Discrete-Time Systems

1. Why do We Celebrate Birthdays Once a Year?

A Ramakalyan, P Kavitha and S Harini Vijayalakshmi

In this article we introduce digital control and contrast it with analog control. Traditionally, control systems have been designed using analog techniques. But today, because of the explosive growth of digital technology, controllers are typically implemented as programmable digital hardware or as programs on digital computers. The concepts and techniques for analysis and design of digital controllers are the central concerns of this article. In this part we introduce discrete-time signals and provide examples. We also describe discrete-time systems in terms of difference equations and look at their solutions via transform methods.

Ramakalyan is currently with the Department of Instrumentation and Control Engineering at REC Tiruchi. Kavitha and Harini were final year students in the same Department at the time of writing this article.

Kavitha is now with Hindustan Computers Ltd. at Chennai.

Harini is now at Texas A & M University doing MS with DSP as major.

Introduction

In the article on Systems and Control Engineering[1], we have seen a Proportional plus Derivative (PD) Controller. The actual physical implementation of that need not concern us, but some remarks are in order. Assuming that error and its derivative can be measured, it is necessary to take a linear combination of these in order to determine the actuating signal $u(t)$. Such combinations are readily carried out by circuits built out of devices called Operational Amplifiers. Refer [2] for more details on operational amplifiers.

A more modern alternative, especially for larger systems, is to convert position and velocity to digital form and to use a computer to calculate the necessary controls. For use by a digital computer, continuous-time signals known as analog signals need to be *sampled* and then converted to discrete-time signals by an operation called analog-to-digital (A/D) conversion. After processing, the discrete-time signals should be converted back to continuous-time by a digital-to-analog (D/A) converter. Such a configuration is called a *Sampled-data System*. *Figure* 1 shows a standard

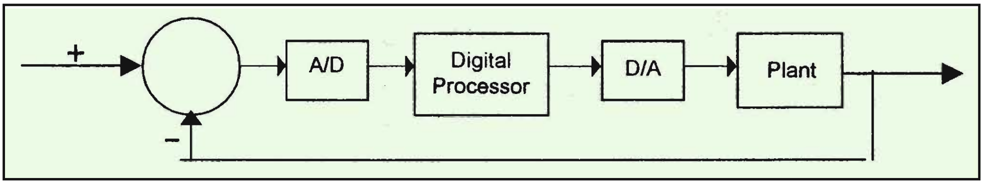

Figure 1.

sampled data feedback control system.

Although the expression *digital control system* refers to a control system with some part in digital form (and that generally means a sampled -data digital control system), occasionally we may have a plant completely described by digital mathematics, and when this plant is controlled by a digital compensator, the system is entirely digital.

Discrete-Time Signals

A signal $f(t)$ is said to be a continuous-time signal or analog signal when it is defined at every instant of time i.e., $\forall\ t \in \Re$.

On the other hand, if the signal is defined only at discrete instants of time and not elsewhere i.e., t takes on *only* the discrete values $t = kT$ for some range of integer values of k, the signal $f(kT)$ is said to be a discrete-time signal. Discrete time signals arise in many areas of science, engineering, business and economics.

One of the most common ways in which discrete-time signals can be obtained is in sampling continuous time signals. As illustrated in *Figure* 2, suppose that $f(t)$, $(t \in \Re)$ is applied to an electronic switch that is closed for a moment every T seconds; the output of the switch can then be viewed as a discrete-time signal $f(kT)$. The resulting discrete-time signal is called the *sampled version* of the original continuous-time signal $f(t)$. The instants at which data appear in a discrete-time signal are called the sampling in-

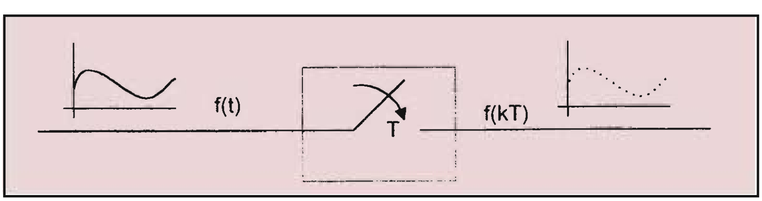

Figure 2.

stants; the time interval T between two subsequent sampling instants is called the sampling interval. The amount of time during which the switch is closed is (assumed to be) much smaller than T. We emphasize here that the sampling process described is called uniform sampling since T is constant. Non-uniform or multi-rate sampling is also used in some applications but is beyond the scope of this article.

