NBS TECHNICAL NOTE 1047

U.S. DEPARTMENT OF COMMERCE / National Bureau of Standards

Deconvolution of Time Domain Waveforms in the Presence of Noise

NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, and the Institute for Computer Sciences and Technology.

THE NATIONAL MEASUREMENT LABORATORY provides the national system of physical and chemical and materials measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; conducts materials research leading to improved methods of measurement, standards, and data on the properties of materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

Absolute Physical Quantities[2] — Radiation Research — Thermodynamics and Molecular Science — Analytical Chemistry — Materials Science.

THE NATIONAL ENGINEERING LABORATORY provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

Applied Mathematics — Electronics and Electrical Engineering[2] — Mechanical Engineering and Process Technology[2] — Building Technology — Fire Research — Consumer Product Technology — Field Methods.

THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

Programming Science and Technology — Computer Systems Engineering.

[1]Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted; mailing address Washington, DC 20234.
[2]Some divisions within the center are located at Boulder, CO 80303.

Deconvolution of Time Domain Waveforms in the Presence of Noise

N. S. Nahman
M. E. Guillaume

Electromagnetic Technology Division
National Engineering Laboratory
National Bureau of Standards
Boulder, Colorado 80303

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, Secretary
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director

Issued October 1981

NATIONAL BUREAU OF STANDARDS TECHNICAL NOTE 1047

Nat. Bur. Stand. (U.S.), Tech. Note 1047, 122 pages (October 1981)
CODEN: NBTNAE

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1981

CONTENTS

Page

FOREWORD

The work reported here is the result of a continuing collaboration between the Centre National D'Etudes Des Telecommunications, France, and the National Bureau of Standards.

From October 1978 through September 1979, Dr. Michele Guillaume was a Guest Scientist in the Time Domain Metrology Group of the Electromagnetic Technology Division, NBS, Boulder Laboratories, Boulder, Colorado. At that time Drs. Guillaume and Nahman worked together to produce the experimental results reported here. It has taken some time for the authors to complete their analyses of the experimental data, to provide some analytic bases for understanding the deconvolution and filtering processes, and to prepare this manuscript. All of these latter activities were accomplished with the authors located in France and the USA, respectively. Thus, there has been some delay in publishing this work.

DECONVOLUTION OF TIME DOMAIN WAVEFORMS IN THE PRESENCE OF NOISE

N. S. Nahman and M. E. Guillaume †
Electromagnetic Technology Division
National Bureau of Standards
Boulder, Colorado 80303

Deconvolution or inverse filtering was used to determine the impulse response of a system using noisy input and output time domain sequences (discrete data). Frequency and time domain methods were studied along with the synthesis of the filters required to obtain stable and smooth results. For the methods studied it was concluded that the superior technique was provided by an optimal frequency domain method implemented via the FFT. Also, it is pointed out that the time domain methods are only in their infancy and still retain the promise of avoiding transform domain filtering. Examples are presented in which the impulse responses are determined in the presence of varying degrees of noise for a coaxial transmission line, a wave-shaping filter, and a broad-band antenna.

Key words: deconvolution; impulse response; inverse filtering; noise contamination; time domain; waveforms.

1. INTRODUCTION

The analogy between a physical linear system and the convolution equation is well known. If $\bar{x}(t)$ is the input waveform to a linear device whose impulse response is $\bar{h}(t)$, the output waveform is described using the convolution equation,

$$\bar{y}(t) = \int_{-\infty}^{+\infty} \bar{x}(\tau)\, \bar{h}(t-\tau)dr = \int_{-\infty}^{\infty} \bar{x}(t-\tau)\bar{h}(\tau)\, d\tau \qquad (1\text{-}1)$$

or in simplified notation

$$\bar{y}(t) = \bar{x}(t) * \bar{h}(t) \qquad (1\text{-}2)$$

where "*" denotes convolution. The Laplace transform of $\bar{y}(t)$ is given by

$$\bar{Y}(s) = \bar{X}(s)\, \bar{H}(s) \qquad (1\text{-}3)$$

which shows that the convolution process transforms to a product operation in the s-domain. The deconvolution process is defined as the solution of the convolution integral equation (1-1) when either $\bar{x}(t)$ or $\bar{h}(t)$ is unknown, i.e.,

$$\bar{h}(t) = \bar{y}(t)\ (1/*)\ \bar{x}(t) \qquad (1\text{-}4)$$

$$\bar{H}(s) = \bar{Y}(s) \div \bar{X}(s) \qquad (1\text{-}5)$$

† Centre de Microelectronique de Grenoble, Centre National d'Etudes de Telecommunications, Meylan, France.

or

$$\bar{x}(t) = \bar{y}(t)\ (1/*)\ \bar{h}(t) \tag{1-6}$$

$$\bar{X}(s) = \bar{Y}(s) \div \bar{H}(s) \tag{1-7}$$

Note that (1/*) which denotes deconvolution transforms to a division operation. Both cases, i.e., the solution for $\bar{h}(t)$ or $\bar{x}(t)$, are basically the same with respect to the deconvolution problem.

The above convolution integral relationship is defined for continuous waveforms. In the problems considered here the waveform data are not continuous but are discrete; the data are obtained from a sampling process followed by an analog to digital conversion.

For the purposes of this present work, only discrete variables or sequences are considered. When the question of error is considered, it is relative to the differences between one sequence and another. The error due to representing a continuous function by a sequence is not considered here. However, it should be recognized that such an error (aliasing) can be very significant when spectral information is desired [1,2].[1]

A sequence of N-values for x, x(0), x(1), x(2), ... x(N-1), is formally denoted as $\{x(k)\}$ where the k-th number in the sequence is x(k),

$$x = \{x(k)\} \tag{1-8}$$

For convenience, here the notation x(k) will denote the sequence of values for x.

For discrete variables, i.e., sequences, the convolution and deconvolution equations (1-1) through (1-7) are respectively replaced by

$$y(k) = \sum_{i=o}^{k-1} x(i)h(k-i) = \sum_{i=o}^{k-1} x(k-i)h(i) \tag{1-9}$$

$$y(k) = x(k)*h(k) \tag{1-10}$$

$$Y(s) = X(s)\ H(s) \tag{1-11}$$

$$h(k) = y(k)\ (1/*)\ x(k) \tag{1-12}$$

$$H(s) = Y(s) \div X(s) \tag{1-13}$$

$$x(k) = y(k)(1/*)h(k) \tag{1-14}$$

$$X(s) = Y(s) \div H(s) \tag{1-15}$$

In the remainder of this report, for the most part, the analyses will be carried out in terms of discrete variables. Where continuous variables are used, the distinction will be clearly made. However, in the figures, for convenience in drawing the discrete waveforms and spectra, they will be drawn as continuous functions, but labeled with the appropriate discrete variables.

