DIGITAL PROCESSING OF SIGNALS

THEOEY AND PRACTICE

MAURICE BELLANGER

DIGITAL PROCESSING OF SIGNALS

THEORY AND PRACTICE

MAURICE BELLANGER

Head of Telephone Transmission Department
T.R.T Le Plessis-Robinson, France

Foreword by
Professor Pierre Aigrain
Secretary of State for Research, France

Translated by
Joan McMullan

A Wiley-Interscience Publication

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto · Singapore

Originally published under the title
Traitement Numérique Du Signal—
Théorie Et Pratique by Maurice Bellanger

Library of Congress Cataloging in Publication Data:

Bellanger, Maurice.
Digital processing of signals.

Translation of: Traitement numérique du signal.
'Wiley–Interscience publication.'
Includes index.
1. Signal processing—Digital techniques. I. Title.
TK5102.5.B3613 1984 621.38'043 83-17035

ISBN 0 471 90318 3

British Library Cataloguing in Publication Data:

Bellanger, Maurice
Digital processing of signals
1. Signal processing—Digital techniques
I. Title II. McMullan, Joan
III. Traitement numérique du signal. *English.*
621.38'043 TK5102.5

ISBN 0 471 90318 3

Typeset by Mid-County Press, London
Printed by The Pitman Press Ltd., Bath, Avon.

FOREWORD

In technology, the most important advances and those with the most significant consequences are not always obvious to the eventual users of the product. Modern methods for the digital processing of signals fall into this category as their consequences are still not generally recognized, and they do not hit the headlines in the popular press.

It is interesting to reflect briefly on how these techniques have evolved. Signal processing by digital calculation in the widest sense of the term is certainly not a new idea in itself. When Kepler developed his laws of planetary motion from the series of observations by his brother-in-law, Tycho-Brahé, he was actually studying digital signal processing, where the signal was the time series of Tycho-Brahé's observations of the planetary positions. However, it is only over the last few decades that digital signal processing has become a discipline in itself. The novel feature is that electrical signals can now be processed in real time using digital methods.

Before this could become possible, a sequence of technical developments had to occur in many fields. The most important of these was undoubtedly the fundamental need for the data to be available as an electrical signal. This led to the development of a wide range of devices that are sometimes called data transducers. These range in complexity from radar to the simple strain gauge, which required a considerable amount of research into the physics of solids before it became possible.

It was then necessary to develop the technological tools which allow arithmetic operations to be carried out at the very high rates required by real time processing. This was achieved through extensive progress in microelectronics. Today, it is quite common to use a microprocessor weighing only a few grammes and consuming milliwatts of electrical power, and with a mean time between failures of more than 10 years, to perform calculations which early computers took several hours to achieve, and often with frequent interruptions due to machine breakdowns. Indeed, the first electronic computer, ENIAC, was built only some forty years ago.

Finally, programming methods had to be refined and optimized because, whatever the impressive calculation capacities of modern microprocessors, their potential should not be wasted in performing redundant operations. The development of algorithms for fast Fourier transforms is one of the most striking examples of the importance of improved programming methods.

This convergence of technological progress in widely varying fields, whether dependent on physics, electronics or mathematics, has not occurred accidentally. To a certain degree, advances in one field created new demands which were met by advances in other disciplines. An examination of this relationship will undoubtedly have great implication for future studies in the history and epistemology of science and technology.

The immediate consequences of these developments are considerable. Analogue processing of electrical signals preceded digital processing, and will continue to be important in certain applications. However, the advantages of digital processing, paraphrased in the words 'accuracy' and 'reliability', have themselves encouraged applications far beyond the electronics and telecommunications fields from which they sprang. To quote only one example, X-ray tomography (or scanning) is based on a theorem derived by Radon in 1917. However, it was not until the developments outlined above were made that this new diagnostic medical tool could become a practical reality. It is certain that, in the future, digital signal processing techniques will be involved in a wider range of products, including those used by the general public, who, while benefiting from the resulting advantages in price, performance and reliability will not always be aware of the considerable infrastructure of research, technology and invention that was involved. This stage has already been reached in the domestic television market.