The sampling theorem suggests the sampling period required to reconstruct the original continuous-time signal from its sampled version.

Shannon's Sampling Theorem

Let $f(t)$ be a continuous-time signal whose frequency spectrum is band limited to ω (in radians per second). Then, $f(t)$ can be reconstructed from its sampled sequence $f(kT)$ if the sampling period T is smaller than $\frac{\pi}{\omega}$. In other words, the sampling rate must be at least twice the bandwidth or the highest frequency component in the system.

Smaller sampling rates are used in communication systems and larger sampling rates are used in control systems. Typically, audio signals are sampled at 44.1 kHz.

In a discrete-time signal, the amplitude can assume any value in a continuous range. A digital signal is a discrete-time signal with quantized amplitude i.e., the amplitude can assume values only from a finite set. Thus, a digital signal is discretized in time and quantized in amplitude, whereas a discrete-time signal is discretized in time but need not be quantized in amplitude. *Discrete-time* is frequently used in theoretical study, while *digital* is used in connection with hardware or software realizations. In practical usage, the terms *discrete-time* and *digital* are often interchanged. In the present article, we refer to *discrete-time*. Refer to *Figure* 3 for different signals.

A very common example of a discrete-time signal is our age. We do grow older every moment but we celebrate our birthdays only once in a year (and increment our age by one), i.e., the sampling period is one year! The plots of *age(t)* and *age(kT)* are shown in *Figure* 4. In this example, we intend

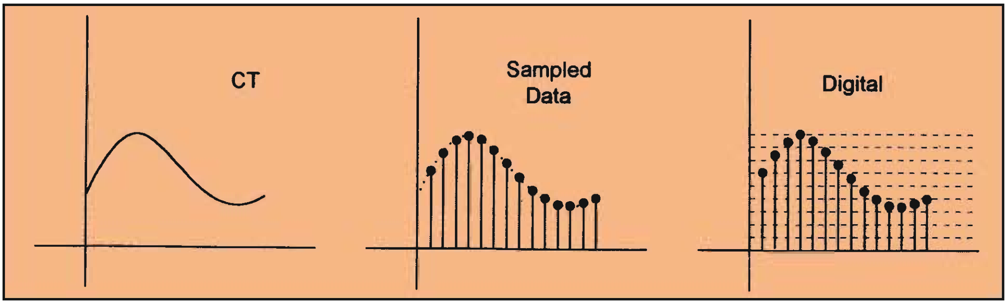

Figure 3.

to emphasize the notion of *biological growth* that generally prevails in individuals. An answer to the question "how old are you?" is always a positive integer!

Other examples of discrete-time signals are daily temperature ($T = 1$ hour), banking transactions ($T = \frac{1}{4}$ year), census (T = 5 years) etc. Whatever may be the sampling interval, it is apparent that we are much more comfortable dealing with discrete-time signals. One reason might be, perhaps, that there is always an inherent delay in processing the signals.

Notice that $f(kT)$ simply becomes a sequence of numbers. The standard convention (which is followed here) is not to show the dependence on T in the notation for the sampled signal. Equivalently, we may assume that the sampling period is normalised to unity.

Discrete-Time Systems

A discrete-time system is defined as a system whose inputs and outputs are discrete-time signals. In Part 2 of [1], we have discussed several properties of systems and the convolution integral description of continuous-time systems.