[1] Numbers in brackets indicate literature references at the end of this paper.

In the experimental situation the imperfections of the acquisition device and the measurement system noise which is added to the signal make the available data not representative of the original continuous waveform but of something close to it. The convolution relationship still holds; but, in general, the deconvolution may no longer be stable. That is to say, the solution which comes from the direct brute-force application of the mathematical equations may not represent the desired signal. Furthermore, an exact solution does not exist and must be replaced by an estimate. The problem is to find an approximate solution which is stable under small changes in the initial data. To do that one uses a regularizing operator or filter. This filter depends on the characteristics of the processed waveforms and on the method used to perform the deconvolution. Particular emphasis is given to the fact that the experimenter controls the operation and is the judge of the quality of the deconvolution result.

In this report after a general presentation of the deconvolution problem, two possible solution methods are examined. In each method the reasons for the instability are inspected and lead to the synthesis of the corresponding regularizing function. The experimental results include the evaluation of the quality of the waveform obtained as a solution of the deconvolution process.

Specifically, this report is divided into six chapters, the first being the present introduction. Chapter 2 is devoted to a general presentation of the deconvolution problem and possible approaches for solution.

Chapter 3 describes the experimental equipment used to obtain waveform data and the various experimental quantitative criteria which may be used to measure the quality of a solution to a deconvolution problem.

Chapter 4 discusses frequency domain deconvolution methods and presents detailed examples of solutions to deconvolution problems using frequency domain processing which include the synthesis of regularization filters.

Chapter 5 discusses time domain deconvolution methods and presents several examples, using time domain processing, one illustrating a time domain solution to a deconvolution problem which was also solved in Chapter 4 using frequency domain methods. The synthesis of regularizing filters for time domain deconvolution is also discussed.

Chapter 6 presents a summary of the deconvolution methods considered and some general conclusions about the solution to the deconvolution problem.

2. GENERAL PRESENTATION OF THE DECONVOLUTION PROBLEM

2.1 An Exact Solution Exists for Exact Waveforms

Consider the linear system shown in figure 2-1; if the input waveform $x_o(k)$ and the measurement system impulse response $h_o(k)$ are assumed to be exactly known, then the exact output waveform $y_o(k)$ is also known and is given by the convolution equation,

$$y_o(k) = x_o(k) * h_o(k). \tag{2.1-1}$$

Furthermore, if $h_o(k)$ is the unknown, the deconvolution

$$h_o(k) = y_o(k)\ (1/*)\ x_o(k) \tag{2.1-2}$$

is a unique solution which exists only because $x_o(k)$ and $y_o(k)$ represent exact, error free waveforms (sequences).

2.2 Experimental Situation Yields Inexact Waveforms

Practically, the exact waveforms $x_o(k)$ and $y_o(k)$ are not available from measured experimental data, because the exact values are corrupted by system noise and sampling errors. Consequently, the experimental data x(k) and y(k) replace the exact data $x_o(k)$ and $y_o(k)$ respectively. However, each representation, x(k) or y(k), belongs to a family of waveforms defined as "close" to the exact waveform; that is to say, x(k) belongs to the family [X],

$$x \in [X], \tag{2.2-1}$$

such that for all elements x of this family,

$$\rho(x, x_o) \leq \delta_x \tag{2.2-2}$$

where ρ being defined for any kind of waveforms w and w_o by the relation,

$$\rho(w,w_o) = \left[\frac{1}{N} \sum_{1}^{N} (w_i - w_{oi})^2\right]^{1/2} \tag{2.2-3}$$

The parameter δ may be interpreted as a measure of the error which is made when the waveform w_o is represented by w. Then for

$$x \in [X],\ \rho(x,x_o) \leq \delta_x \tag{2.2-4}$$

$$y \in [Y],\ \rho(y,y_o) \leq \delta_y \tag{2.2-5}$$

These two relationships define families of waveforms such that "small changes" are allowed between the elements of each family.

The direct use of the deconvolution equation (1-3) when applied to the y(k) and x(k) waveforms leads to a result h(k) which, in general, is not close to $h_o(k)$ the exact solution, i.e., if [H] is the family of waveforms close to $h_o(k)$, then

$$h(k) \notin [H] \tag{2.2-6}$$

and

$$\rho(h, h_o) > \delta_h \tag{2.2-7}$$

hence, the solution is not stable under small changes in the initial data.

A strict mathematical solution is just not possible because, in general, the result is not stable [3]. The problem is then to find an approximate solution stable under small changes in the initial data. The requirement of stability has to be satisfied since it is associated with the physical determinancy of the phenomenon described by the convolution equation. A rule has to be chosen for the definition of the family [H] of possible solutions. In practice this choice is arbitrary; consequently, one needs to use supplementary information concernin the solution.

However, it is known that the convolution itself is stable under small changes in the initial data. If h(k) is the obtained approximate solution, its convolution with the known input waveform x(k) must give a result which belongs to the family [Y] of waveforms close to $y_o(k)$. If

$$c(k) = h(k) * x(k) \tag{2.2-8}$$

and

$$\rho(c,y) \leq \delta_y;\ c \in [Y] \tag{2.2-9}$$

then

$$e(k) = y(k) - c(k) \tag{2.2-10}$$

defines an error criterion whose properties will be used to judge the quality of the representation of $h_o(k)$ by h(k). Moreover, it may be assumed that the approximate solution has to be smooth because the solution represents a continuous deterministic signal or system (in sampled data discrete form.) This qualitative information will be used in addition to the error criterion to control the process which leads to the accepted solution.

2.3 The Regularization Operator

Another way which employs linear algebra, i.e., matrix methods, may be used to describe the convolution process. The convolution

$$y(k) = h(k) * x(k)$$

may be written as

$$y(k) = A\ x(k) \tag{2.3-1}$$

where A is an operator which maps the family [X] into the family of waveforms [Y]. It has already been pointed out that the inverse operation corresponding to the operator A^{-1}, which was expected to map [Y] in [X], is not stable. Hence, A^{-1} has to be replaced by a new operator R which really maps [Y] in [X] such that

$$x(k) = R\ y(k). \tag{2.3-2}$$

R is called the "regularized operator". Its expression depends upon the properties of the two families of waveforms [X] and [Y]. These properties include the precision in the knowledge of a waveform when compared to the exact waveform (parameter δ) and the shape of the waveform itself. On the other hand, the method used to perform the deconvolution determines the practical effect of the noise and influences the definition of the regularized operator. From the user's point of view, the properties of R must be easy to modify in such a way that a progressive improvement of the regularization finally leads to the accepted estimate solution.

Functionally,

$$R \equiv R\ \{[X],\ [Y],\ \alpha(\delta)\} \tag{2.3-3}$$

where α is a set of parameters which are used to take into account the different possible noise figures.

