SAMPLED-DATA CONTROL SYSTEMS

John R. Ragazzini Gene F. Franklin

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

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

This book deals with the theory of sampled-data systems, a subject which has been of increasing interest and importance to engineers and scientists for the past decade. The science and art of communications have profited from the realization and application of the fact that intelligence can be transmitted and stored in discrete pieces or as a sequence of numbers spaced in real time. As we hope to have shown in this book, the control systems field can similarly benefit by the utilization of this concept. Even though we treat sampled-data systems primarily from the viewpoint of the control function, it is not surprising that many concepts have been borrowed from the communications field. Control systems are essentially power devices which respond to intelligence that has been processed in subsystems similar to those in the communications field. Furthermore, the same body of theory can be used to describe the over-all performance of the control system, even though its primary function is the controlled actuation of power elements and processes.

Sampled-data systems are characterized by the fact that the signal data appear at one or more points in the system as a sequence of pulses or numbers. A central problem in the theory of such systems is that of describing the response of linear continuous elements, or pulsed filters, as they are sometimes called, to pulse sequences applied to their input. The use of the z transformation and the all-important pulse transfer function of the pulsed filter makes this problem relatively straightforward. A unique component found in sampled-data control systems is the digital controller, which is a computer that accepts a sequence of numbers at its input, processes it in accordance with some logical program, and applies the resultant sequence to the controlled element. In view of the operation of this type of controller, it is possible to implement it by means of a conventional digital computer or its equivalent in the form of a mixed or wholly analogue computer. If the numerical process programmed in the computer is linear, it can be expressed mathematically in terms of a recursion formula which is transformed into a generating function having similarity to the pulse transfer function of a pulsed linear filter. It is not unexpected to find the same general theory applying equally well to pulsed linear filters and to the description of linear numerical processes. Sampled-data theory which is developed in this book serves as a common base for the analysis and synthesis of linear digital systems, pulsed continuous systems, and their combinations often found in practice.

Contributions to the theory of sampled-data systems have been made by scientists, mathematicians, and engineers throughout the world. An examination of the list of references and bibliography in this book will reveal papers from many countries, including England, France, the U.S.S.R., and the United States. As in the case of all new fields, the research papers listed are not equally significant. The philosophy we have used in writing this book is that a major responsibility of the author is to sift, evaluate, and interpret the significant contributions. This is particularly important when a book is among the first, if not the first, in its field, for all too often its coverage tends to set the pattern for subsequent books. It would have been far easier for us to write a book which is merely an organized compendium of the papers in the field. It has been much more difficult to be selective, and we fully expect that others who are well-versed in this field may not agree with our choice of material.

As a result of the application of this philosophy, this is a rather short book. We have tried to avoid overwhelming our readers with verbiage or confusing them with a large number of disconnected items which might have been included for the sake of completeness. We have directed this book to readers who are mature technically and who are capable of referring to the literature when necessary. To make this easier, the book is documented as fully as possible.

This is not a book for beginners in the field of control systems. It is assumed that the reader is a graduate student, practicing engineer, or scientist who has had a thorough training in differential equations, the Laplace transformation and its applications, linear feedback control theory, and the elements of probability and statistics. On the other hand, it is an introductory text, and the reader need have had no prior contact with the theory of sampled-data systems or numerical processes. While specifically directed to control systems, there is much material in this book which has general application. This includes the z transformation, data-reconstruction theory, applications of transform methods to numerical processes, and the theory of sampled random time functions.

The level of presentation is such that the book can be used as a text for a graduate course on the subject. Depending on the preparation of the students, this could be a one-semester course of three hours per week or a two-semester course of two hours per week. In exceptional cases, where the students have had a thorough grounding in linear systems, feedback control, and the Laplace transformation, it is possible to use this book in a senior course. The material has also been used as the basis of a seminar in which literature study was the main element of the course.

The authors have been engaged in the study of sampled-data systems for a number of years and have supervised doctoral research in this field at Columbia University and presently at Stanford University as well. A strong impetus to the advancement of this activity was and is now provided by the United States Air Force Office of Scientific Research, under whose auspices much of the research in sampled-data systems was done at Columbia University. This support is gratefully acknowledged by us, our colleagues, and graduate students.

