7

DISCRETE-TIME SIGNALS AND LINEAR DIFFERENCE EQUATIONS

7.0 Introduction

Thus far our study of signals and systems has emphasized voltages and currents as signals, and electric circuits as systems. To be sure, we have pointed out various analogies to simple mechanical systems, and we have explored in at least an introductory way systems composed of larger blocks than elementary R's, L's, and C's. But, mathematically, all of our signals have been specified as functions of a continuous variable, t, and almost all of our systems have been described by sets of linear, finite-order, total differential equations with constant coefficients (or equivalently by system functions that are rational functions of the complex frequency s).

With this chapter we shall begin the process of extending our mathematical models to larger classes of both signals and systems. Such extensions are interesting in part because they will permit us to analyze and design a wider range of practical systems and devices. But an equally important goal is to learn how to brush aside certain less fundamental characteristics of our system models so that we may concentrate on the deep, transcendent significance of their linearity and time-invariance.

Specifically, we shall consider in the next few chapters systems in which the signals are indexed *sequences* rather than functions of continuous time. We shall identify such sequences by $x[n]$, $y[n]$, etc., where the square brackets indicate that the enclosed index variable, called *discrete time*, takes on only integral values: $\ldots, -2, -1, 0, 1, 2, \ldots$. Discrete-time (DT) signals, like continuous-time (CT) signals, may be defined in many ways—by bar diagrams as in Figure 7.0–1, by formulas such as $x[n] = 2^n$, by tables of values, or by combinations of these.

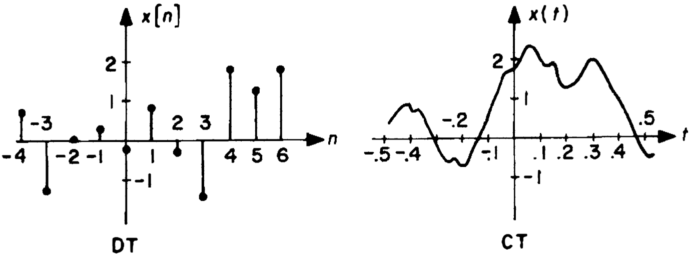


Figure 7.0–1. Comparison of DT and CT signals.

Many systems inherently operate in discrete time. Some examples are certain banking situations ("regular monthly payments"), medical therapy regimes ("two pills every four hours"), econometric models utilizing periodically compiled indices, and evolutionary models characterizing population changes from generation to generation. Moreover, a variety of regular structures in space, rather than time—such as cascaded networks, tapped delay lines, diffraction gratings, surface-acoustic-wave (SAW) filters, and phased-array antennas—lead to similar mathematical descriptions. In other cases, DT signals are constructed by periodic sampling of a CT signal. If the sampling is done sufficiently rapidly and the signal is sufficiently smooth, the loss in information can be small, as we shall demonstrate in Chapter 14. Motion picture and television images are examples of this kind—the image in two space dimensions and continuous time is sampled every 30–40 msec to yield a sequence of frames.

Sometimes the reason for the transformation from continuous to discrete time is to permit time-sharing of an expensive communications or data-processing facility among a number of users. Examples range from a ward nurse measuring patient temperatures sequentially to telemetry systems transmitting interleaved samples of a variety of data from scattered oil wells or weather stations or interplanetary space probes. But the most common reason today for replacing a CT signal by an indexed sequence of numbers is to make it possible for the signal processing to be carried out by digital computers or similar special-purpose logical devices. Examples include such disparate areas as speech analysis and synthesis; radio, radar, infrared, and x-ray astronomy; the study of sonar and geophysical signals; the analysis of crystalline and molecular structures; the interpretation of medical signals such as electrocardiograms, CAT scans, and magnetic resonance imaging; and the image-enhancement or pattern-recognition processing of pictures such as satellite or space probe photos, x-ray images, blood smears, or printed materials. Processing by computers requires not only that the signal be discrete-time, but that the numbers representing each sample be rounded off or quantized—a potential additional source of error. In exchange, however, one gains great flexibility and power. For example, once the signals in a complex radar receiver have been sampled and quantized, all further processing involves only logical operations, which are inherently free of the parameter drift, sensitivity, noise, distortion, and alignment problems that often limit the effectiveness of analog devices. Thus digital filters can process extremely low-frequency signals that in analog filters would be hopelessly corrupted by the effects of aging and drift. Moreover, digital computers can accomplish certain tasks, such as the approximate solution of large sets of non-linear differential equations, that are virtually impossible to do in any other way. And a change in the task to be carried out requires only a change in instructions, not a rebuilding of the apparatus. Thus computer simulation is increasingly replacing the "breadboard" stage in complex system design because it is faster, cheaper, and permits more flexible variation of parameters to optimize performance. Indeed, sometimes (as in the design of integrated circuits) a simulated "breadboard" may be more accurate than one constructed with "real" elements such as lumped transistors that are not the ones that will ultimately be used in the actual device.

A natural vehicle for describing a system intended to process or modify discrete-time signals—a *discrete-time system*—is frequently a set of *difference equations*. Difference equations play for DT systems much the same role that differential equations play for CT systems. Indeed, as we shall see, the analysis of linear difference equations reflects in virtually every detail the analysis of linear differential equations. The next few chapters will thus also serve as a review of much of our development to this point. In addition, DT systems are in certain mathematical respects simpler than CT systems. The extension from difference/differential equation systems to general LTI systems is thus easiest if we first carry it out for DT systems, as we shall in Chapter 9.

7.1 Linear Difference Equations

A linear N^{th} order constant-coefficient difference equation relating a DT input $x[n]$ and output $y[n]$ has the form*

$$\sum_{k=0}^{N} a_k y[n+k] = \sum_{\ell=0}^{N} b_\ell x[n+\ell]. \tag{7.1–1}$$

Some of the ways in which such equations can arise are illustrated in the following examples.

Example 7.1–1

A \$50,000 mortgage is to be retired in 30 years by equal monthly payments of p dollars. Interest is charged at 15%/year on the unpaid balance. Let $P[n]$ be the unpaid principal in the mortgage account just after the n^{th} monthly repayment has been made. Then

$$P[n+1] = (1+r)P[n] - p, \qquad n \geq 0 \tag{7.1–2}$$

where $r = 0.15/12 = 0.0125$ is the monthly interest rate. Initially $P[0] = 50,000$, and we seek the value of p such that $P[360] = 0$.

We shall return to this problem in Example 7.3–1, but for the moment notice that (7.1–2) has the form of (7.1–1) with $N = 1$ (first order) and

$$\begin{aligned} y[n] &= P[n] & \qquad x[n] &= p, \quad 0 < n \leq 360 \\ a_0 &= -(1+r) & b_0 &= -1 \\ a_1 &= 1 & b_\ell &= 0 \text{ otherwise} \\ a_k &= 0 \text{ otherwise.} \end{aligned}$$

▶▶▶

*There is no loss of generality in assuming the same number $(N+1)$ of terms on each side in (7.1–1), since we may always set certain of the coefficients to zero.