2.4 Possible Methods of Deconvolution

There are many different methods for solving the deconvolution equation [4,5]. In this section three methods are briefly discussed in order to provide a broader perspective of deconvolution methods. In the main part of this report only two of the three methods will be discussed. In practice, all methods of deconvolution (including the three) have a common feature in that the amount of information that can be extracted from the empirical curves (physical data) is limited. Consequently, some form of approximation is a necessary part of the solution.

2.4.1 Model Fitting

Model fitting consists in finding an analytical expression to represent the exact waveforms $y_o(k)$ and $x_o(k)$ before the deconvolution is performed e.g., Prony's Method [6,7]. In this case, from the inexact available data it is necessary to obtain a mathematical expression which is assumed to represent the exact solution. This operation has to be made for each waveform with the additional condition that the two obtained waveforms have to be compatible in the convolution process, i.e., that y'_o (assumed exact representation of y(k) must be the image of x'_o (assumed exact representation of x(k)) in the convolution operation. This method, model fitting, is not examined in this present experimental study.

2.4.2 Time Domain Processing

The problem is solved using only the time domain representation. The regularizing operator is also defined in the time domain. This class of methods includes the inversion of the convolution matrix and the iteration technique.

2.4.3 Transform Domain Processing

Transform domain processing methods include the method of moments (statistical) [4], the homomorphic deconvolution [8] and the Fourier transform deconvolution. The latter method is considered here. Generally, in transform domain processing the available data are transformed into another representation (moments, frequency, etc.) and the regularizing operator is defined in this new domain before the return to the time domain.

2.5 Loss of Information in the Convolution Process

The details of the observed waveform are smoothed by the convolution with the apparatus function. It is impossible to distinguish between the effects of the noise and very small irregularities due to the signal itself. Moreover, the amount of information available in the output waveform must be comparable to the amount of information needed for the description of the input waveform (or impulse response). This last point is very important and may be explained using a frequency domain presentation.

First of all, when the duration of the unknown impulse response waveform is smaller than that of the input waveform, figure 2.2, the corresponding discrete frequency domain spectrum magnitudes are as shown in figure 2.3. It is evident from figure 2.3 that the discrete spectra X(n) and Y(n) do not contain enough high frequency components to allow a complete description of H(n). It is then impossible to determine H(n) or h(t) unless additional information is used.

When the duration of the unknown impulse response is larger than that of the input pulse, x(t) and y(t) contain enough information to determine h(t), figures 2-4 and 2-5. Consequently, a possible way to remove the fundamental difficulty illustrated by figures 2-2 and 2-3, is to use an input signal x(t) of lesser duration. When the input pulse duration is less than that of the system impulse response, the only problem which remains is to find a way to manage the effects of the noise which always is present. Such management is called a regularizing process or regularization.

In conclusion of this chapter on the general presentation of the deconvolution problem, it is important to reiterate (in slightly different words) a point made in an earlier study [5]:

> "A strict mathematical solution cannot be obtained because of the ever present noise in the original data, but the correlation between the mathematics and the physical data must always be kept in mind to reach an acceptable estimated solution".

The effect on the final solution of the uncertainty of the initial data due to noise will depend upon the method used for the deconvolution. Two methods will now be studied in detail: the Fourier transform processing method and the time domain processing method. In the frequency domain, it will be seen that the disturbing effect of the noise appears in the high frequency range. In the second method, i.e., in the time domain method, it will be seen that the noise increases the value of the difference between two consecutive samples, and this random deviation from the exact values causes instability in the solution. Furthermore, the time domain method is very sensitive to the D.C. value or offset of the waveform; this leads to difficulty in using the method.

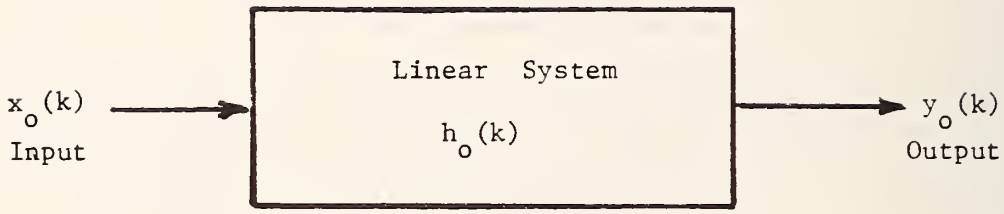


Figure 2.1 Linear System Possessing an Impulse Response $h_o(k)$.

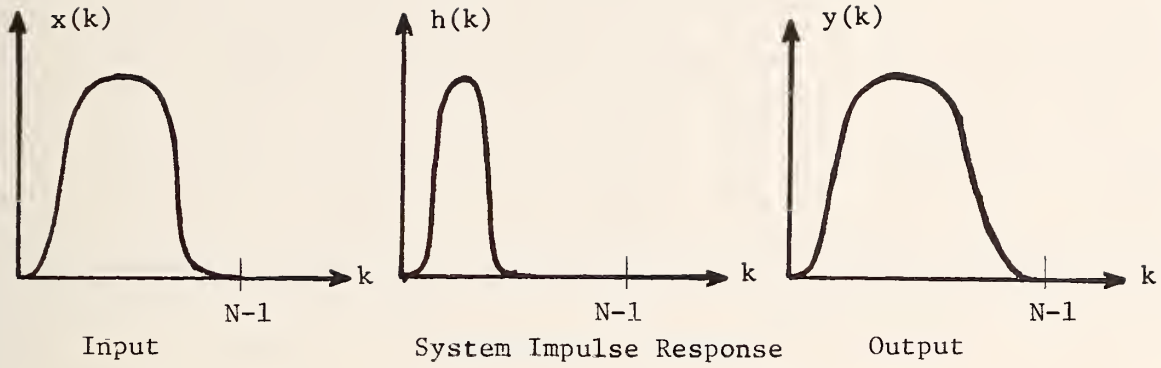


Figure 2.2 The unknown h(k) convolved with x(k) where the duration of h(k) is less than that of x(k).

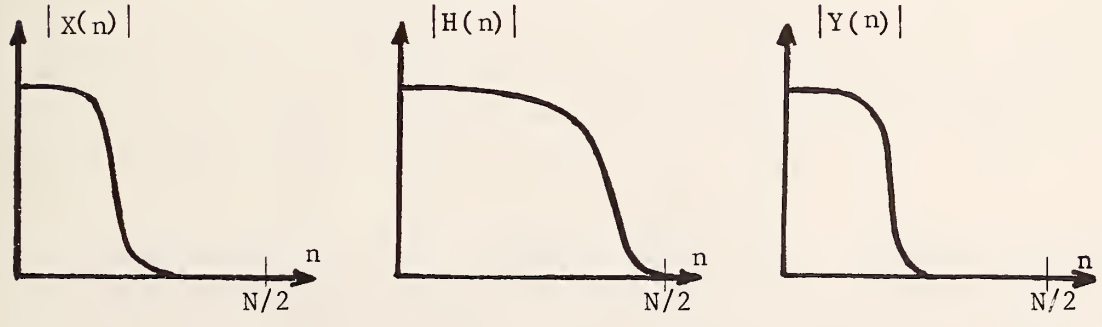


Figure 2.3 The frequency domain spectrum magnitudes corresponding to x(k), h(k) and Y(k) of figure 2.2.