An attempt to list all those individuals who have been of assistance to us in one way or another would surely lead to the embarrassment of having omitted some. Risking this, however, we shall mention a few and recognize their many suggestions with thanks. Included are Professors Lotfi A. Zadeh, John E. Bertram, George M. Kranc, and Bernard Friedland, Dr. Rudolph E. Kalman, and the many graduate students and research assistants with whom we have been associated.

JOHN R. RAGAZZINI
GENE F. FRANKLIN

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

$$f(0), f(T), f(2T), f(3T), \ldots, f(nT), \ldots \tag{1.1}$$

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

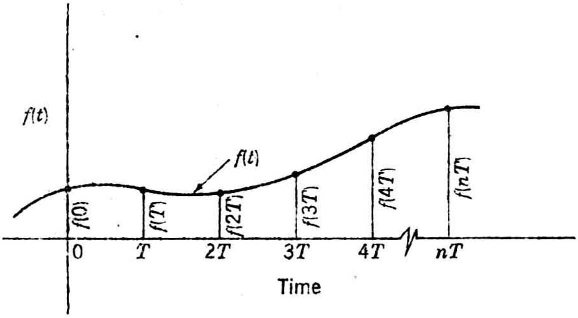

FIG. 1.1. The sampling operation.

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.

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 dynamical sampled-data systems is that physical plants and dynamical devices are basically analogue, or continuous, in form. For instance, in a control system, the actuator may be an electric motor which responds to a continuous signal input and delivers a continuous output. If such a motor is incorporated into a sampled-data feedback control system, continuous signal at its input must somehow be reconstructed to obtain satisfactory

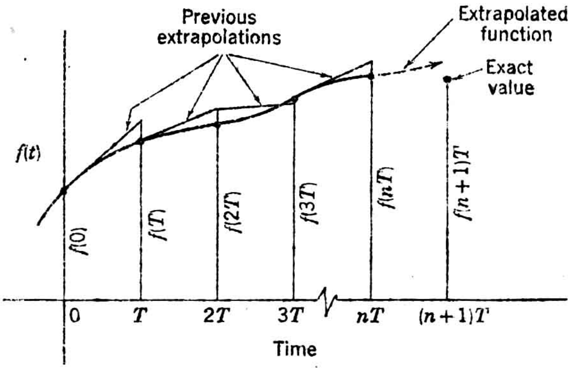

FIG. 1.2. Data reconstruction process.

operation. The devices which reconstruct continuous data from a sequence of samples or numbers are generally called *data holds, extrapolators, desampling filters,* or some similar descriptive name. They all have the same function and, from the practical viewpoint, they are made as simple as possible. In obtaining this physical simplicity, accuracy of extrapolation is often sacrificed.

1.3 Open-loop Sampled-data Systems

A sampled-data system is an interconnected group of dynamical elements in which the signal data appear at one or more points in the system as a sequence of numbers. Figure 1.3 shows the simplest form of

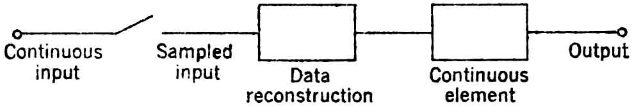

FIG. 1.3. Typical open-cycle sampled-data system.

open-cycle sampled-data system. As a result of a sampling operation, the continuous input signal is converted into a sequence of numbers equally spaced in time by a sampling interval T. The operation of sampling is shown schematically by a switch which is presumed to close momentarily at each sampling instant. The sequence of samples emerging from the switch is reconstructed into an approximation of the input function before being applied to the continuous element. The output of this element is the useful output of the system.

The schematic representation of Fig. 1.3 is intended to show only a possible sequence of operations, not necessarily the physical elements themselves. For instance, this system could represent a pulse-code communications system in which the sampling and coding operation is symbolized by the switch. The quantizing aspect of the operation is ignored here, since it is assumed that the input is quantized infinitely fine. Thus, the input amplitude is presumed to be perfect in this representation. The data-reconstruction element reconstructs a continuous signal from the sequence of samples as well as is practical. Usually, this can be relatively crude, and the physical device takes the form of a simple clamp or boxcar circuit. The continuous element is the device which is being driven by the reconstructed signal, and its output is the useful signal.