Example 7.1–2

The numerical integration of differential equations typically involves difference equations as an intermediate step resulting from replacing derivatives by formulas involving differences, such as

$$\dot{x}(t) = \frac{dx(t)}{dt} \approx \frac{x(t+\Delta t) - x(t)}{\Delta t}$$

where Δt is the *step size* , a small time increment that we assume fixed. Thus consider Figure 7.1–1, which is the circuit of Example 1.3–3 with two of the sources set to zero.

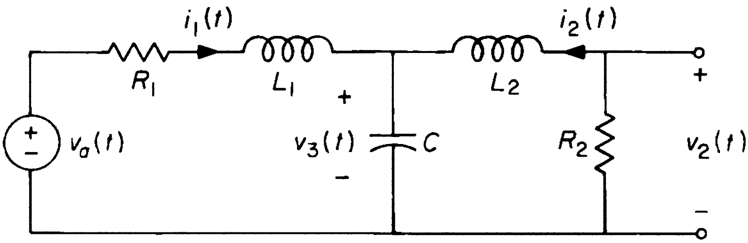


Figure 7.1–1. Circuit for Example 7.1–2.

From Example 1.3–3, the dynamic equations in state form are

$$\begin{aligned}
\frac{di_1(t)}{dt} &= -\frac{R_1}{L_1} i_1(t) - \frac{1}{L_1} v_3(t) + \frac{1}{L_1} v_a(t) \\
\frac{di_2(t)}{dt} &= -\frac{R_2}{L_2} i_2(t) - \frac{1}{L_2} v_3(t) \\
\frac{dv_3(t)}{dt} &= \frac{1}{C} i_1(t) + \frac{1}{C} i_2(t)
\end{aligned}$$

with the output equation

$$v_2(t) = -R_2 i_2(t).$$

Replacing derivatives as above yields

$$\begin{aligned}
i_1(t+\Delta t) &\approx \left(1 - \frac{R_1 \Delta t}{L_1}\right) i_1(t) - \frac{\Delta t}{L_1} v_3(t) + \frac{\Delta t}{L_1} v_a(t) \\
i_2(t+\Delta t) &\approx \left(1 - \frac{R_2 \Delta t}{L_2}\right) i_2(t) - \frac{\Delta t}{L_2} v_3(t) \\
v_3(t+\Delta t) &\approx v_3(t) + \frac{\Delta t}{C} i_1(t) + \frac{\Delta t}{C} i_2(t) \\
v_2(t) &= -R_2 i_2(t).
\end{aligned}$$

Substituting $t = n\Delta t$, $i_1(t) = i_1(n\Delta t) = i_1[n]$, etc., and using the numerical values $R_1 = R_2 = 1\ \text{k}\Omega$, $L_1 = L_2 = 0.1$ H, $C = 0.2\ \mu$F, $\Delta t = 10\ \mu$sec, yields a set of three simultaneous difference equations

$$\begin{aligned} i_1[n+1] &= 0.9i_1[n] - 10^{-4}v_3[n] + 10^{-4}v_a[n] \\ i_2[n+1] &= 0.9i_2[n] - 10^{-4}v_3[n] \\ v_3[n+1] &= v_3[n] + 50i_1[n] + 50i_2[n] \\ v_2[n] &= -10^3 i_2[n]. \end{aligned}$$

We shall return to this example in Example 7.3–2.

▶▶▶

There are many different ways to approximate a set of differential equations by a set of difference equations as above; these correspond to different integration algorithms. The choice in Example 7.1–2 (called the *forward Euler algorithm*) is perhaps the most straightforward, but it is not usually the best in terms of minimizing the approximation error for a given step size. Indeed, for many systems (including this one—see Example 8.3–2) the forward Euler algorithm is *numerically unstable* in that the total accumulated error in the approximate solution grows with time unless the step size is sufficiently small. This difficulty can be overcome with various *implicit* integration algorithms such as the *backward Euler algorithm*, which substitutes $\dot{x}(t+\Delta t) \approx \dfrac{x(t+\Delta t) - x(t)}{\Delta t}$ and yields in the present case the equations

$$\begin{aligned} 1.1\, i_1[n+1] + 10^{-4}\, v_3[n+1] &= i_1[n] + 10^{-4} v_a[n+1] \\ 1.1\, i_2[n+1] + 10^{-4}\, v_3[n+1] &= i_2[n] \\ -50i_1[n+1] - 50i_2[n+1] + v_3[n+1] &= v_3[n] \end{aligned}$$

which must be solved at each iteration because new values of $i_1[n]$, $i_2[n]$, and $v_3[n]$ at $n+1$ depend on one another (this is why the backward Euler algorithm is called "implicit"). Other examples of simple implicit algorithms are given in Problem 7.1. For a comprehensive discussion of the issues involved in selecting among such algorithms, see, for example, L. O. Chua and P.-M. Lin, *Computer-Aided Analysis of Electronic Circuits* (Englewood Cliffs, NJ: Prentice-Hall, 1975).

Example 7.1–3

As an example of a case in which the index variable n corresponds to space rather than time, consider the uniform cascaded ladder network, a section of which is shown in Figure 7.1–2.

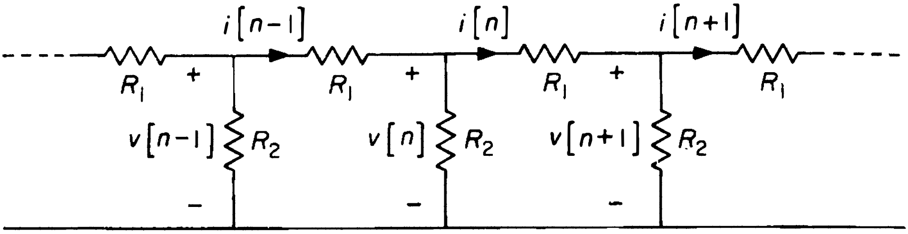


Figure 7.1–2. Uniform ladder network for Example 7.1–3.

The voltages and currents along the ladder satisfy the equations

$$v[n+1] = v[n] - i[n]R_1$$
$$i[n+1] = i[n] - \frac{v[n+1]}{R_2}.$$

The first equation can be solved for $i[n]$ and substituted into the second for $i[n]$ and $i[n+1]$ to yield a single second-order homogeneous difference equation for $v[n]$:

$$R_2 v[n+2] - (2R_2 + R_1)v[n+1] + R_2 v[n] = 0. \qquad (7.1\text{–}3)$$

This example is continued in Example 7.4–1.
▶ ▶ ▶

7.2 Block Diagrams and State Formulations for DT Systems

Circuit and block diagrams provide useful alternatives to differential equations as descriptions of CT systems. Similar diagrams can be devised for the description of DT systems. One such diagram is directly analogous to the integrator-adder-gain class of CT block diagrams introduced in Section 1.4; it employs a DT system called an *accumulator* in place of the integrator. Another kind of DT block diagram—on the whole, a more useful one—uses *delay* elements rather than accumulators as the memory elements. Both types of block diagrams provide significant insights into DT system behavior—particularly the notion of state—and their study yields opportunities for further practice with difference equations.