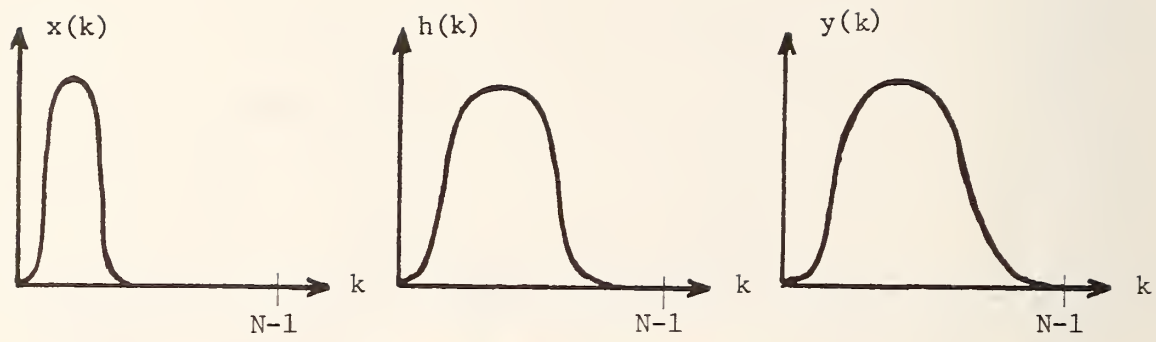


Figure 2.4 The unknown h(k) convolved with x(k) where the duration of h(k) is greater than that of x(k).

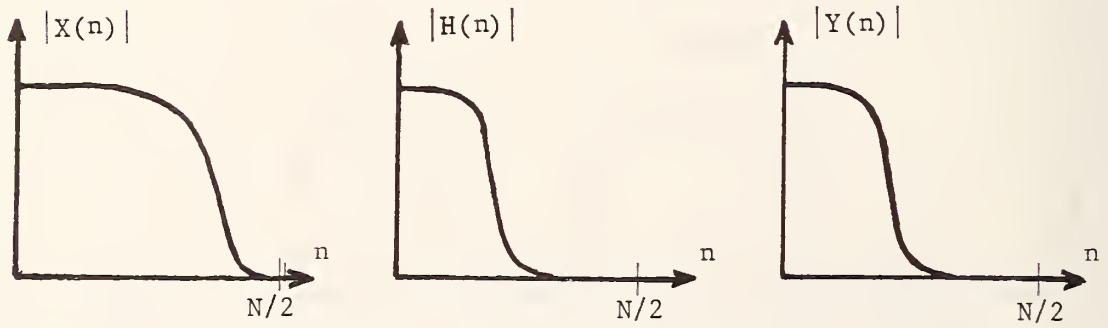


Figure 2.5 The frequency domain spectrum magnitudes corresponding to x(k), h(k) and y(k) of figure 2.4.

3. EQUIPMENT AND METHODS USED IN THE EXPERIMENTS

3.1 Acquisition of Waveforms

A minicomputer controlled sampling oscilloscope system was used to implement the deconvolution experiments. The specific system employed was the NBS Automatic Pulse Measurement System (APMS) [9,10]. The oscilloscope bandwidth is very large (dc-18 GHz) and permits the observation of picosecond domain signals. A block diagram of the APMS is shown in figure 3.1.

In operation, a repetitive signal synchronized to the oscilloscope time base is applied to the sampling head. The time at which the samples are taken is controlled by the minicomputer (using the digital to analog converter). The amplitude of each sample is digitized and stored in the memory. The effect of the noise which always accompanies the signal may be reduced by using either analog or digital averaging techniques. The APMS can only acquire a limited finite amount of data (points); consequently, a time window of finite extent is used to observe the waveform. The acquisition process is fast and provides a large number of data points (1024 for example) which are stored in an array whose format is well adapted for further digital processing. The CRT terminal provides a powerful tool which allows the operator to implement, control, and monitor signal acquisition, conditioning and deconvolution processing.

3.2 Interactive Processing

Because no exact solution exists, it is difficult to make the deconvolution process entirely automatic. For that, one would have to use a general mathematical criterion whose definitions could hardly take into account each particular configuration of available data. The resulting solution would not be so precise as the one which is obtained if the judgment of the operator is taken as a part of the process. With the operator involved in the process the graphical capability of the CRT display is extensively used to maintain the contact between the operator and the digital processing. Each time a decision is made its effect appears in the display permitting a measurement of its effectiveness or suitability.

3.3 Qualitative Measurement of the Quality of One Solution

The operator uses all the available qualitative information. He may have a preconceived idea of the shape of the expected waveform. In this case he may recognize that the obtained result is close or far off the solution he expects. Often, the only usable criterion is the smoothness of the deconvolved waveform which results from the deconvolution process.

3.4 Quantitative Criteria

Because different methods may be used for the achievement of the deconvolution, the quality of the result has to be measured using a quantitative error function. This criterion permits a choice between several possible solutions. The error function [e] is defined by the relationship, figure 3-2.

$$[e] = [y] - \{[d] * [x]\} \tag{3.4-1}$$

[y] and [x] are the two known waveforms (output and input), and [d] is the approximate solution of the deconvolution

$$[d] = [y] \ (1/*) \ [x]. \tag{3.4-2}$$

If the deconvolution process is perfect (exact solution) all samples of the error function [e] will be equal to zero.

The use of this error function is justified by the following remarks. Remember that the convolution operator is stable under small changes in the initial data. If the obtained result [d] belongs to the family [H] of waveforms close to the exact solution [h_o], the result [y'] of its convolution by the input waveform must belong to the family [Y] of waveforms close to the exact output. But the exact output [y_o] is not known. Then [Y] is used as the reference for the definition of the error function (3.4-1).

The error function must be carefully interpreted. Four parameters will be used for its characterization, the mean value $\bar{e}$, the standard deviation σ, the maximum value e_{max} and the minimum value e_{min}. The mean value is given by

$$\bar{e} = \frac{1}{N} \sum_{k=1}^{N} e(k) \tag{3.4-3}$$

Ideally $\bar{e}$ should be zero, hence in practice a value as close to zero as possible is sought. The standard deviation is given by

$$\sigma = \left[\frac{1}{N} \sum_{k=1}^{N} [e(k) - \bar{e}]^2 \right]^{1/2} \tag{3.4-4}$$

and must be small when compared to the maximum value of the reference waveform [y]. A typical acceptable value for σ is when it is the same order of magnitude as the noise on the output waveform [y]. Experimentally, it has been shown that the maximum and minimum values of the error, e_{max} and e_{min} occur for some samples associated with the largest values of [y], figure 3.2; also, for a well chosen regularizing function $|e_{max}|$ and $|e_{min}|$ will be of the same order of magnitude. Finally, it should be kept in mind that in the course of the deconvolution process, the goal is to make the error parameters $\bar{e}$, σ, e_{max} and e_{min} very small compared to the maximum value of the reference waveform [y].