Almost inevitably, such a technological revolution leads to a further difficulty—user training. For the user, not only is a new tool involved, but also often a new way of thinking. If care is not taken, this training stage could become a major obstacle in the introduction of new techniques. This is why this book by M. Bellanger is important. It is based on a course given for several years at the École Nationale Supérieure des Télécommunications and at the Institut Supérieur d'Électronique in Paris. With its textbook style and its extensive exercises, this book will be frequently used, and will undoubtedly play a role in speeding the evolution of this important technological advance.

P. AIGRAIN

PREFACE

Innovation requires an engineer to give material form to his knowledge and to determine the potential offered by new techniques which have been discovered and developed in research laboratories. The use of digital techniques in signal processing opens up a wide range of potential advantages such as precise systems design, equipment standardization and stability of performance characteristics, as well as ease of monitoring and control. However, these techniques involve a certain degree of abstraction and their practical application demands a basis of theoretical knowledge, which is often considered to be more familiar, or at least more readily accessible, to the research worker than to the engineer. This can represent an obstacle to their utilization. It is the aim of this book to overcome this obstacle and to facilitate access to digital techniques by relating theory and practice, and by placing the most valuable results in this field within the reach of the working engineer.

The book is based upon a course given for several years at the École Nationale Supérieure des Télécommunications and the Institut Supérieur d'Électronique in Paris. The author has attempted to give a clear and concise discussion of the technical principles of digital processing, to compare the merits of the various techniques, and to present the most valuable results in a form suitable for immediate implementation in system design and for rapid evaluation of a project which is to be developed on a short time scale. Theoretical treatment has been reduced to that which is strictly necessary for good comprehension and for correct application of the results. Additional material can be found in the references to each chapter. At the end of each chapter there are several exercises which are often based on concrete examples. These allow the reader to test his understanding of the material in the chapter and to become familiar with its use.

The book is also intended for research workers, for whom it not only gives a collection of useful results, but indicates possible directions for their research by displaying the constraints of technological reality. Some results are also given which are taken from research papers by the author and his colleagues. In order to establish a dialogue with research workers and to be able to communicate their discoveries to the technology as quickly as possible, the engineer should be involved with the scientific work and should make his own contribution to research. Through his daily contact with the practical aspects of implementation, he can not only evaluate and consolidate the results of research, but can also open up new paths.

The work upon which this book is based was carried out with the support of the Centre National d'Études des Télécommunications, to whom the author would like to express his gratitude. He would also like to thank Mr J. Daguet, the Technical Director of the Société Télécommunications Radioélectriques et Téléphoniques, for his guidance and encouragement over many years. The author is also grateful to his colleagues for their contributions and for their constant support.

INTRODUCTION

Pocket calculators can be found today in every child's schoolbag. Thanks to a tiny electronic circuit, they are capable of performing the four basic numerical operations of addition, subtraction, multiplication and division. They can also store data in a memory, find percentages or even perform running totals. When such simple and familiar functions can be performed by tiny and inexpensive machines, it is not surprising that they should extend their influence from their original area of information processing to every other aspect of technology. Instrumentation systems and industrial production equipment are examples of how the search for precision and safety has led to digitally controlled machine tools. However, it is in signal processing that one of the best illustrations of the opportunities for digital techniques is to be found. Indeed, these techniques have led to simplifications in design and implementation that are essential for the widespread development of this supposedly difficult and abstract field. They reduce the most complicated functions to a sequence of elementary operations.

The signal is the carrier of intelligence in these systems. It transmits the commands in the control equipment, and it conveys words or images over information networks. It is particularly fragile and has to be handled with great care. The processing that it undergoes is intended to extract information, modify the message that it carries, or adapt it to the method of transmission. This is where digital techniques are involved. If a signal can be replaced by a set of numbers which are stored in a memory and represent its amplitude at suitably chosen instants of time, then the processing, even in its most complicated form, reduces to a sequence of arithmetic and logic operations on this set of numbers.