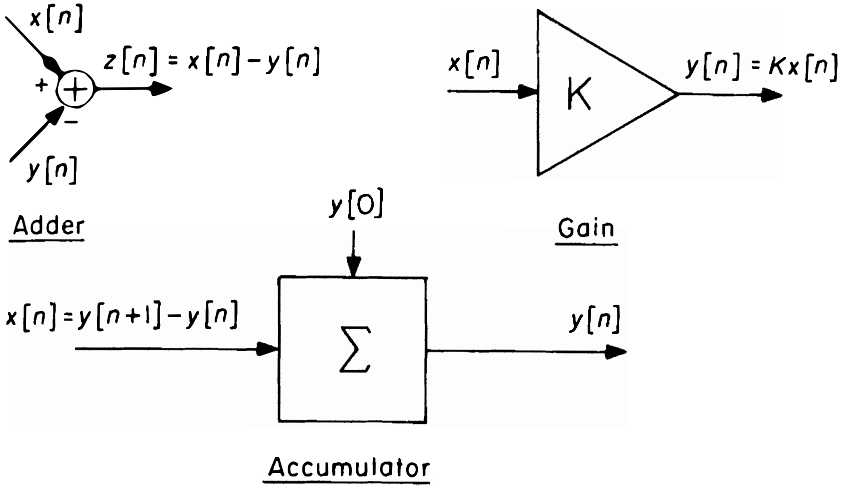


Figure 7.2–1. Some basic discrete-time block-diagram elements.

The elements described in Figure 7.2–1 are those used in what are called accumulator-adder-gain DT block diagrams. The difference equation in this figure defines an accumulator* in the same implicit way one might define an integrator by stating that its input is the derivative of its output. In Example 9.1–1 we shall show that this version of an accumulator can be defined explicitly by

$$y[n] = y[0] + \sum_{m=0}^{n-1} x[m]$$

which clarifies both the name "accumulator" and its relation to a CT integrator. Note that specification of the initial output value $y[0]$ is necessary to yield a unique solution to the accumulator difference equation, which is why it is explicitly shown as a separate input in Figure 7.2–1.

Indeed, the accumulator output $y[n_0]$ at time $n = n_0$ defines the state of the accumulator at that time. That is, knowledge of $y[n_0]$ together with the input $x[n]$ for $n \geq n_0$ determines the output $y[n]$ for $n > n_0$. Hence, if a DT system is described by a block diagram of interconnected accumulators, adders, and gain elements, the current state of the system is defined by the current values of the outputs of the accumulators. To derive the dynamic difference equations for the system in state form, it is sufficient to choose the accumulator outputs as variables and to determine from the block diagram how these variables are combined to yield the input to each accumulator. The procedure parallels exactly the corresponding procedure for CT systems composed of integrators, adders, and gain elements, and is most easily explained through an example.

*As implied in Section 7.1, there are many possible DT analogs of a CT integrator. The selection made here implements the forward Euler algorithm as before. See Problem 8.7.

Example 7.2–1

The CT circuit of Example 7.1–2 led (for the parameter values given in that example) to the differential state equations

$$\frac{di_1(t)}{dt} = -10^4 i_1(t) - 10v_3(t) + 10v_a(t)$$
$$\frac{di_2(t)}{dt} = -10^4 i_2(t) - 10v_3(t)$$
$$\frac{dv_3(t)}{dt} = 5\times10^6 i_1(t) + 5\times10^6 i_2(t)$$

with the output equation

$$v_2(t) = -10^3 i_2(t).$$

This circuit is thus equivalent to the block diagram of Figure 7.2–2. Each of the state equations corresponds directly to the equation describing the adder at the input to the corresponding integrator.

In Example 7.1–2, DT difference equations approximating this system were derived by replacing the derivatives by differences,

$$i_1[n+1] = 0.9i_1[n] - 10^{-4}v_3[n] + 10^{-4}v_a[n]$$
$$i_2[n+1] = 0.9i_2[n] - 10^{-4}v_3[n]$$
$$v_3[n+1] = v_3[n] + 50i_1[n] + 50i_2[n]$$

and

$$v_2[n] = -10^3 i_2[n].$$

These equations also describe the accumulator-adder-gain block diagram shown in Figure 7.2–3. Again each difference equation corresponds directly to the equation describing the output of each adder multiplied by $\Delta t = 10^{-5}$. Not surprisingly, the block diagrams in Figures 7.2–2 and 7.2–3 are identical except that each integrator is replaced by a gain Δt in cascade with an accumulator. The difference equations above are said to be in *state form* because they provide an explicit way to calculate the next values of the *state variables*, $i_1[n+1]$, $i_2[n+1]$, and $v_3[n+1]$, from the present values, $i_1[n]$, $i_2[n]$, and $v_3[n]$, and the system input $v_a[n]$.

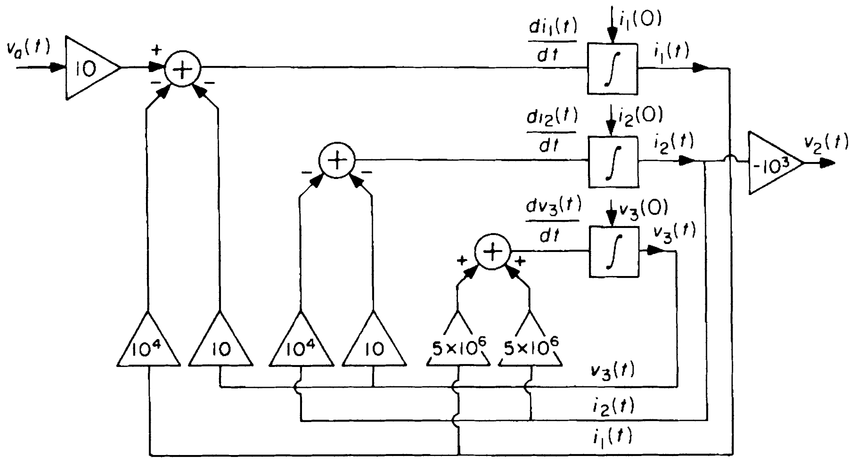


Figure 7.2–2. Block diagram for the circuit of Example 7.1–2.