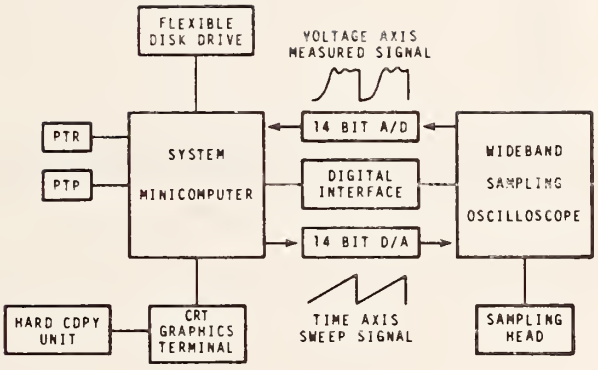


Figure 3.1 The NBS Automatic Pulse Measurement System (PTP and PTR are paper-tape-punch and paper-tape-reader, respectively).

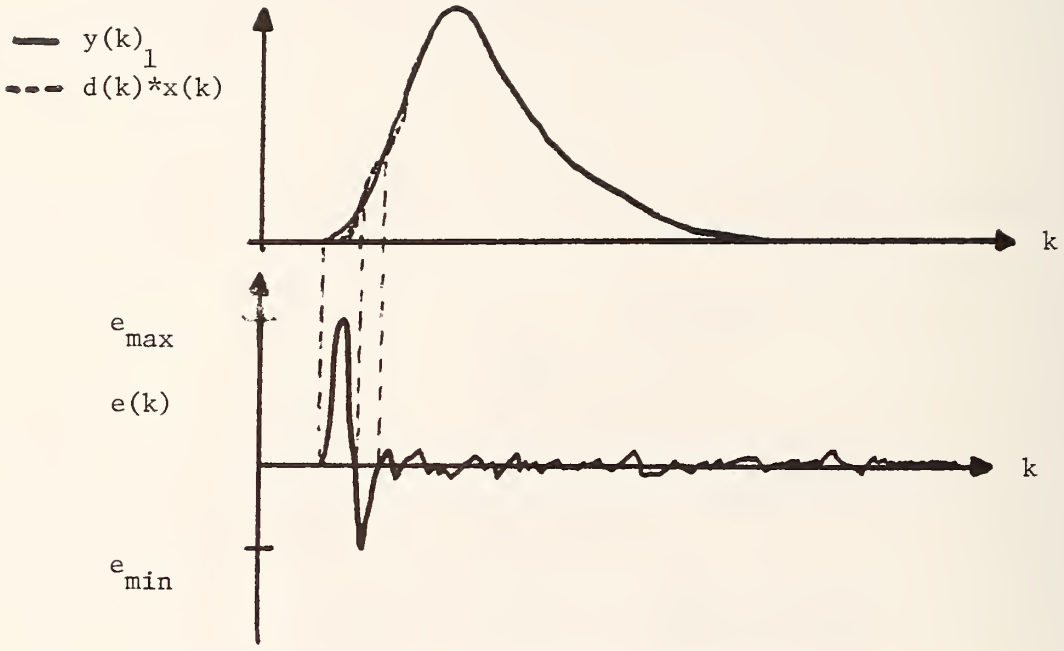


Figure 3.2 The error e(k) is the difference between the reference function y(k) and the estimated function d(k) * x(k).

4. FREQUENCY (TRANSFORM) DOMAIN DECONVOLUTION

4.1 Introduction

Frequency domain deconvolution methods operate indirectly on the time domain sequence x(k); they operate on the transform sequence $\overline{X}(n)$ which is the Discrete Fourier Transformation (DFT) [11,12] of the time domain sequence. The DFT and its inverse IDFT are defined in the following way. The sequence X(k), i.e.,

$$x(kT) = x(0), x(T), x(2T), \ldots x[(N-1)T] \tag{4.1-1}$$

is sequence of values spaced apart by an interval T. The DFT of x(k) is a sequence of complex values given by

$$X(n\,\Omega) = \sum_{k=0}^{N-1} x(kT)\, e^{-j\Omega\, Tkn}, \quad n = 0, 1, 2 \ldots N-1 \tag{4.1-2}$$

where each value is spaced apart by an interval Ω, where $\Omega T = 2\pi/N$. The inverse transformation IDFT is given by

$$x(kT) = \frac{1}{N} \sum_{n=0}^{N-1} X(n\,\Omega)\, e^{j\,\Omega\, Tkn}, \quad k = 0, 1, 2 \ldots N-1 \tag{4.1-3}$$

As indicated above, the notation x(k) and X(n) implies x(kT) and X(n Ω), respectively.

Errors introduced into the deconvolution process stem from three sources:

1. In the time domain - erroneous or inconsistent data and computation errors in computing the error.
2. Transform Domain - filtering and computation errors.
3. Transformation Operation - computation errors.

As summarized in Chapter 2, noise or errors in the time domain data cause the data to be inconsistent in the sense that the input x(k) and output y(k) data (i.e. the two known quantities) are not related to each other through a causal stable impulse response. For example, if for a given causal stable system $x_o(k)$, $y_o(k)$, and $h_o(k)$ are the input, output and system impulse response sequences, respectively, then the sequences are uniquely related by the convolution equation (1-8).

$$y_o(k) = x_o(k) * h_o(k) \tag{4.1-4}$$

Consequently, $x_o(k)$ and $y_o(k)$ together provide the data to exactly characterize the system $h_o(k)$, figure 4.1. Now assume that in the measurement of $x_o(k)$ and $y_o(k)$ the independent noise sources, n_x and n_y, corrupt $x_o(k)$ and $y_o(k)$, respectively, to yield x(k) and y(k), figure 4.2. Clearly, x(k) and y(k) are different from $x_o(k)$ and $y_o(k)$, respectively, Furthermore, in general x(k) and y(k) are not related to each other through (4-4) and cannot provide $h_o(k)$ through the deconvolution process as specified by

$$h_o(k) = y_o(k)(1/*)\, x_o(k) \tag{4.1-5}$$

which specifically employs $x_o(k)$ and $y_o(k)$ and not x(k) and y(k), respectively. However, the impulse response

$$h(k) = y(k)(1/*)\, x(k) \tag{4.1-6}$$

can be calculated. This result may be stable or unstable, and if stable, it may resemble $h_o(k)$ or a very noisy $h_o(k)$.