The conversion of a continuous analogue signal to a digital form is either performed by processors operating on recordings of the signal, or carried directly in the equipment which sends or receives the signal. The operations which follow this conversion are accomplished on suitably programmed digital computers. These techniques are applied in very different fields, and they are encountered in automation, industrial processes, aeronautics, radar systems, telecommunications, telemetry, medical instrumentation and geophysics.

Before introducing the various chapters of this book, it is appropriate to give a more precise discussion of the type of processing that is involved. The term digital signal processing is used to describe the complete set of operations, arithmetic calculations and numerical manipulations which are performed on the group of numbers representing the signal to be processed, in order to produce another set of numbers representing the processed signal. Many different functions can be performed in this way: spectral analysis, linear or non-linear filtering, transcoding, modulation, detection, and estimation and extraction of parameters. Digital computers are used in all of these applications and the processing obeys the laws of discrete systems. The numbers on which they operate can, in certain cases, be derived from a discrete process. However, more often they represent the amplitudes of samples of a continuous signal, and in this case, the computer follows an analogue–digital converter and, eventually, precedes a digital-analogue converter. Signal digitization is of fundamental importance in the design of such systems and in the study of their operation, and it is necessary to examine sampling and coding. Distribution theory forms a concise approach, which is simple and effective for this analysis. After giving some of the background to Fourier analysis, distributions and the representation of signals, the first chapter presents the results which are of most use in the sampling and coding of a signal.

Digital processing began with the development of algorithms for the fast calculation of the discrete Fourier transform. This transform is the basis for the study of discrete systems and is the digital equivalent of the Fourier transform in the analogue case. It is the mechanism for passing from a discrete time space to a discrete frequency space. It appears naturally in spectral analysis using a step in the frequency domain which is a factor of the sampling frequency of the signal to be analysed.

Fast calculation algorithms improve performance to such an extent that they allow real time calculations to be performed in many applications, provided certain basic conditions are fulfilled. Thus, the discrete Fourier transform not only forms a basic tool for determining the characteristics of a process and in its effect on the signal, but further, it gives a basis for the design of suitable equipment whenever a spectral analysis is involved, for example, in systems involving banks of filters or when, through the power of its algorithms, it leads to an advantageous approach for a filtering circuit. Chapters 2 and 3 are devoted to this subject. They present the elementary properties of the fast algorithms and the mechanics of their application. They also present a unified treatment of the algorithms and discuss possible developments. As for the system, the discrete Fourier transform computer is a discrete linear system, and is invariant in time.

Most of the present text is based on the study of one-dimensional time-invariant discrete linear systems. From the point of view of digital signal processing, these are the most important, the most readily accessible and the most useful. Multi-dimensional systems, in particular those with two dimensions, have undergone a certain degree of development. They are applied, for example, to image analysis; however, their properties are generally deduced from those of one-dimensional systems, of which they are often only simple extensions. Non-linear systems or systems which are variable in time either contain a large sub-set which has the properties of linearity and time invariance, or one which can be analysed using the same techniques.

Linearity and time invariance involve the existence of a convolution relation which governs the operation of the system, or involve a filter with the same properties. This convolution equation is defined using the response of the system to an elementary impulse, the impulse response. Thus, if $x(t)$ denotes the signal to be filtered and $h(t)$ is the impulse response of the filter, the filtered signal $y(t)$ is given by the equation:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\,\mathrm{d}\tau$$

This type of equation, even though it directly represents the real operation of the filter, is of limited interest in practice. This is because it is difficult to détermine the impulse response from the criteria which define the intended filtering operation. Also, an equation involving an integral does not allow the behaviour of the filter to be readily discovered or verified. The design is much more easily obtained in frequency space because the Laplace or Fourier transforms allow access to a transformed space in which the convolution equations of amplitude–time space become simple products of functions. Fourier transformation produces the frequency response which corresponds to the impulse response of the system, and filtering amounts to producing this frequency response through the Fourier transform, or the spectrum, of the signal to be filtered.