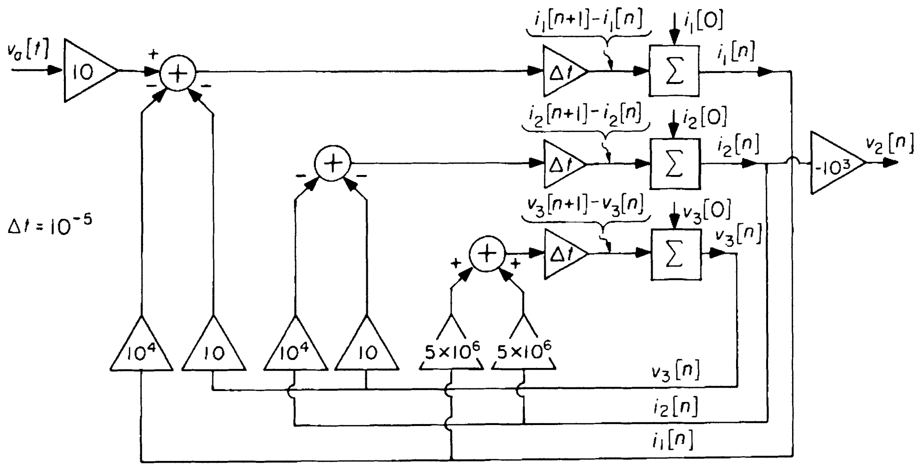


Figure 7.2–3. Block diagram for the difference equations of Example 7.1–2.

▶▶▶

As Example 7.2–1 shows, it is easy to go from an accumulator-adder-gain block diagram to a set of difference equations in state form, choosing as state variables the outputs of the accumulators. The inverse statement is also true: It is generally possible to go from a set of difference equations or a higher-order input-output difference equation to a block diagram in accumulator-adder-gain form.

However, as implied above, an alternative DT block-diagram representation in which delay elements replace accumulators as the memory elements is also general and often much more convenient. The *unit delay element* is represented schematically as in Figure 7.2–4 and is characterized by the property that its present output is equal to its input one unit of time earlier. In a block diagram made up of delay-adder-gain elements the state of the system is characterized by the present values of the delay-element outputs. Although delay elements can be combined with adders and gains in many configurations, the most useful arrangements seem to result if the delay elements are cascaded—the output of one becoming the input to the next—to form what is sometimes called a DT *tapped delay line*. The following example illustrates some of the things one can do with tapped delay lines.

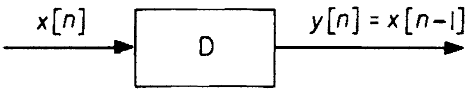


Figure 7.2–4. Unit delay element.

Example 7.2–2

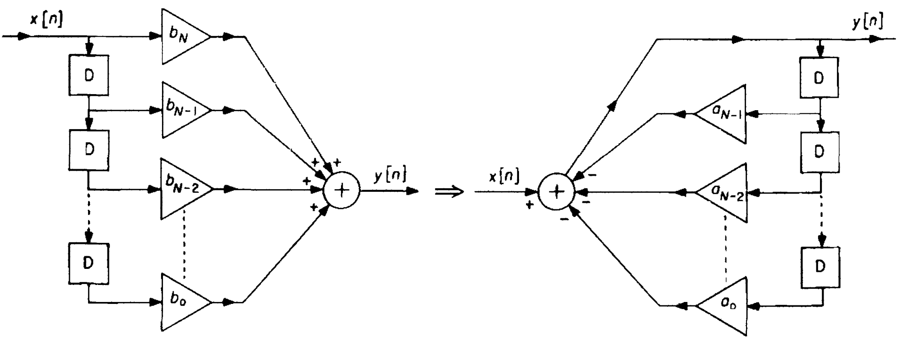


Figure 7.2–5. Feedforward (a) and feedback (b) DT delay-line systems.

The simple feedforward or *transversal* structure of Figure 7.2–5(a) and the feedback or *recursive* arrangement of Figure 7.2–5(b) realize two special cases of the general difference equation (7.1–1). Thus, combining the signals of the adders, it is evident that Figure 7.2–5(a) is characterized by the difference equation

$$y[n] = b_o x[n-N] + \cdots + b_{N-2}x[n-2] + b_{N-1}x[n-1] + b_N x[n] \tag{7.2–1}$$

or, equivalently,

$$y[n+N] = b_0 x[n] + \cdots + b_{N-2}x[n+(N-2)] + b_{N-1}x[n+(N-1)] + b_N x[n+N] \tag{7.2–2}$$

which corresponds to (7.1–1) with $a_N = 1$ and $a_i = 0$, $0 \leq i < N$. Similarly, Figure 7.2–5(b) is characterized by

$$y[n] = -a_{N-1}y[n-1] - a_{N-2}y[n-2] - \cdots - a_0 y[n-N] + x[n] \tag{7.2–3}$$

or, equivalently,

$$y[n+N] + a_{N-1}y[n+(N-1)] + a_{N-2}y[n+(N-2)] + \cdots + a_0 y[n] = x[n+N] \tag{7.2–4}$$

which corresponds to (7.1–1) with $a_N = b_N = 1$ and $b_i = 0$, $0 \leq i < N$. Notice, however, that if the output of the feedforward system of Figure 7.2–5(a) is made the input to the system of Figure 7.2–5(b), that is, if the two systems are cascaded as implied by the double arrow in Figure 7.2–5, then the overall system is characterized by

$$y[n+N] + a_{N-1}y[n+(N-1)] + \cdots + a_0 y[n] = b_0 x[n] + b_1 x[n+1] + \cdots + b_N x[n+N] \tag{7.2–5}$$

which is precisely (7.1–1) except for the (non-constraining) choice $a_N = 1$. Thus the input-output behavior of *any* DT system can be simulated in this way.

An even more interesting block diagram results if the order in which the systems are cascaded in Figure 7.2–5 is reversed, as in Figure 7.2–6. Such a reversal does not change the overall input-output ZSR behavior. (We are not quite in a position to prove this statement easily at this point—see Section 8.3. The analogous result for CT systems was discussed in Section 5.1.) But observe that if we label the signal at the midpoint $w[n]$, then the signals at successive points in each of the delay lines are the same—$w[n-1]$, $w[n-2]$, ..., $w[n-N]$. Hence only one of the two lines is really necessary; an equivalent arrangement is shown in Figure 7.2–7 and is called the *canonical* form because it employs the minimum possible number of delay elements. The representation of the general input-output difference equation (7.1–1) in the form of Figure 7.2–7 is particularly convenient because the coefficients in the difference equation are directly identifiable as corresponding gains in the block diagram. A similar block diagram for CT systems was derived in Problem 1.6.

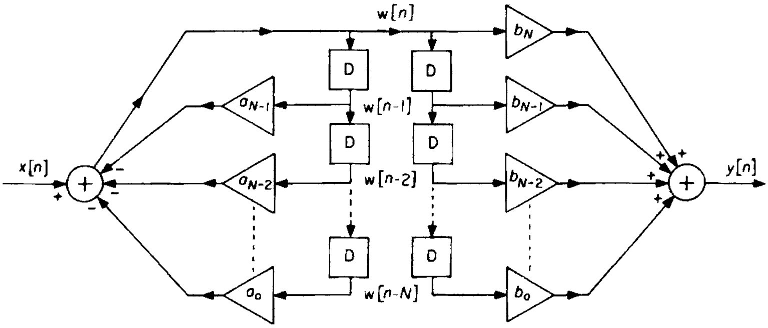


Figure 7.2–6. Cascade of systems of Figure 7.2–5 in reversed order.