The impulse response of a physical system is causal (zero for time less than zero, the time at which the impulse is applied) and stable (returns to zero after the passage of time), $d_1(t)$ in figure 4.3.

By operating on the data sequences with the DFT the discrete Fourier transforms of x(k) and y(k) can be obtained; these transforms are the complex sequences,

$$\text{DFT}[x(k)] = X(n) \tag{4.1-7}$$

$$\text{DFT}[y(k)] = Y(n) \tag{4.1-8}$$

Also, it is true that the exact processes (4.1-4) and (4.1-5) have the DF transforms

$$Y_o(n) = X_o(n)\, H_o(n) \tag{4.1-9}$$

$$H_o(n) = Y_o(n)/X_o(n) \tag{4.1-10}$$

but the ratio Y(n)/X(n) still does not represent $H_o(n)$, i.e.

$$H_o(n) \neq Y(n)/X(n) \tag{4.1-11}$$

Furthermore,

$$H(n) = Y(n)/X(n) \tag{4.1-12}$$

which leads to the conclusion

$$H(n) \neq H_o(n) \tag{4.1-13}$$

Consequently, the DF transformation in itself does not remove the errors that change $X_o(n)$ and $Y_o(n)$ into X(n) and Y(n), respectively.

The amount of error in erroneous or noise corrupted data depends upon the signal to noise voltage ratio, SNR, e.g., $[x(k)]_{max}/\sigma_x$, where σ_x is the standard deviation (3.4-4) of the zero-mean measurement system noise present while x(k) was being measured. For a very large SNR, the noise has a very small effect upon the resultant (sampled) data sequence. With decreasing SNR's, the noise becomes increasingly dominant. Typically, for the most part, the noise spectrum is uniformly distributed (constant spectrum amplitude) and is weighted or filtered by the network properties of the sampled-data measurement system. The net result is that the higher frequency components of a low pass signal well within the measurement system passband are increasingly corrupted by noise with increasing frequency. Consequently, the corrupted DFT's X(n) and Y(n), respectively, with increasing corruption at the higher frequencies. Furthermore, their ratio gives the corrupted transfer fuction H(n), (4.1-12), whose IDFT yields the corrupted impulse response

$$h(k) = \text{IDFT}\,[H(n)] = \text{IDFT}\left[\frac{Y(n)}{X(n)}\right] \tag{4.1-14}$$

which may or may not be stable, and which may not even resemble the exact or true impulse response sequence $h_o(k)$; stable and unstable responses are shown in figure 4.3.

To reduce the influence of noise at the higher frequencies on H(n) and ultimately on h(k), H(n) is filtered with a low pass filter, R(n), which is judiciously chosen so as to provide a stable and smooth impulse response, d(k). R(n) operates on H(n) to give D(n),

$$D(n) = R(n) : H(n) \tag{4.1-15}$$

where : denotes that R(n) operates on H(n). In the time domain,

$$d(k) = \text{IDFT}\, D(n) \tag{4.1-16}$$

The error e(k) due to the estimate d(k) is then obtained from (3.4-1) and characterized as described in Section 3.4.

Too much filtering will yield an impulse response which varies too slowly, while too little filtering may yield an unstable result, figure 4.3. The choice of the low pass filter parameters is in fact arbitrary because the filtered result giving the impulse response, d(k), is subject to the interpretation of the viewer. Accordingly, the filtering process introduces an intrinsic error even though filtering may be necessary to achieve a stable result, i.e., a causal stable impulse response.

In addition to the noise corruption of the time domain sequence and the transform domain filtering distortion, the DFT , IDFT, and other computations can introduce numerical errors which also affect the choice of the estimate d(k). The DFT and IDFT are usually implemented using one of the fast Fourier transform (FFT) algorithms [13]. DFT, FFT and other operations can be analyzed to estimate the resultant signal to noise ratios where the noise is due to numerical truncation and rounding required by finite computing register length [14]. However, the errors due to computations are relatively small compared to those due to noise and filtering.

In summary, the overall deconvolution processes for time domain sequences using transform domain filtering possesses three sources of error whose effects can only be reduced but not eliminated. Hence, it is clear that the deconvolution process is inexact, i.e., the true answer, $h_o(k)$ cannot be obtained as has been pointed out earlier (Section 2.5), and reiterated here.

However, because "hope springs eternal in the human spirit" techniques have been and continue to be developed by which x(k) and y(k) are operated upon in order to "determine" $h_o(k)$. The job cannot be done in an exact sense, that is, $h_o(k)$ cannot be determined, but it can be estimated; the result is an estimate for $h_o(k)$. In many practical situations the results are good enough to be very useful. To be good enough and useful, the results reasonably approximate some preconceived view of the system behavior derived from a physically-based mathematical model. The argument can be made to appear more plausible by saying that the effects of the computation noise and the noise components $n_x(k)$ and $n_y(k)$, figure 4.2 are relatively small; but a little noise can ultimately be just as devastating to the deconvolution process as a large amount of noise. In fact, in practice, techniques for noise reduction or enhancement of the signal to noise ratio are first used; then, that noise which remains will still be potent enough to prevent exact deconvolution. The direct effect of noise can be seen in the analysis of time domain deconvolution processes; analyses of the effect are presented in Chapter 5.

4.2 Methods Considered

Two frequency domain methods are considered in the remainder of this chapter; they are:

1. A two parameter filtering method
2. A one parameter filtering method.

The common objective of the two methods is to filter out the noise which inhibits the deconvolution process from producing a smooth stable result. The filtering process is called a regularizing operation and the filter, a regularizing filter, is denoted by R(n).

In the first method there are two parameters which fix the characteristics of the filter. Its simple mathematical form enables the viewer to easily envision and choose a trial set of values for the two parameters based upon spectrum amplitude data in dB, $20 \log|H(n)|$, versus frequency $n\,\Omega$. After the two parameters are selected, the viewer observes the time domain response (IDFT), d(k). of the filtered H(n), D(n). The choice of the filter structure in this method was not based upon any strong a priori conditions, but rather upon the weak condition to attenuate the noisy portion of the spectrum for H(n) and thus reduce the error (3.4-1).

In the second method there is only one filter parameter; its value governs the gradual on-set of filtering to provide a smooth time domain response. Again, the viewer selects a trial value for the filter parameter and observes the time domain response, d(k) of the filtered H(n), D(n). However, in

this method the filter structure is based upon the a priori conditions to minimize a performance measure consisting of a weighted sum of the error power and the smoothness power (power being the energy per period or time window.)