The theory which underlies the performance of this system should take into account two deteriorating aspects: the quantizing effect and the sampling effect. Both of these tend to distort or deteriorate the signal in some way. It is much easier to take into account the effect of sampling since it will be shown that this can be described by means of linear difference equations. On the other hand, the quantizing effect is much more difficult to account for, since it is described by nonlinear equations. All the theory in subsequent sections will deal with the linear problem, on the assumption that the quantization of the variables is made fine enough to produce negligible effect. Generally, the theoretical objectives which apply to systems of the type shown in Fig. 1.3 are to obtain the output sequence or continuous output in terms of the input sequence and the system parameters.

1.4 The Sampled-data Feedback System

If the system configuration includes elements which feed the output variable back to the input and if a sampling operation is included, the system is referred to as a *sampled-data feedback system*. If the objective of the system is to control one or more variables in the system so that they have a desired functional relationship with the inputs and disturbances, the qualifying term *control* is included in the name.

A simple sampled-data feedback control system is shown in Fig. 1.4. In this system the error signal is sampled and is reconstructed before being applied to the continuous element. The latter may be the plant or process which is being controlled, including amplifiers, instruments, and actuators. This error-sampled system can be compensated by the addition of networks in the continuous element, just as in the case of ordinary continuous feedback systems. The problem of designing such network is considerably more complex, however, because of the presence of th sampling operation.

A configuration which is unique to sampled-data systems is one i which a digital controller is used, as shown in Fig. 1.5. In this systen the controller accepts a sequence of numbers and processes them, usuall linearly, to produce an output number sequence. The latter sequence i

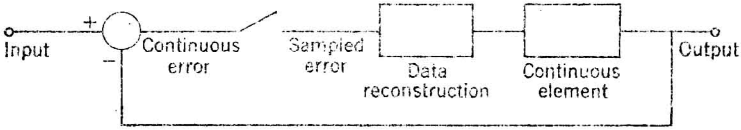

FIG. 1.4. Typical error-sampled feedback control system.

reconstructed into a continuous command signal and is applied to th plant. If the linear program of the digital controller is properly designec the over-all system can be stabilized and its dynamical performanc made to conform to fairly rigid specifications. The digital controlle

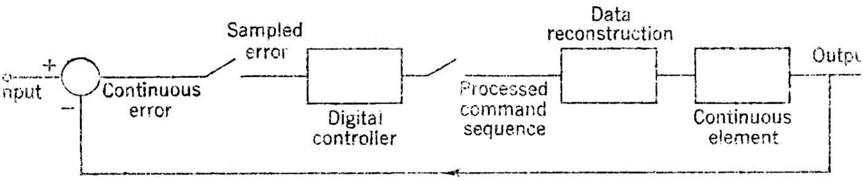

FIG. 1.5. Sampled-data feedback control system using digital controller.

may be implemented by digital-computer techniques or it may emplo a mixture of analogue and digital components. Its main requiremer is that it be capable of receiving a sequence of numbers equally space in time and of processing them in real time into a command signal. I will be shown later that controllers of this type can produce system whose performance cannot be duplicated by all-continuous systems.

The problems which must be studied in sampled-data feedback contr systems include all those encountered in continuous systems. First, criterion for stability must be derived and adapted for application t physical problems. Second, a means for relating the input and outpu which is as direct and simple as the Laplace transform in continuou systems must be developed, along with a means for shaping and compen sating the system. A unique property of sampled-data systems is tha the output will contain a small periodic output component which is th result of intermittency in the signal within the system caused by th sampling operation. This periodic variation is known as *ripple*, an methods for analyzing this component and reducing or controlling it magnitude are required.

There are many possible configurations possible in addition to those shown in Figs. 1.4 and 1.5. Sampled signals may exist at several points in the system as well as in the error line. There may be dynamical elements in the feedback line, and there may be multiple loops. The transform methods for sampled-data systems must be applicable to all possible configurations.

1.5 The *Z* Transformation

Continuous linear dynamical systems are described mathematically by a set of linear differential equations. While their solution can be carried out by classical methods, the use of the Laplace transformation organizes and simplifies the process. What is even more important, inversion of the transform of the variable of interest is rarely necessary in order to deduce the important characteristics of the system and their relation to the system constants. Mapping techniques on the complex plane in the form of transfer loci or root loci further clarify the properties of the system. Certainly, the value of the Laplace transform as a tool for the analysis and synthesis of linear continuous systems is indisputable.