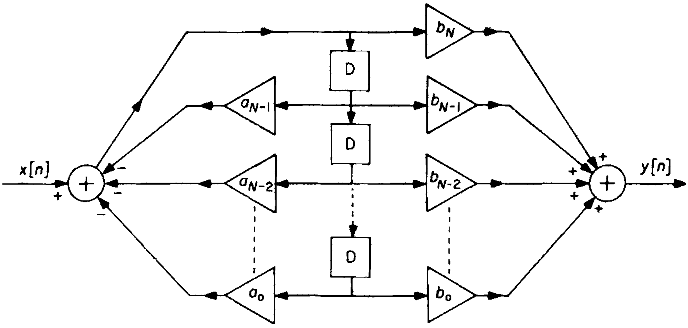


Figure 7.2–7. The canonical form of delay-line realization of the input-output behavior of a general DT LTI system.

Difference equations in state form are readily derived for the system of Figure 7.2–7 by taking the outputs of the delay elements as state variables:

$$\begin{aligned}
\lambda_0[n+1] &= \lambda_1[n] \\
\lambda_1[n+1] &= \lambda_2[n] \\
&\vdots \\
\lambda_{N-2}[n+1] &= \lambda_{N-1}[n] \\
\lambda_{N-1}[n+1] &= -a_{N-1}\lambda_{N-1}[n] - a_{N-2}\lambda_{N-2}[n] - \cdots \\
&\qquad \cdots - a_0\lambda_0[n] + x[n].
\end{aligned}$$

The output equation is then

$$\begin{aligned}
y[n] &= b_N\lambda_{N-1}[n+1] + b_{N-1}\lambda_{N-1}[n] + b_{N-2}\lambda_{N-2}[n] + \cdots + b_0\lambda_0[n] \\
&= (b_{N-1} - b_N a_{N-1})\lambda_{N-1}[n] + (b_{N-2} - b_N a_{N-2})\lambda_{N-2}[n] + \cdots \\
&\qquad \cdots + (b_0 - b_N a_0)\lambda_0[n] + b_N x[n].
\end{aligned}$$

▶ ▶ ▶

The canonical form of Figure 7.2–7 is well-adapted to exploit the capabilities of a class of cheap monolithic MOS devices called *charge-transfer devices* (CTD). Several categories of CTD's are available—such as bucket-brigade devices (BBD) or charge-coupled devices (CCD)—that are broadly similar but differ in details of structure and performance characteristics. These devices are driven by a periodic chain of clock pulses whose rate typically may range from perhaps 100 pulses per second (pps) to 10^7 pps or more. On each clock pulse, the analog voltage on the input capacitor of each of a succession of cascaded MOS stages is transferred to become the voltage on the input capacitor of the next stage. Charge transfer requires part of the interpulse period; during the rest of the period, the voltage remains nearly constant. The voltage at each stage can be sampled (through buffers), weighted, and added to the weighted outputs of other stages as in the canonical system structure (fixed weights can be built into the chip; adjustable weighting requires external potentiometers), or even fed back to add to the voltage of earlier stages (general recursive devices). Hundreds or even thousands of stages can be fabricated on a single chip. Some of the uses of CTD's as filters, delay lines, memory devices, correlators, frequency changers, Fourier transformers, etc., will be discussed in later chapters and problems.

7.3 Direct Solution of Linear Difference Equations

As with linear differential equations, linear difference equations describe the input-output behavior of a DT system only implicitly—they must be "solved" to find the response to a specific input. And again as in the CT case, the difference equation and the input in the semi-infinite interval $0 \leq n < \infty$ yield only a partial description of the output for $n \geq 0$. It is necessary to provide N additional pieces of information for an N^{th}-order system, corresponding to the initial conditions or the initial state in the continuous case. However, unlike the situation with differential equations, the solution to the difference-equation problem is trivial if the information given is the first N values of the output, since it is always possible to rewrite (7.1–1) in the form*

$$y[n+N] = \sum_{\ell=0}^{N} b_\ell x[n+\ell] - \sum_{k=0}^{N-1} a_k y[n+k]. \qquad (7.3\text{–}1)$$

We may thus immediately find $y[N]$ from the known input and the known values of $y[0]$, $y[1]$, …, $y[N-1]$. Iteratively, we find $y[N+1]$ from the input and $y[1]$, $y[2]$, …, $y[N]$, and so on. This technique is illustrated in the following examples.

*Without loss of generality, we have again set $a_N = 1$.

Example 7.3–1

Continuing Example 7.1–1, observe that (7.1–2) already has the form of (7.3–1), that is,

$$P[n+1] = (1+r)P[n] - p \tag{7.3–2}$$

with $P[0] = \$50,000$. That is all the information we need, since the equation is first-order. We readily calculate

$$\begin{aligned} P[1] &= (1+r)P[0] - p \\ P[2] &= (1+r)P[1] - p \\ &= (1+r)^2 P[0] - (1 + (1+r))p \\ P[3] &= (1+r)^3 P[0] - (1 + (1+r) + (1+r)^2)p \end{aligned}$$

from which we deduce the general term

$$P[n] = (1+r)^n P[0] - \frac{(1+r)^n - 1}{r} p, \quad n \geq 0\,. \tag{7.3–3}$$

To write $P[n]$ in closed form we have used the important formula for the partial sum of a geometric series,

$$1 + a + a^2 + \cdots + a^m = \frac{1 - a^{m+1}}{1-a}. \tag{7.3–4}$$

For a further discussion of (7.3–4), see Example 8.1–4.

From (7.3–3) we can directly compute the value of p necessary to pay off this mortgage in 30 years. Setting

$$P[360] = 0$$

gives

$$p = \frac{r(1+r)^{360}}{(1+r)^{360} - 1} P[0]\,.$$

For $r = 0.0125$ (15%/year) and $P[0] = \$50,000$ this becomes

$$p = \$632.22\,.$$

The overall return to the bank on the \$50,000 mortgage is $360p = \$227,599.20$, which illustrates rather clearly why banks make loans.

▶ ▶ ▶

Example 7.3–2

The simultaneous difference equations replacing the differential state equations for the circuit of Example 7.1–2 are already in the form of (7.3–1) and can readily be solved by iteration if $i_1[0]$, $i_2[0]$, and $v_3[0]$ (characterizing the initial state) are known. The results for $i_1[0] = i_2[0] = v_3[0] = 0$ and $v_a[n] = 1$, $n \geq 0$, are shown in Figure 7.3–1 for several values of Δt. Further features of this example are considered in Example 8.3–2.