4.3 The Two Parameter Method

4.3.1 The Filter Structure

The filter specifically operates on the real and imaginary parts of H(n),

$$H(n) = H_R(n) + j\, H_I(n) \tag{4.3-1}$$

to yield

$$D(n) = D_R(n) + j\, D_I(n) \tag{4.3-2}$$

As it was indicated for (4.1-14), the notation R(n):H(n) denotes that R(n) operates on H(n) which in general implies that the operation need not be the product R(n)H(n). In the two parameter method, the operation is not a product and operates on $H_R(n)$ and $H_I(n)$ to yield

$$D_R(n) = \begin{cases} H_R(n), & 0 \le n < n_o \\ |H(n_o)| \exp(-Mn), & n_o \le n \le N-1 \end{cases} \tag{4.3-3a}$$

$$D_I(n) = \begin{cases} H_I(n) & ; \quad 0 \le n < n_o \\ o & , \quad n_o \le n \le N-1 \end{cases} \tag{4.3-3b}$$

The magnitude of D(n) is given by

$$|D(n)| = \begin{cases} [H_R^2(n) + H_I^2(n)]^{1/2} & , \quad o \le n < n_o \\ |H(n_o)| \exp(-Mn) & , \quad n_o \le n \le N-1 \end{cases} \tag{4.3-4}$$

and

$$|D(n)|_{dB} = \begin{cases} 20 \log[H_R^2(n) + H_I^2(n)]^{1/2} & , \quad 0 \le n < n_o \\ 20 \log|H(n_o)| - 20(\log e)Mn & , \quad 0 \le n \le N-1 \end{cases} \tag{4.3-5}$$

The constant M is the cut-off slope of filtered H(n), i.e., $|D(n)|_{dB}$, figure 4.4 and is given by

$$M = \frac{-100 + |H(n_o)|_{dB}}{(A-1)n_o} \tag{4.3-6}$$

The rationale for the structure of the filter was to provide a simple means for attenuating magnitude |H(n)| response for n greater than some value n_o (i.e., in the noisy region), while simultaneously totally eliminating the noisy phase function, $\theta(n)$,

$$\theta(n) = \begin{cases} \mathrm{Tan}^{-1}[H_I(n)/H_R(n)], & 0 \le n < n_o \\ 0 , & n_o \le n \le N-1 \end{cases} \tag{4.3-7}$$

4.4 Application of the Two Parameter Method

4.4.1 Procedural Summary

In practice, the viewer inspects the data $|H(n)|_{dB}$ and selects the acceptable range, n_o, of the data (relatively noiseless data); this specifies n_o and $|H(n_o)|$. Then, the viewer selects the cut-off coefficient, A, which then completely specifies the cut-off slope, M, and D(n). Finally, the viewer calls for the IDFT of D(n) which provides the estimated or deconvolved impulse response, d(k). If the resultant d(k) and its associated error criteria, (3.4-1) through (3.4-4), are not satisfactory in the judgment of the viewer, the process is repeated using different values of n_o and A.

The procedure starting with the time domain sequences x(k) and y(k) and ending with the deconvolved time domain impulse sequence d(k) progresses as follows:

1. Offset correction for x(k) and y(k). Eliminate the offset, i.e., insure that x(o) = x(N-1) = 0 and y(o) = y(N-1) = 0. If x(k) and y(k) are step-like waveforms so that x(N-1) and y(N-1) are not zero, they must be transformed to impulsive waveforms. See subsection 4.6.2.
2. DF transform x(k) and y(k) to obtain X(n) and Y(n); display $|X(n)|_{dB}$ and $|Y(n)|_{dB}$.
3. Calculate the complex ratio H(n) = Y(n)/X(n); display $|H(n)|_{dB}$.
4. Select the regularization filter coefficients n_o and A; then filter H(n) to yield D(n), (4.3-3) display $|D(n)|_{dB}$.
5. IDF transform D(n) to obtain the time domain deconvolved result, d(k). (3.4-2), (4.1-16); display d(k).
6. Calculate the error criteria (3.4-1), (3.4-3), and (3.4-4) display e(k) and the values for $e_{max.}$, $e_{min.}$, $\bar{e}$, and σ.
7. Judge the quality of the deconvolved impulse response estimate, d(k), in terms of the error criteria; if unsatisfactory, return to step 4 and change the filter parameters.
8. If the quality is satisfactory, print the d(k) graphic display, related error criteria, and the numerical values.

Examples of the two parameter deconvolution method are given below for the determination of the impulse response of (1) 300 meter coaxial transmission line and (2) a very broadband antenna.

4.4.2 The Estimation of the Insertion Impulse Response of 300 Meters of RG 58 C/U Coaxial Cable

In this example the insertion input and output waveforms of a 300 meter length of RG 58 CU coaxial transmission line, x(k) and y(k), respectively, were recorded, figure 4.5; the signal to noise ratios for these waveforms were relatively large, SNR-x 75 dB and SNR-y 46 dB. Deconvolution to obtain the insertion impulse response was then performed using varying degrees of regularization with the two parameter filter. The DFT of the insertion impulse response is the scattering insertion parameter $S_{12}(n)$. [17]. The discrete transformation and inverse transformation operations were implemented using the FFT and IFFT algorithms.

Next, various combinations of SNR-x and SNR-y were created by adding noise (a pseudorandom sequence) into the original high SNR data recorded above. For each combination of SNR's, two parameter deconvolution was performed, again using varying degrees of regularization.

Specifically, the following deconvolution experiments were performed:

1. SNR-x: 75 dB, SNR-y: 46 dB
   No regularization: (N-1) Ω = 250 MHz
   Moderate regularization = n_o Ω = 100 MHz, A = 5
   Strong regularization: n_o Ω = 74 MHz, A = 6

2. SNR-x: 75 dB, SNR-y: 36 dB
   No regularization: (N-1) Ω = 250 MHz
   Strong regularization: n_o Ω = 74 MHz, A = 6

3. SNR-x: 45 dB, SNR-y: 46 dB
   No regularization: (N-1) Ω = 250 MHz
   Moderate regularization: n_o Ω = 100 MHz, A = 5

4. SNR-x: 40 dB, SNR-y: 40 dB
   No regularization: (N-1) Ω = 250 MHz
   Strong regularization: n_o Ω = 74 MHz, A = 6

The results are summarized in figures 4.6 through 4.8 and more detailed information is presented in figures 4.9 through 4.23.

Figure 4.6 brings together the results for the no-regularization cases of each of the experiments. In each of these cases, H(n) was directly inversely transformed by the IDFT to obtain the time domain sequence h(k). Also, shown in figure 4.6 are the $|H(n)|_{dB}$ corresponding to the various h(k) responses. Generally, it is seen that the noisier a $|H(n)|_{dB}$ is, the noisier is the corresponding h(k). Furthermore, a change in the SNR-y of the output waveform is more critical than that of the input waveform with regards to the noise characteristics of $|H(n)|_{dB}$ and h(k). For example, a 10 dB reduction in SNR-y yields very noisy results, figure 4.6b, while a 30 dB reduction in SNR-x yields a much smaller noise increase in $|H(n)|_{dB}$ and h(k), figure 4.5c. This follows from the fact that in the latter case the transmission properties of H(n) attenuates the effect of the noise components introduced at the input.