In digital systems, which are discrete in character, convolution is performed by summation. The filter is defined by a set of numbers which defines its impulse response. Thus, if the set to be filtered is $x(n)$, the filtered set $y(n)$ is expressed by the following summation, where n and m are integers:

$$y(n) = \sum_{m} h(m)x(n-m)$$

Two cases are then possible. Firstly, the summation can be performed over a finite number of terms, that is to say, all the $h(m)$ are zero except for a finite number of values of the integer variable m. The filter is said to have a finite impulse response. By reference to its construction, it is also said to be non-recursive, because its implementation does not require a feed-back loop from the output to the input. It has a finite memory, as it only stores an elementary signal, such as an impulse, for a short time. The numbers $h(m)$ are called the coefficients of the filter, and define it completely. They can be calculated in a very simple direct manner, for example, by performing the Fourier series expansion of the frequency response to be achieved. This type of filter has some very important basic characteristics. For example, it is possible to create a strictly linear phase response, i.e., a constant group delay. Signals whose components are in the pass band of the filter are not deformed when passing through it. This potential is used, for example, in data transmission systems or in spectral analysis.

Secondly, the summation can be performed over an infinite number of terms, as there are an infinite number of non-zero $h(m)$. The filter is said to have an infinite impulse response or to be of the recursive type, because it requires a feed-back loop from the output to the input. Its operation is governed by an equation according to which an element of the output set $y(n)$ is calculated by the weighted summation of a certain number of elements of the input set $x(n)$ and of a certain number of elements of the preceding output set. For example, if L and K are integers, the operation of the filter can be defined by

$$y(n)=\sum_{l=0}^{L} a_l x(n-l)-\sum_{k=1}^{K} b_k y(n-k)$$

The numbers a_l $(l=0, 1, \ldots, L)$ and b_k $(k=0, 1, \ldots, L)$ are the coefficients. As with analogue filters, this type of filter cannot generally be analysed directly. It is necessary to change to a transformed space using, for example, the Laplace or Fourier transforms. However, for discrete systems, a more suitable equivalent transform, the Z-transform, is available. A filter is described by its Z-transfer function, generally denoted by $H(Z)$, which involves the two sets of coefficients:

$$H(Z)=\frac{\sum_{l=0}^{L} a_l Z^{-l}}{1+\sum_{k=1}^{K} b_k Z^{-k}}$$

The frequency response of the filter is obtained by replacing the variable Z in $H(Z)$ by the following expression, where f denotes the frequency and T is the sampling period of the signals:

$$Z=e^{j2\pi fT}$$

In this operation, the imaginary axis in the Laplace plane corresponds to a circle of unit radius, centred on the origin in the Z-space. It is readily apparent that the frequency response of the filter defined by $H(Z)$ is a periodic function with the sampling frequency as its period. Another representation of the function $H(Z)$ is useful for the design of filters and involves the roots of the numerator, called the zeros of the filter, Z_l $(l=1, 2, \ldots, L)$, and the roots of the denominator, called the poles, P_k $(k=1, 2, \ldots, K)$:

$$H(Z)=a_0\frac{\prod_{l=1}^{L}(1-Z_l Z^{-1})}{\prod_{k=1}^{K}(1-P_k Z^{-1})}$$

The term a_0 is a scale factor which defines the gain of the filter. The condition for the stability is expressed very simply by the constraint that all the poles should be inside the unit circle. The positions of the poles and the zeros relative to the unit circle allow a very simple and commonly used representation of the filter characteristics.