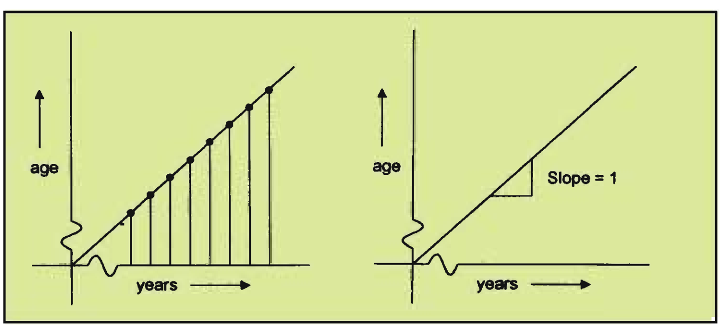

Figure 4.

All those concepts transfer directly from the analog world to discrete-time systems with a little modification as follows.

As depicted in *Figure* 2, the A/D conversion itself may be viewed as a system so that $f(k)$ is the convolution of $f(t)$ and a train of impulses separated by T units of time.

$$f(k) = \sum_{k=-\infty}^{\infty} f(t)\,\delta(t - k). \tag{1}$$

Further if we assume the properties of linearity, time invariance and causality we can see that

$$y(k) = \sum_{i=0}^{k} h(k - i)\,u(i)$$

where u( ) is the input sequence, h( ) is the impulse response, and y( ) is the output sequence.

Example 1: Consider a savings account in a bank. Assume that the interest is compounded monthly at $r\%$.

If we deposit $u(0)$ rupees at the beginning of a month, then $y(0) = u(0)$. Hence, $y(1) = y(0) + \frac{r}{100}\,u(0)$ is the balance at the beginning of next month. It is now easy to see that

$$y(k) = \sum_{i=0}^{k} (1 + (\frac{r}{100}))^{k-i}\,u(i)$$

is the total amount in the account if Rs.$u(i)$ is deposited at the beginning of every month.

We can also write the same convolution summation using induction and arrive at the following equation. We leave the derivation to the interested reader.

$$y(k+1) - (1 + \frac{r}{100})\,y(k) = u(k+1) \;\; ; k = 0, 1, 2, \ldots\ldots \tag{2}$$

This is called a first order linear difference equation with constant coefficients. Linear difference equations are to discrete systems as linear differential equations are to continuous systems. Differential equation has initial conditions as derivative terms evaluated at say $t = 0$ while a difference equation has initial conditions as shifted terms, such as $y(0)$, $y(-1)$, and $y(-2)$ for a third order difference equation, for example. Owing to several reasons such as number of coefficients, number of operations and the presence of initial conditions, the difference equation is preferable to convolution describing a Linear Time Invariant Lumped (LTIL) system. Notice that a similar argument is applicable to differential equations.

Example 2 : The difference equation

$$y(k+1) = (1+b-d)\, y(k) + u(k)$$

describes the population model of any country where b is the birth rate, d is the death rate and $u(k)$ is the number of net immigrants entering the country in year k.

We leave it to the reader to model the vacillating mathematician problem [3] as a discrete-time system.

General Forms of Difference Equations

An n^{th} order difference equation may be written, typically, either as

$$y(k+n) + a_{n-1}\, y(k+n-1) + \quad + a_0\, y(k) =$$
$$b_m\, u(k+m) + b_{m-1}\, u(k+m-1) + . \; . + b_0\, u(k) \quad (3)$$

which is called the *advanced form*, or as

$$y(k) + a_{n-1}\, y(k-1) + \quad + a_0\, y(k-n) =$$
$$b_m\, u(k+m-n) + \quad + b_0\, u(k-n) \quad (4)$$

which is called the *delayed form* where n and m are fixed integers and k ranges over $-n, \; -n+1, \qquad 0, \; 1, \; 2,$ The coefficients a_i and b_i are real constants, not necessarily non-zero. We assume that the coefficient of $y(k+n)$ in (3) or $y(k)$ in (4) has been normalized to one. If $m > n$, it can be verified that the system is not causal. Thus, for causal systems, we require $n \geq m$.

Both the forms are widely used in applications. The advanced form is used in control and filter design, the delayed form is used in estimation, identification and DSP. To find the response $y(k)$, $k = 0, 1, 2...$ excited by an input $u(k)$, we need n initial conditions $y(j)$, $j = -1, -2, \dots -n$. (It is often assumed that $u(k) = 0$ for $k < 0$). Once the initial conditions are given, the response can be obtained by direct substitution. The solution, however, is not in closed form and is difficult to infer from the general properties of the system.