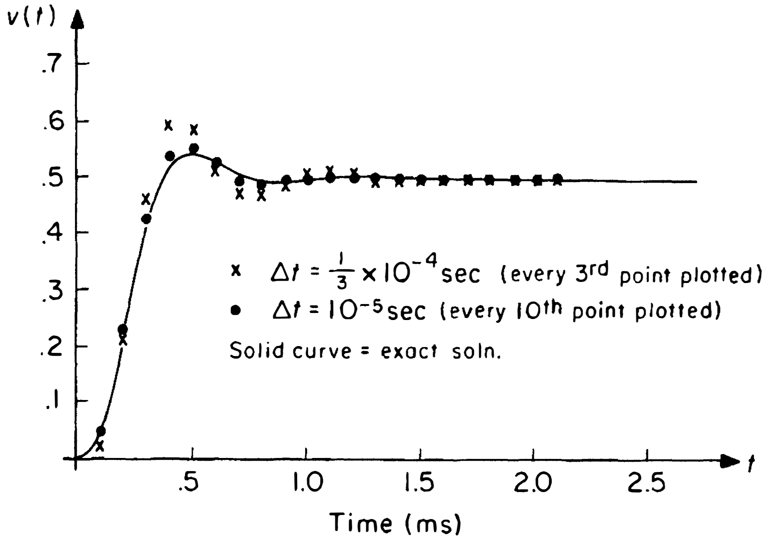


Figure 7.3–1. Solution of the equations of Example 7.1–2 for several values of Δt.

In using difference equations to obtain approximate solutions to differential equations, the choice of sampling interval or step size, Δt, is critical. If Δt is too large, the accuracy of the discrete approximation is poor; if Δt is too small, the number of iterations required to describe the interesting range of the output becomes large. In general, Δt must be significantly smaller than the interval over which the CT state changes substantially. Sometimes it is useful to change Δt during the evolution of the solution to better match the speed at which the state is changing.

▶ ▶ ▶

Unfortunately, the boundary conditions are not always given in a form that is convenient for the direct solution method discussed in this section. And even when they are, identifying a closed-form expression for $y[n]$ is not always easy (although this is not necessarily a serious problem if the equations are being solved by a computer). There are, however, alternative solution schemes that parallel the ZSR and ZIR procedures for CT linear differential equations and that provide much insight as well as workable solutions for many problems. We shall begin to explore these in the next section.

7.4 Zero Input Response

The ladder network of Example 7.1–3 is a typical situation in which the direct method of Section 7.3 fails because the boundary conditions describe a global rather than a local characteristic of the solution.

Example 7.4–1

Continuing Example 7.1–3, suppose we seek the input resistance to, and voltage distribution along, the semi-infinite ladder structure of $1\,\Omega$ resistors shown in Figure 7.4–1.

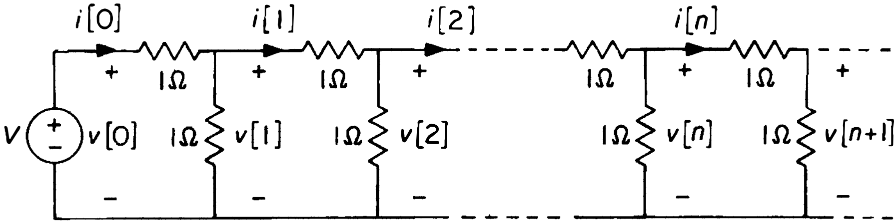


Figure 7.4–1. Resistance ladder network.

In Example 7.1–3 we concluded that $v[n]$ satisfies the homogeneous difference equation

$$v[n+2] - 3v[n+1] + v[n] = 0 \tag{7.4–1}$$

with boundary conditions $v[0] = V$ and $v[n] \to 0$ as $n \to \infty$. Since the difference equation is 2^{nd}-order, whereas only one boundary condition is given for n near zero, the direct-solution method is inapplicable.

Constant-coefficient linear homogeneous difference equations, such as (7.4–1), have solutions of the form

$$v[n] = Az^n \tag{7.4–2}$$

where z is an appropriate (generally complex) number; that is, the homogeneous solutions generally are *discrete-time exponentials*. To show this in the case at hand, substitute (7.4–2) into (7.4–1) to obtain

$$Az^{n+2} - 3Az^{n+1} + Az^n = Az^n(z^2 - 3z + 1) = 0\,.$$

If $v[n] = Az^n \neq 0$, then z must be a root of

$$\left(z^2 - 3z + 1\right) = \left(z - \frac{3+\sqrt{5}}{2}\right)\left(z - \frac{3-\sqrt{5}}{2}\right) = 0$$

or

$$z = \frac{3+\sqrt{5}}{2}, \frac{3-\sqrt{5}}{2}.$$

Since (7.4–1) is linear, the most general solution is the sum of two terms such as (7.4–2):

$$v[n] = A\left(\frac{3-\sqrt{5}}{2}\right)^n + B\left(\frac{3+\sqrt{5}}{2}\right)^n.$$

Since $(3-\sqrt{5})/2 = 0.38 < 1$, whereas $(3+\sqrt{5})/2 = 2.62 > 1$, the second term grows with n and the first dies away. For a semi-infinite resistive ladder driven only at $n = 0$, then, we must have $B = 0$ and $A = V$; the desired solution to our problem is thus

$$v[n] = V\left(\frac{3-\sqrt{5}}{2}\right)^n, \quad n \geq 0.$$

To find the current $i[0]$, and hence the input resistance $V/i[0]$, we observe that

$$V = i[0]\cdot 1 + v[1] = i[0] + V\left(\frac{3-\sqrt{5}}{2}\right)$$

so that*

$$\frac{V}{i[0]} = \frac{2}{\sqrt{5}-1} = \frac{1+\sqrt{5}}{2} = 1.62\,\Omega.$$

▶ ▶ ▶

The lesson of Example 7.4–1 can be readily generalized and summarized as follows. The zero input response (ZIR) of the system characterized by the linear constant-coefficient difference equation

$$\sum_{k=0}^{N} a_k y[n+k] = \sum_{\ell=0}^{N} b_\ell x[n+\ell] \tag{7.4–3}$$

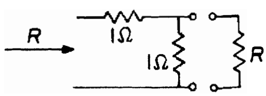
is the solution to the homogeneous equation

$$\sum_{k=0}^{N} a_k y[n+k] = 0. \tag{7.4–4}$$

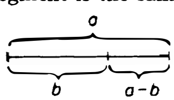
This solution has the form

$$y[n] = \sum_{k=1}^{N} A_k z_k^n = 0 \tag{7.4–5}$$

*The input resistance (if that is all we seek) can be obtained more directly as follows. If R is the input resistance to the infinite ladder, then adding one more section should not change its value, so that R must be the resistance looking into one section terminated in R.



Thus, $R = 1 + \frac{1\times R}{1+R}$, or $R^2 - R - 1 = 0$. The roots of this equation are $R = \frac{1\pm\sqrt{5}}{2}$; the negative root is extraneous in the physical situation being analyzed. It is interesting that the value of R is the same as the "golden section" that divides a line in such a way that the ratio of the whole to the longer segment is the same as the ratio of the longer segment to the shorter.



$$\frac{a}{b} = \frac{b}{a-b} = \frac{1+\sqrt{5}}{2}$$

This same ratio appears in the study of Fibonacci numbers (see Problem 7.2).

where $\{z_k\}$ is the set of N roots of the characteristic equation*

$$\sum_{k=0}^{N} a_k z^k = 0 \tag{7.4–6}$$

and $\{A_k\}$ is a set of N constants chosen to match the initial conditions. Further examples will be given in succeeding chapters.

