.

(b) Sketch as accurately as possible the response of the above system to a unit impulse.

(c) The values tabulated below represent the output of the system to an unknown input signal. Assume both output and input were sampled at 1 Hz. Determine the corresponding values of the input signal at the times displayed in the table below:

Time, t (s)	0	1	2	3	4	5	6	7
$y(t)$	−4	3	8	−2	6	−7	−5	1

P7.4. Combine the IPFM model, implemented in SIMULINK as `IPFM.slx`, with the model of circulatory control introduced in Section 5.5.1 (`rsa.slx`), so that the extended “RSA” model will generate simulated “R-waves,” similar to the ECG spikes that accompany each heart beat. Then, using the successive intervals between adjacent R-waves, produce plots of heart period variability similar to those displayed in Figure 5.15. Generate such plots for the “normal,” “+atropine,” and “+propranolol” conditions. By resampling these R-to-R interval time series with a uniform sampling interval of 0.5 s and applying `rsa_tf.m` to the resulting time series, determine if the corresponding frequency responses are similar to the plots displayed in Figure 5.16.

BIBLIOGRAPHY

Bayly, E.J. Spectral analysis of pulse frequency modulation in the nervous system. *IEEE Trans. Biomed. Eng.* BME-15: 257–265, 1968.

Chiu, H-W, and T. Kao. A mathematical model for autonomic control of heart rate variation. *IEEE Eng. Med. Biol.* 20: 69–76, 2001.

Oppenheim, A.V., and R.W. Schafer. *Discrete-Time Signal Processing*, 3rd edition, Pearson, New York, 2009.

Oppenheim, A.V., and A.S. Willsky. *Signals and Systems*, 2nd edition, Pearson, New York, 1996.