SAMPLED-DATA CONTROL SYSTEMS

John R. Ragazzini
Dean, College of Engineering
New York University

Gene F. Franklin
Assistant Professor of Electrical Engineering
Stanford University

McGRAW-HILL BOOK COMPANY, INC.
New York Toronto London
1958

SAMPLED-DATA CONTROL SYSTEMS

Library of Congress Catalog Card Number 58-8856

III

51145

THE MAPLE PRESS COMPANY, YORK, PA.

CONTENTS

CHAPTER 1

INTRODUCTION

The trend of the past few decades has been toward dynamical systems that operate with variables which are in the form of a sequence of numbers. These variables are generally quantized in amplitude and are available only at specified instants of time, which are usually equally spaced. By contrast, a continuous, or analogue, system has variables which are continuous functions of time, that is, their values are known at all instants of time. Both types of system can have imperfections in the amplitude of the signal variables. For instance, the discrete system, in which the variables are sequences of numbers, may operate with these variables quantized so that even if there is no other source of amplitude error, there is the uncertainty in the magnitude equal to one quantum. In continuous systems, imperfections in the data-transmission and transducing devices, as well as unwanted noise, produce uncertainties in the amplitude of the system variables which are similar to those of the discrete systems. The major point of difference between analogue and discrete systems lies in the fact that analogue, or continuous, systems have variables which are known at *all* instants of time, whereas discrete systems have variables which are known only at *sampling instants.*

A system in which the data appear at one or more points as a sequence of numbers or as pulses is known as a *sampled-data system.* A system in which the data are everywhere known or specified at all instants of time is known as a *continuous,* or *analogue,* system. This book deals with sampled-data systems, the theory underlying their operation, and the synthesis of systems of this type which fulfill certain practical objectives.

1.1 The Sampling Operation

In any dynamical system found in nature, there exist dependent and independent variables which are related to each other by linear or nonlinear differential equations. In the systems approach, independent variables are referred to as *inputs* and dependent variables as *outputs.* In complex systems there are also intermediate variables, which are considered as being internal in the system, although they can be brought out as outputs should the necessity arise. Assuming for purposes of discussion that $f(t)$ is a variable of interest, it is plotted in Fig. 1.1 as a continuous function of time. The plot or some analytic expression for $f(t)$ will describe the function completely as a function of time.

If, now, the value of $f(t)$ is read or sampled at equal intervals of time T so that the function is described by the sequence of numbers

$$f(0), f(T), f(2T), f(3T), \ \ldots \ , f(nT), \ \ldots \qquad (1.1)$$

it is seen that a limited description of the function $f(t)$ has been given. For instance, the value of $f(t)$ at $f(1.5T)$ is not available, so that a certain amount of information has been lost in the process of expressing $f(t)$ as a number sequence given by (1.1). On the other hand, if the function is well-behaved, the intermediate values of $f(t)$ can be interpolated between samples with acceptable accuracy. If the function is not well-behaved, it means that large and unpredictable variations in $f(t)$ have occurred between sampling instants. The number sequence such as that of (1.1) then gives only a poor approximation of the variable.

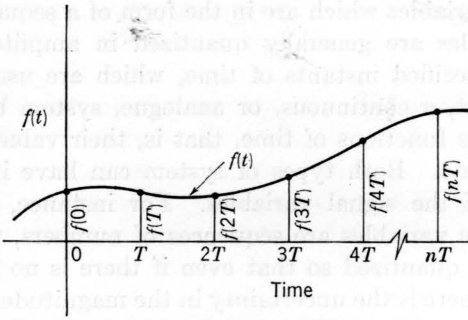

FIG. 1.1. The sampling operation.

It is seen from this simple qualitative discussion that the sampling frequency must be related to the characteristics of the function being sampled, lest important information be lost in the sampling process. At the same time, if the sampling frequency is well chosen relative to the characteristics of the time function being sampled, only negligible information is lost in the sampling process. In the latter circumstance, the use of more samples would merely burden the system by carrying unessential information that could have been obtained by the simplest of interpolative processes.

Considerations such as these suggest that continuous systems are capable of carrying and transmitting far more information than is required or justified by the dynamical-system characteristics. In the frequency domain, this is equivalent to stating that a capability of some components of the system to carry and transmit excessively large bandwidths is not justifiable if some of the cascaded components transmit restricted bandwidths. If there are practical advantages to be gained by transmitting and processing only a sequence of numbers as opposed to a continuous variable, then a proper selection of sampling frequency and the use of a sampled-data system seems desirable.

There are situations when the data-gathering devices themselves are capable of producing only discrete sets of numbers rather than a continuous variable. For instance, a scanning search radar will generate a fix on a target only once every scan. In some large-scale radars, this might occur only once every 10 or 15 sec. Between these scans, or "looks," no information exists as to the variations in target position. Another possibility is the use of time-shared data links in which information can be transmitted only once every cycle time. In such situations, a system which incorporates one of these devices as an element is, of necessity, a sampled-data system. On the other hand, it will be shown later that there are certain advantages to be gained by deliberately converting a continuous feedback control system into a sampled-data system. The use of sampled-data controllers results in systems having dynamical performance which cannot be matched by the continuous system from which they are derived.