7.5 Summary

Linear difference equations bear much the same relationship to DT signals and systems as linear differential equations bear to CT signals and systems. In particular, they can be described by block diagrams in which accumulators become the DT equivalents of integrators whose outputs describe the state of the system. A more convenient block diagram for representing the input-output behavior of any LTI DT system uses delay elements instead of accumulators. Difference equations can be solved in basically the same ways as differential equations. Thus the solutions of homogeneous difference equations (zero input responses) are sums of DT exponentials,

$$y[n] = \sum_{k=1}^{N} A_k z_k^n$$

with $\{z_k\}$ equal to the set of roots of the characteristic equation and the set $\{A_k\}$ chosen to match boundary conditions. The forced response to exponential drives can also be found as in the CT case—substitute an assumed solution having the same form as the drive and match coefficients. But, also as in the CT case, it is more convenient to solve driven problems by transform methods. The DT analog of the $\mathcal{L}$- transform is the *Z-transform*, which will be the topic of the next chapter.

*If the characteristic equation (7.4–6) has multiple-order roots, then terms of the form $n^P z_k^n$ must be added to (7.4–5) in precise analogy to the continuous-time case.

EXERCISES FOR CHAPTER 7

Exercise 7.1

A savings bank advertises "deposit \$100 each month for 12 years and we'll pay you \$100 a month forever!" Treat the bank as a DT system where the input $x[n]$ is your monthly payment (positive) or withdrawal (negative) and the output $y[n]$ is the principal in your account just after the n^{th} monthly payment has been received or made. Assume that no withdrawals are made during the first 12 years and that only \$100/month is withdrawn thereafter.

a) Sketch $x[n]$ in accordance with the terms of the advertisement.

b) Describe the general shape of $y[n]$ if \$100/month is precisely the largest payment that can be guaranteed forever given a monthly interest rate r.

c) Find a difference equation relating $y[n]$ and $x[n]$.

d) Determine directly the first few terms of $y[n]$. Infer the general formula

$$y[n] = 100\frac{(1+r)^{n+1}-1}{r}$$

during the first 12 years, $0 \leq n < 144$, and show that it satisfies the difference equation.

e) Show that the yearly interest rate implied by this scheme is approximately 5.8%.

PROBLEMS FOR CHAPTER 7

Problem 7.1

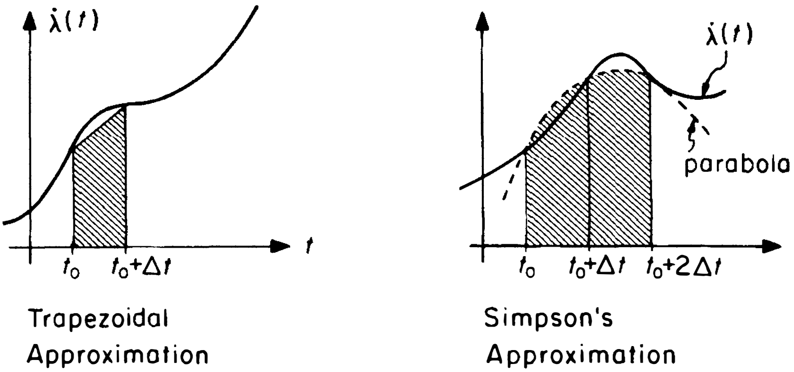
Two other common simple integration algorithms are:

TRAPEZOIDAL RULE:

$$\frac{\lambda(t+\Delta t)-\lambda(t)}{\Delta t} \approx \frac{1}{2}\left[\dot{\lambda}(t)+\dot{\lambda}(t+\Delta t)\right]$$

SIMPSON'S RULE:

$$\frac{\lambda(t+2\Delta t)-\lambda(t)}{2\Delta t} \approx \frac{1}{6}\left[\dot{\lambda}(t)+4\dot{\lambda}(t+\Delta t)+\dot{\lambda}(t+2\Delta t)\right]$$



a) Show that the trapezoidal rule is equivalent to approximating

$$\int_{t_0}^{t_0+\Delta t} \dot{\lambda}(t)\,dt = \lambda(t_0+\Delta t)-\lambda(t_0)$$

as the area of the trapezoid shown in the figure above to the left.

b) Show that Simpson's rule is equivalent to approximating

$$\int_{t_0}^{t_0+2\Delta t} \dot{\lambda}(t)\,dt = \lambda(t_0+2\Delta t)-\lambda(t_0)$$

as the area under the parabola passing through the points $\dot{\lambda}(t_0)$, $\dot{\lambda}(t_0+\Delta t)$, and $\dot{\lambda}(t_0+2\Delta t)$, as shown in the figure above to the right.

c) Use the trapezoidal rule to convert the differential equations of Example 7.1-2 into a set of difference equations. (Simpson's rule is an excellent algorithm for computing the area under a curve, but it leads to an unstable set of difference equations if applied to the approximate solution of differential equations—see Problem 8.7.)

Problem 7.2

a) The first few terms of a sequence of numbers called the *Fibonacci numbers* are

$$\{y[n]\} = \{0, 1, 1, 2, 3, 5, 8, \cdots\}, \quad n \geq 0.$$

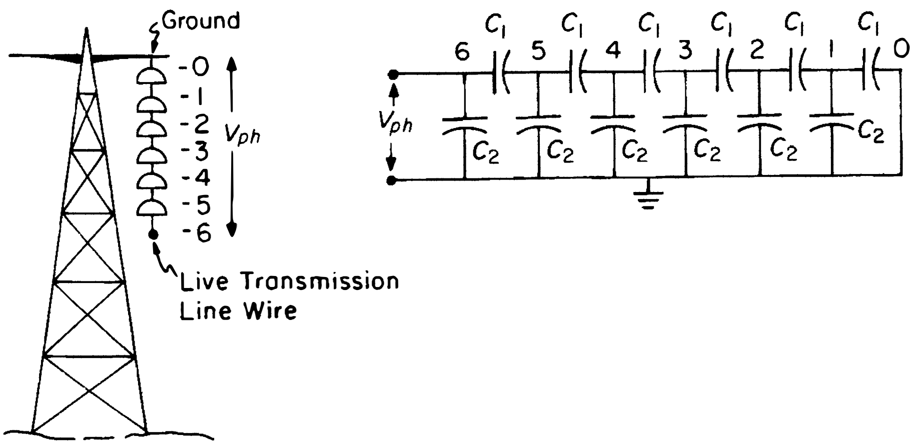
This sequence is generated by the difference equation

$$y[n+2] = y[n+1] + y[n]$$

with initial conditions $y[0] = 0$ and $y[1] = 1$. Treating the sequence as the ZIR of an LTI discrete system, find an explicit formula for $y[n]$.

b) Suppose that newborn rabbits become fertile at the age of one month. Assume that the gestation period for rabbits is also one month and that each litter consists of precisely two rabbits—one male and one female. Suppose further that once a pair becomes fertile it will continue to produce a new litter each month indefinitely. Starting with one newborn pair of rabbits, how many rabbits will there be one year later (assuming that all survive)?