Linear sampled-data dynamical systems are shown to be described by a set of linear difference equations, provided that all the samplers in the system are synchronous, that is, their sampling periods are equal or related by integers. Some of this earlier work, as reported by Oldenbourg and Sartorius,[46] was motivated by the use of intermittent error-sensing devices such as the chopper-bar galvanometer, shown schematically in Fig. 1.6. In this type of device, a small error voltage or current is applied to the galvanometer coil. While the chopper bar is raised, the sensitive galvanometer movement is free and the coil responds with a large displacement in response to the weak signal. Periodically, the chopper bar is lowered and the projecting galvanometer needle causes a bell crank to be rotated more or less proportionately to the deflection angle θ. The bell crank causes the output shaft to rotate with a torque capacity determined by the chopper-bar drive rather than the galvanometer-coil drive.

The main point of interest here is that a datum is stored in the output shaft just once per cycle of the chopper-bar drive. In a sense, the intermittency of the output signal has been accepted in return for a high sensitivity of the system. The early work by Oldenbourg and Sartorius generalized systems of this type into the form of the sampled-data block diagrams of Figs. 1.4 and 1.5. It was shown that these systems could be described by a set of linear difference equations whose solution could be obtained by classical methods. The linear sampled-data system was therefore placed in the same status as the continuous system, using classical methods to solve the differential equations.

In the field of mathematics, Demoivre and Laplace[9,33] developed a form of transform calculus which could be applied to the solution of linear difference equations. This approach was adapted to the solution of pulsed filters and sampled-data systems by Hurewicz,[17] who laid much

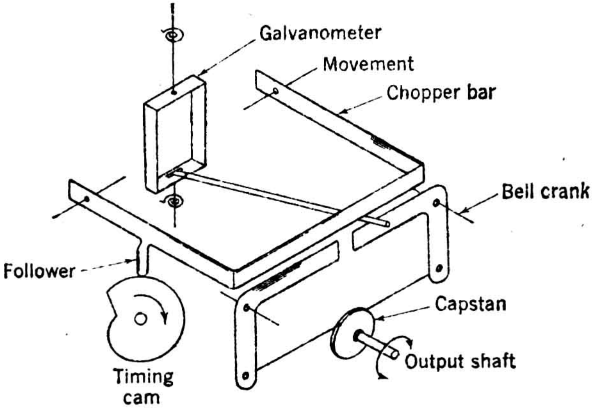

FIG. 1.6. Sketch of chopper-bar galvanometer.

of the basic groundwork for the transform method of analysis of sampled-data systems. Subsequent investigations[1,2,34,47] further extended this initial work. The result of these efforts was the development and refinement of the so-called z *transformation* and its application to the analysis and synthesis of sampled-data systems.

The z transformation is entirely analogous to the Laplace transformation and its application to continuous systems. It turns out that, for systems having lumped constants, that is, those which are described by linear difference equations with constant coefficients, the z transformation gives expressions which are rational polynomial ratios in the variable z. This variable is complex and is related to the complex frequency s used in the Laplace transform by the relation $z = e^{Ts}$. In z-transform theory, such concepts as the transfer function, mapping theorems, combinatorial theorems, and inversion bear the same powerful relation to sampled-data systems as does the Laplace transformation to continuous systems.

Without going into detail at this point, the general concept of the z transformation as applied to systems is shown in Fig. 1.7. Here the output number sequence of the system is related to the input number sequence by a linear difference equation. If the sampled output is $c^*(t)$ and the input is $r^*(t)$, and if the z transforms of these sequences are $C(z)$ and $R(z)$, respectively, a pulse transfer function $G(z)$ can be found which relates them in the following manner:

$$C(z) = G(z)R(z) \tag{1.2}$$

The form and constants of the pulse transfer function $G(z)$ are a property of the system and can be found in terms of the system constants. These relations will be rigorously derived in later chapters.

The relation expressed in (1.2) is dependent on the fact that the input and output samples are taken with the same sampling instants.