A group of four chapters is devoted to the study of the properties of these digital filters. Chapter 4 discusses the properties of discrete time-invariant linear systems, reviews the principal properties of the Z-transform, and gives the necessary elements for studying the filters. Chapter 5 discusses finite impulse response filters. Their properties are considered, together with the techniques for calculating the coefficients, and the structures for their realization. As infinite impulse response filters are generally formed as a cascade arrangement of fundamental elements of the first and second order, Chapter 6 describes these elements and their properties. These elements considerably simplify the approach to this type of system and provide a set of results which are of use in practical applications. Chapter 7 describes the methods for calculating the coefficients of infinite impulse response filters and discusses the problems of their realization, with the limitations that are implied, and their consequences, in particular for round off noise.

As infinite impulse response filters have properties comparable to those of continuous analogue filters, it is natural to consider their realization in structures of the same type as are presently used in analogue filtering. This is the subject of Chapter 8, which describes ladder structures. A digression is made into switched capacitor devices, which are not digital in the strict sense of the term, but which are nevertheless of the sampled type and are very useful complements to digital filters. As a guide for the user, a resumé of the relative merits of the structures described is given at the end of the chapter.

Some equipment, as used for example in spectral analysis or in the telecommunications field, involves signals represented by a set of complex numbers. Analytic signals are an important example of particular practical interest. Their properties are discussed in Chapter 9, along with the design of devices for the production or processing of such signals. Some complementary concepts in filtering, such as the minimum phase shift condition, are also discussed in this chapter.

If the machines for digital signal processing are to operate in real time, then they must operate at a rate which is closely related to the sampling frequency of the signals. Thus, their complexity depends on the number of operations to be performed and on the time interval available for their completion. The signal sampling frequency at the input or the output of the system is generally imposed by other constraints, but it can be varied within the system itself to match the properties of the signal and the processing, thus reducing the number of operations and the calculation rate. One simplification to the machines that can be of great importance is obtained by changing the sampling frequency to match the band-width of the useful signal throughout the processing. This is multirate filtering, which is discussed in Chapter 11. The effects on the processing characteristics are described, together with methods for their realization. Rules for their use and evaluation are provided. This technique yields particularly interesting results for narrow pass band filters or for banks of filters. In this latter case, the system combines the discrete Fourier transform computer with a set of phase shift circuits.

Filters can be determined from a time specification, as would be the case, for example, when modelling the behaviour of a system. As the properties vary with time, the filters should adapt to match the evolution of the system to be modelled. This adaptation depends on an approximation criterion and is carried out at a rate up to the sampling rate of the system. Such a filter is said to be adaptive. Chapter 11 is devoted to simple adaptive filters, of the most common type, in which the coefficients vary according to the gradient algorithm. The approximation criterion is the least mean square of the error signal. After a brief review of the properties of random signals, and particularly their auto-correlation matrix whose eigenvalues play an important role, the gradient algorithm is introduced and the convergence conditions are studied. The two major adaptation parameters, the time constant and the residual error, are then analysed, together with the arithmetic complexity. Several implementation structures are discussed, and finally the important case of linear prediction is discussed.

Circuits and technological considerations are dealt with in Chapter 12. The elements required are an arithmetic unit to perform the calculations, an active memory to store the data, the intermediate results and the variable parameters, a read-only memory for fixed parameters such as the coefficients, and a control unit which coordinates the operation of the other components. The architecture of the system depends heavily on the arithmetic operator. This can perform operations on numbers presented either in serial form, with the binary digits (bits) representing the numbers at each input and output appearing one after the other, or in parallel form, where the bits are presented together simultaneously. The circuits are described in this chapter with examples of their construction using large-scale integration techniques for circuits which are specifically designed for digital signal processing. As the technology advances, the trend is towards the concentration of the functions and the use of programmable processors. Complexity parameters are introduced to give guidance in systems design or in the development of specific projects.

Telecommunications equipment represents a preferred area for the application of digital processing, as transmission networks are progressively digitized and as digital computers are introduced at the terminals. Chapter 13 presents a group of typical applications illustrating the theory and the techniques presented in the book as a whole.

CONTENTS