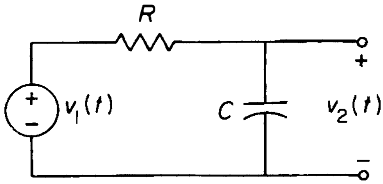
Problem 7.3

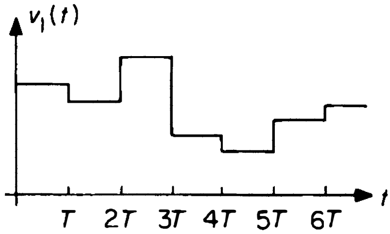


Assume that a string of six insulators supporting a transmission-line wire from a tower can be represented by the equivalent circuit of capacitors shown above for purposes of calculating the voltage distribution along the string. For simplicity, assume that $C_1 = C_2$. Find an expression for the voltage across the k^{th} insulator. If $V_{ph} = 76$ kilovolts, show explicitly how to evaluate the arbitrary constants.

Problem 7.4

a)





If the input to the CT circuit above is the *staircase function* $v_1(t)$, which is piecewise-constant over intervals of length T,

$$v_1(t) = v_1(nT) = v_1[n], \quad nT \le t < (n+1)T,$$

show that it is possible to describe $v_2(t)$ at times $t = nT$ by the difference equation

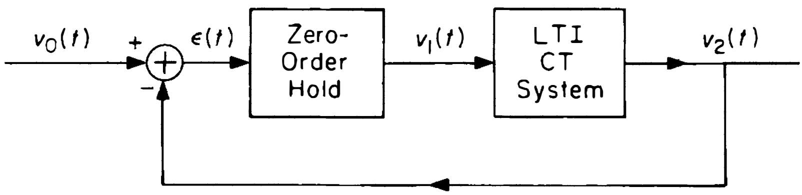
$$v_2[n+1] = \alpha v_2[n] + \beta v_1[n],$$

where

$$v_2[n] = v_2(nT).$$

Find the constants α and β.

b)



The CT circuit above is connected in a feedback arrangement as shown. The *zero-order hold* block has the property that

$$v_1(t) = v_1(nT) = v_1[n] = \epsilon(nT), \quad nT \le t < (n+1)T.$$

Show that for any continuous waveform $v_0(t)$

$$v_2[n+1] = \gamma v_2[n] + \delta v_0[n]$$

where $v_0[n] = v_0(nT)$. Find γ and δ in terms of α and β.

Problem 7.5

This problem deals with a simple model for the price adjustment process in a single commodity that is traded at discrete times. Suppose the *supply* $s[k]$ at time k is determined by the *price* $p[k-1]$ at time $k-1$ (reflecting a delay due to production time) according to the supply law

$$s[k] = s_0 + bp[k-1]$$

and let the *demand* $d[k]$ at time k be determined by the demand law

$$d[k] = d_0 - ap[k]$$

(where a and b are constants that measure the sensitivities of consumers and producers, respectively, to price changes).

a) At each trading time, market forces cause the price $p[k]$ to adjust so that demand equals supply. Use this fact to determine a first-order difference equation that governs price evolution from time $k-1$ to k.

b) Show that a particular solution of this difference equation is

$$p[k] = \frac{d_0 - s_0}{a+b} = \text{constant}.$$

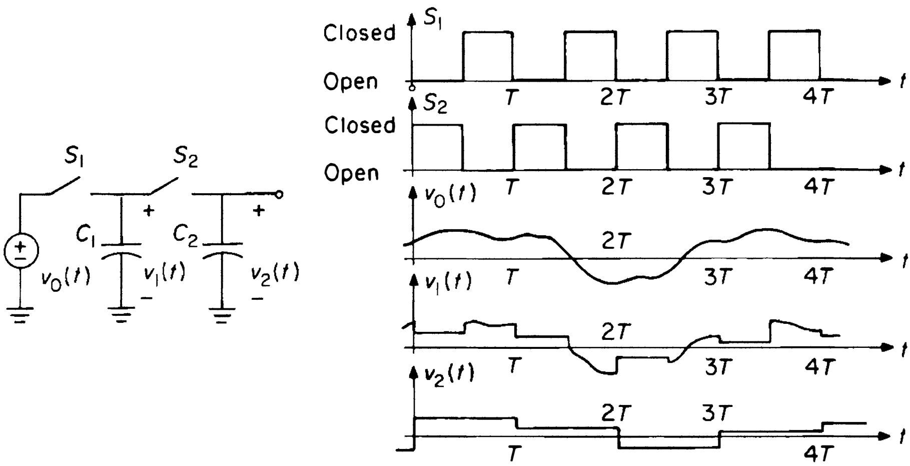
c) Suppose that the initial price has some arbitrary value $p[0]$. Derive an expression for $p[k]$, $k \geq 0$, if $d_0 = 4$, $s_0 = 1$, $a = 2$, $b = 1$. Sketch the result.

d) Repeat (c) for $d_0 = 4$, $s_0 = 1$, $a = 1$, $b = 2$.

e) Observe that in one case the price evolves toward an equilibrium value, whereas in the other case it does not. Determine the condition on a and b required for convergence to an equilibrium value.

Problem 7.6

The following diagrams describe a *commutating* or *switched-capacitor filter*. In practice, the switches are MOS gates driven by a clock with period T, so that they are alternately open and closed as shown below. Assume that the circuit and switch resistances are so small that the capacitor charging times are essentially zero compared with the clock period. Hence, during the time that S_1 is closed, $v_1(t) = v_0(t)$; and during the time that S_2 is closed, $v_2(t) = v_1(t) =$ constant. Moreover, when S_2 is open, C_2 holds its charge.



a) Define $v_0[n] = v_0(nT)$, $v_2[n] = v_2((n+0.5)T)$. Show that $v_2[n]$ and $v_0[n]$ satisfy a difference equation of the form

$$v_2[n] = a\,v_0[n] + b\,v_2[n-1]$$

and find the values of the constants a and b in terms of C_1 and C_2.

b) Suppose that $C_2 = 4C_1$ and that both capacitors are uncharged at $t = 0$. Let $v_0(t) = V_0 u(t)$ (assume the step rises just after S_1 opens). Show that

$$v_2[n] = V_0\left[1 - \left(\frac{4}{5}\right)^n\right] u[n] .$$

c) Let $T = 1$ msec. Sketch $v_2(t)$. On the same scale, sketch $v_2(t)$ if T were 0.25 msec.

d) Let $v_2'(t)$ be the ZSR to some input $v_0'(t)$. Similarly, let $v_2''(t)$ be the ZSR to some other input $v_0''(t)$. Is it true that $v_2(t) = v_2'(t) + v_2''(t)$ will be the ZSR to the input $v_0(t) = v_0'(t) + v_0''(t)$? That is, does this device satisfy the superposition principle?