The Z-Transform

Now let us look at a transform called the Z-transform to study the system. Recall, the Laplace transform of a signal $f(t)$ given by

$$F(s) = \int_0^{\infty} f(t)\, e^{-st}$$

for $t \geq 0$. Now that the continuous-time signal is discretized, the integral simply reduces to a summation over the index k, i.e,

$$F(s) = \sum_{k=0}^{\infty} f(kT)\, e^{-skT}$$

Let us define a new complex variable

$$z = e^{sT} \tag{5}$$

Accordingly the transformation becomes

$$F(z) = \sum_{k=0}^{\infty} f(kT)\, z^{-k} \tag{6}$$

and $F(z)$ is called the Z-transform of $f(kT)$. Several properties analogous to the Laplace transform such as linearity, frequency shifting, time shifting, initial value theorem, final value theorem etc. are also applicable for the Z-transform. We may apply the Z-transform to the difference equation and obtain a closed form solution. We have given some important properties of the Z-transform and the Z-transform of some signals in *Box* 1. Analogous to continuous-time system discussed in [1], we get the zero input response (due to initial conditions alone) and the zero state response (due to externally applied input alone). We also have the notions of

Box 1. Z-Transforms.

An extensive list of Z-transforms may be found in [4]. However, a few signals and their Z-transforms are given below.

Unit impulse	$Z\{\delta(k)\} = 1$
Unit step sequence	$Z\{q(k)\} = \frac{z}{z-1}$
e^{ak}	$Z\, e^{ak} = \frac{z}{z-e^{a}}$
$\cos(\omega k)$	$Z \cos(\omega k) = \frac{z(z-\cos(\omega))}{z^2 - 2z\cos(\omega) + 1}$
$f(k-i)$	$z^{-i}F(z) + z^{-i}f(-1) + \ldots + f(-i)$
$f(k+i)$	$z^{i}F(z) - z^{i}f(0) - \ldots - z\, f(i-1)$
initial value $f(0)$	$\lim_{z\to\infty} F(z)$
final value $f(\infty)$	$\lim_{z\to 1}(z-1)F(z)$

transfer function (the ratio of the transform of output to the transform of input with all initial conditions equal to zero), and complete characterization. In other words, Z-transforms are to discrete-time systems as Laplace transforms are to continuous-time systems. However, in the present article, we shall employ a new technique to solve the difference equation. This technique is called the State-Space description of LTIL systems.

Description of Systems in State Space

For the sake of clarity, let us assume that the order of the system is 2 and that all b_is except b_0 are zero in (3). The system is now

$$y(k+2) + a_1 y(k+1) + a_0 y(k) = b_0 u(k). \quad (7)$$

Let us now define 2 variables $x_1(k)$ and $x_2(k)$ as

$$x_1(k) = y(k), \quad x_2(k) = x_1(k+1) \quad (8)$$

so that in terms of these variables the system (7) may be rewritten as

$$x_2(k+1) = -a_0 x_1(k) - a_1 x_2(k) + b_0 u(k). \quad (9)$$

Together with the definitions of the 2 variables in (8) we may write

$$\vec{X}\ (k+1) = A\ \vec{X}\ (k) + B\ u(k), \tag{10}$$

and

$$y(k) = C\ \vec{X}\ (k) \tag{11}$$

where

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\ A = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}\ B = \begin{bmatrix} 0 \\ b_0 \end{bmatrix}\ C = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{12}$$

The following observations may be readily made.

1. An n^{th} order difference equation is split into n first order difference equations.

2. The vector $\vec{X}$ comprising the variables $x_i, i = 1,2,3, ...n$ is called the *state* vector. If x has n state variables, then it has the dimension $n \times 1$. Notice that the initial conditions remain intact. They are simply renamed as $\vec{X}(0)$! Each of the first order difference equations describes the evolution of an internal *state* of the system. Accordingly, the name of the variable is *state variable* and hence the vector $\vec{X}$ is state vector. Equation (10) which involves the dynamics of the state vector is called the *state equation.* It describes the relationship between the input and the state.