1.2 Data Reconstruction

It was stated in the previous section that the continuous function from which the number sequence is obtained can be reconstructed by processes of interpolation or extrapolation. In numerical computation, this is done by using many samples obtained before or after the region of interest. On the other hand, real-time dynamical systems can use only past samples since the future samples are not known. Thus, data reconstruction must be a process of extrapolation using only the preceding set of samples. This process is sketched in Fig. 1.2, where a continuous function is being extrapolated from the latest sampling instant at nT. The extrapolation in real-time systems is carried out for only one sampling interval, extending from nT to $(n + 1)T$. Since the value of the function is known exactly at the next sampling instant $(n + 1)T$, this most recent value can be used as the base for an extrapolation into the next sampling interval. Thus, the extrapolation process is reiterated as each new sample becomes available. There are a number of techniques and extrapolation formulas which can be used to implement this process. In all cases, the objective is to reproduce as well as possible a reasonable facsimile of the actual time function from which the sample or number sequence was derived.

The reason why data reconstruction is important in the field of dynam-

For instance, if a very simple form of data extrapolation in which only the first term, $r(nT)$, is taken, the extrapolator is referred to as a *zero-order* hold since the polynomial generated by this system is of zero order. Similarly, if the first two terms are implemented, the classification is *first-order* since the polynomial which will be extrapolated is of first order. It is recognized that the zero-order data hold is also known as a data clamp and that it operates on the assumption that the value of the function in a given sampling interval is equal to the function at the beginning of that interval. This extrapolation is perfect only for functions which are constants. Practical systems rarely employ data extrapolators which are beyond first order, both for reasons of economy and because if too many back differences are taken, an excessive settling time results and noise effects are increased. These points will be discussed later in more detail.

3.3 The Zero-order Data Hold

As indicated in the previous section, the zero-order data hold includes only the first term of the series as expressed in (3.12). This form of data hold is important from a practical point of view because of its simplicity and the fact that it is readily implemented. A standard electronic clamp circuit will set its output at a level equal to or proportional to the magnitude of an input pulse and then reset itself when a new pulse is applied. Such circuits maintain a constant output between pulses and thus implement the zero-order-hold relationship. Similarly, a digital register will hold a number until a reset pulse is applied and a new number set up. In all cases, the output of the device essentially assumes that the continuous function within a sampling interval is constant and equal to the value of the function at the preceding sampling instant.

The form of the reconstructed function at the output of a zero-order hold is shown in Fig. 3.4. The extrapolation in each case is a constant and is refreshed at each sampling instant. Because of the appearance of the reconstructed function, the data hold is sometimes referred to as a "staircase" or "boxcar" data system.

In order to include the effects of a zero-order data hold in a dynamical system, it is necessary that a mathematical description of its effect be obtained and, if possible, a transfer function derived. It is recalled that the transfer function of a linear system is the Laplace transform of the impulsive response of the system. In the case of a zero-order data hold, it is assumed that the short pulse which is applied to its input is approximated by an impulse of an area equal to the magnitude of the pulse. If an impulse of unit area is applied to the data hold, its response should be a unit-magnitude continuous function which is maintained until refreshed by the next impulse. In view of the assumption that the sampling intervals are all equal to T, the extrapolated value of the function should fall to zero T sec after the application of the unit impulse in order that the next impulse may restore the data hold to its new value.

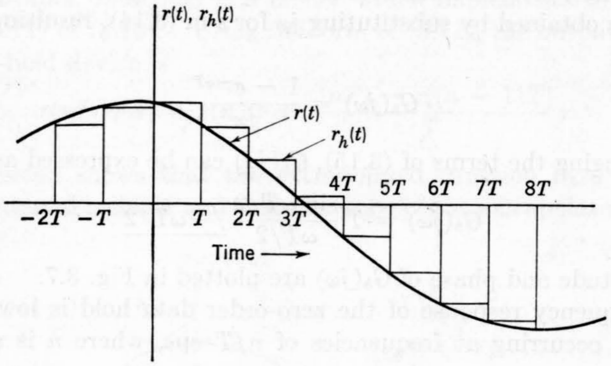

FIG. 3.4. Reconstruction of $r(t)$ by a zero-order data hold.

The impulsive response of the zero-order data hold should therefore appear as shown in Fig. 3.5.

To obtain the transfer function of this system, the impulsive response

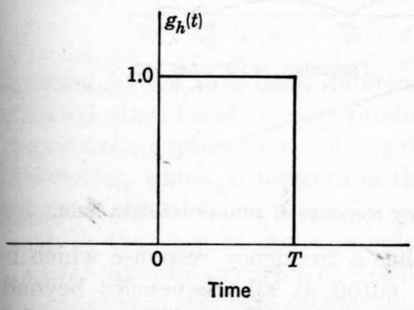

FIG. 3.5. Impulsive response of a zero-order data hold.

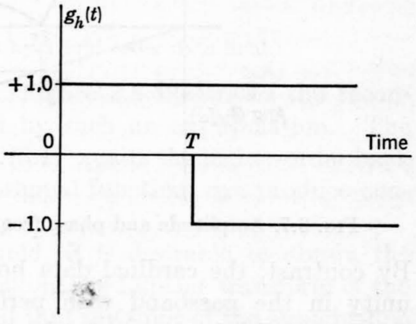

FIG. 3.6. Step-function components of impulsive response of zero-order data hold.

can be decomposed into two unit step functions, as shown in Fig. 3.6. The impulsive response is given by

$$g_h(t) = u(t) - u(t - T) \tag{3.13}$$

where $u(t)$ is the unit step function. The Laplace transform of $g_h(t)$ is

$$G_h(s) = \frac{1}{s} - \frac{1}{s} e^{-Ts} \tag{3.14}$$

This transfer function is useful in the analysis of systems which include