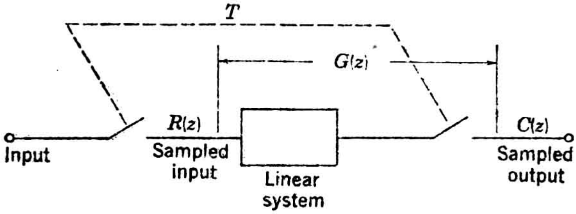

FIG. 1.7. The pulse transfer function $G(z)$.

It is possible to extend the concepts of the z transformation to include the case where the output and input samples are taken at some integral multiple of a basic sampling rate. For instance, if the basic rate is taken as unity, the input and output sampling operations can take place at two and three times the basic rate, respectively. Such systems are referred to as *multirate* sampled-data systems, and suitable modifications in z-transform theory can be made to cover these cases. A typical system is shown in Fig. 1.8, where the input sampler operates with a sampling

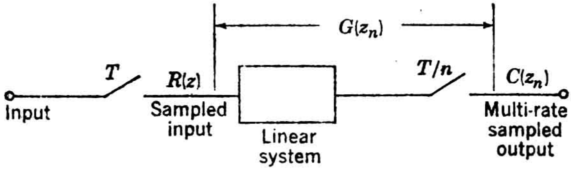

FIG. 1.8. Multirate sampled-data system.

interval T, while the output sampler has a sampling interval T/n. The pulse transfer function relating these sequences at input and output is the multirate pulse transfer function $G(z_n)$. If $R(z)$ is the z transform of the input sequence and $C(z_n)$ is the multirate z transform of the output, then they are related as follows:

$$C(z_n) = G(z_n)R(z) \tag{1.3}$$

Inversion of (1.3) will yield the multirate output sequence.

Sampled-data systems are often subjected to inputs or disturbances which are random. To handle this situation analytically, a number of concepts and definitions analogous to those for continuous systems must be devised. Such terms as auto- and cross-correlation function of sample sequences, sampled power spectra, and cross spectra are used. Relations which give the shaping effects of a linear sampled-data system can be found, just as in the case of continuous systems. Techniques are available to optimize the performance of sampled-data systems based on mean-square criteria used in the design of sampled-data feedback control systems and filters.

1.6 Miscellaneous Uses of Sampled-data Theory

If a linear system contains variables which are actually sampled, analysis by use of the z transformation is exact. Interestingly enough, the same theory can be applied approximately to models of continuous systems in which sampling of the variable is introduced artificially as an aid to analysis. For instance, with the continuous feedback control system shown in Fig. 1.9a, it is often desired to obtain the response of the system to an input in the time domain using ordinary inversion of the Laplace transform of the output variable. In principle, this is very simple and straightforward, but if an accurate solution is desired, the process can be quite laborious, requiring the use of calculating machines.

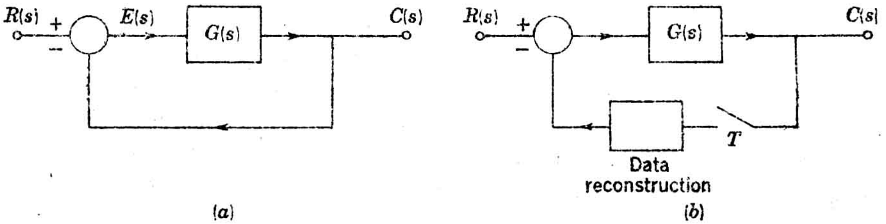

FIG. 1.9. (a) Continuous feedback system. (b) Sampled model of feedback system for computation.

It so happens that one of the techniques for inversion of the z transform is directly accomplished by routine numerical processes. This advantage can be applied to continuous systems by constructing a sampled model which gives solutions with tolerable error. Such a model for feedback systems often takes the form shown in Fig. 1.9b. By selecting the sampling rate high enough and using a sufficiently sophisticated data-reconstruction element, acceptable accuracy can be achieved. As a matter of comparison, the sampling interval is exactly analogous to the quadrature interval which would be selected in the numerical integration of a differential equation. The sampled-data approach has the advantage, however, that a physical interpretation of the process is readily seen. Having selected the sampled model of the continuous system, its analysis becomes one of numerical methods simply carried out by a desk calculator or digital-computer program.

The use of a sampled model in this as well as other applications has the advantage of making clear just where the sampler should be placed