3. For an n^{th} order system, the square matrix A, called the system matrix, has the dimension $n \times n$. (What are the eigen values of A?)

4. If there is a single input, then the dimension of B is $n \times 1$. It is easy to see that the dimension of B becomes $n \times r$ if the system is driven by r inputs, packed into a vector $\vec{U} = [u_1, u_2, ... u_r]^T$

5. If there is a single output, then C is simply a row vector; otherwise its dimension is $m \times n$, where m is the number of outputs of the system. The algebraic equation (11) which gives the output in terms of the state vector is called the *output equation.*

6. Without loss of generality, we may also include u in the output equation with an appropriate weight i.e., $y = c\vec{X} + D\,u$. (When does D come into the picture?)

State-variable equations are also called state-space equations because the state vector forms a linear space. By now, the reader should appreciate the obvious advantage of such a description of a system in terms of state variables. Unlike transfer functions, which are limited to single input single output or SISO systems, state-space equations are readily extendable to multi-input multi-output or MIMO Systems.

Originally, the idea of state space was given by Poincaré. But it was Rudolf Kalman who had exploited this description in the 1960s. This approach introduced the concepts of controllability and observability, which led to a better understanding of the structure of systems, and a number of new results, such as state feedback and state estimator. Refer to Part 1 of [1] for a brief history of modern control.

Applying Z-Transform to State Space Equations: By applying the Z-transform to the state-space equations (10) and (11), we get

$$\vec{Y}(z) = C\,(ZI - A)^{-1}\,[Z\,\vec{X}(0) + B\,\vec{U}(z)] + D\,\vec{U}(z)$$

For a SISO system, we may easily verify that

$$\frac{Y(z)}{U(z)} = H(z) = C\,(ZI - A)^{-1}\,B + D$$

(with the initial conditions assumed to be zero.) This is a fundamental equation relating the state-space equation and discrete transfer function. (What is the characteristic polynomial of the system?)

An important point is that the state space description of a system is not unique. The one presented in this article, called the *controllable canonical form*, is simple and is intuitively appealing. Physically, not all the state variables defined as before may have a meaning. However, mathematically, we may perform similarity transformations to get a variety of models for the same system.

Suggested Reading

[1] A Ramakalyan and J R Vengateswaran, Systems and Control Engineering, *Resonance*, Vol. 4, Nos. 1, 3 & 5, 1999.

[2] Kapil Krishnan Manu and R Ramaswamy, Chaos, *Resonance*, Vol. 3, Nos. 4, 6, 10, 1998.

[3] K B Athreya, The Vacillating Mathematician Problem, *Resonance*, Vol. 2, No. 1, 1997.

[4] C T Chen, *System and Signal Analysis*, Holt, Rinehart and Winston, N.Y., 1994. (Also available from Oxford Univ. Press priced in rupees.)

[5] J G Proakis and D G Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3/e, Prentice Hall of India, 1997.

[6] A V Oppenheim and R W Schafer, *Discrete-time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989. (Also available in Prentice Hall of India).

Address for Correspondence
A Ramakalyan
Department of Instrumentation and Control Engineering,
Regional Engineering College
Tiruchirapalli 620 015, India.
Email:rkalyn@rect.ernet.in

Solution to Crossword Puzzle

(Published in the previous issue)

1 M	E	2 N	D	3 E	L	E	4 E	V		5 X
O		E		M						Y
6 L	A	U	R	I	C		7 E	N	O	L
A		T		T			S			A
8 R	A	R	E		9 I	N	T	10 R	O	N
		O			V			O		
11 B	I	N	12 A	R	Y		13 S	T	A	14 R
O			N			15 A		I		E
16 Y	O	L	K		17 O	L	E	F	I	N
L			L			U		E		A
E		18 P	E	N	U	M	B	R	A	L

(Cross Word Puzzle Contributed by Vinita Krishnamurthy)