Figure 4.7 brings together the results for the regularized cases of each of the experiments. In each of these cases, H(n) is operated on by the two parameter (n_o,A) regularization operator to yield D(n), and the result is inversely transformed by the IDFT to obtain the time domain sequence, d(k). The horizontal scale is expanded to 50 ns/Div. (as compared to 200 ns/Div. in figure 4.6) in order to show the fine structure of d(k). Also, h(k) for the largest SNR case (figure 4.6a) is shown in expanded time scale as figure 4.5a for comparison purposes.

It is evident that the regularization or filtering process produces much smoother estimates, d(k), of the actual coaxial cable impulse response, $h_o(k)$, than do the various non-regularized h(k). The smoothest non-regularized response is shown in figure 4.7a while the other non-regularized responses of lesser smoothness are shown in figure 4.6. The effect of regularization is very dramatic in reducing the noise from the impulse response.

Before commenting any further on figure 4.7, it is necessary to consider the input and output waveforms of figures 4.9 and 4.10, respectively. First, from figure 4.9, it is clear that the input pulse waveform abruptly rises, at say t_o, and is zero prior to t_o; also, the pulse duration was of the order of 10 ns. Second, from figure 4.10, the output pulse waveform possesses some small variations prior to the large abrupt rise of the pulse waveform; also, there is a small variation from the smooth decay of the pulse waveform at about 800 ns. The main rise and subsequent return to zero of the output waveform occurs in about 1000 ns; consequently the input waveform is approxmiately an impulse,

which in turn, says that the output waveform will approximate the impulse response of the coaxial cable, $h_o(k)$. Equipped with these physical details of the input and output waveforms, further comments can be made about the deconvolution results presented in figures 4.6 and 4.7.

Because the output waveform in figure 4.10 was produced by a very short duration input pulse, variations similar to those in the output waveform should be detectable in the deconvolved waveforms, d(k) and h(k). Inspection of figure 4.7 shows different oscillations before the main rise of each waveform. The time scale in figure 4.7 is 50 ns/Div. which is one-quarter of that in figure 4.10. Because figure 4.10 doesn't possess oscillations in the 50 ns time interval prior to the start of the main waveform rise, those appearing in figure 4.7 are due to noise or the filtering process.

For example, refer to figure 4.6a and note the noise level on the baseline of the waveform; then look at figure 4.7a which is an expanded time scale version of figure 4.6a. Consequently, it is clear that the oscillations before the main rise are due to noise.

Next, look at figure 4.7b and note that the regularization has smoothed out the variations due to noise, but that the smoothing very close to the initial rise is not as great as that elsewhere. Stronger regularization increases the magnitude of the oscillations as evidenced in figure 4.7c. This is due to the filtering by the regularization process which introduces the Gibbs phenomena due to the truncation of H(n) when forming D(n) [15]. Such effects can be made negligible by avoiding abrupt or rectangular filtering of the H(n) spectrum. Smooth filtering is used in the one-parameter deconvolution method discussed in Section 4.6.

In Section 4.1 it was pointed out that the FFT and IFFT computations temselves introduce noise or errors due to finite register length. Such errors also cause the imaginary part of the IFFT to be nonzero. For example, say, the real time domain sequence f(k) is given and is then transformed by the FFT to F(n); next, f(k) is recovered from F(n) via the IFFT. However, upon inspection of the data for IFFT {F(n)} it is found that the result is complex with a very small imaginary part while the real part is very close to the values of the original sequence f(k). In practice, such non-zero imaginary parts are always present, and are relatively very small, figure 4.8. In Section 4.5, the nature and origin of the errors affecting d(k) will be discussed.

4.4.3 The Estimation of the Insertion Impulse Response of a Broadband Antenna

As a second example for the application of the two parameter method, the method was applied to experimental data to obtain the insertion impulse response of a broadband antenna. The data was obtained using the NBS Time Domain Antenna Range in conjunction with the NBS APMS. Schematically, the experimental arrangement is shown in figure 4.24.

Figures 4.25 and 4.26 show the input, x(k), and the output, y(k), and the magnitudes of their respective DFT's, $|X(n)|$ and $|Y(n)|$. The magnitude,

$$|H(n)| = |Y(n)| \div |X(n)|$$

is shown in figure 4.27. Deconvolution results are shown in figures 4.28a and 4.28b for two different degrees of filtering; note that the smoother result and most error occurs for the greater degree of filtering.

4.5 On the Nature of the Error, e(k)

With the results presented in Section 4.4 in mind, it is now appropriate to provide a model for the error e(k) encountered in frequency domain deconvolution. The model is not restricted to any specific frequency domain regularization filter; consequently, it is valid for both the two-parameter method (Sections 4.3, 4.4) and the one parameter method (Section 4.6). The discussion will be divided into three parts:

1. Error in the absence of both regularization (filtering) and computation errors (noise).
2. Error in the presence of regularization and absence of computation errors.
3. Error in the presence of both regularization and computation errors.

being contained in subsections 4.5.1 through 4.5.3, respectively.

4.5.1 Error in the Absence of the Filtering Operation and Computation Noise

The error sequence, e(k), as given by (3.4-1) includes the effects of filtering and computation errors in the d(k) term,

$$e(k) = y(k) - x(k) * d(k) \tag{4.5.1-1}$$

In the absence of filtering and computation noise

$$d(k) = h(k) \tag{4.5.1-2}$$

then

$$e(k) = 0 \tag{4.5.1-3}$$

because

$$y(k) = x(k) * h(k) \tag{4.5.1-4}$$

In words, the error sequence, e(k), will be zero **only** when d(k) equals h(k); any other choice for d(k) will give a non-zero result.

4.5.2 Error in the Presence of the Filtering Operation and in the Absence of Computation Errors

Now consider the nature of h(k) for the signals x(k) and y(k) which are both noisy signals, i.e.,

$$x(k) = x_o(k) + n_x(k) \tag{4.5.2-1}$$

$$y(k) = y_o(k) + n_y(k) \tag{4.5.2-2}$$

where $n_x(k)$ and $n_y(k)$ are the noise components associated with the true signals $x_o(k)$ and $y_o(k)$, respectively. The system impulse response is

$$h(k) = [y_o(k) + n_y(k)](1/*)[x_o(k) + n_x(k)] \tag{4.5.2-3}$$

whose DFT is

$$H(n) = \frac{Y_o(n) + N_y(n)}{X_o(n) + N_x(n)} \tag{4.5.2-4}$$

The transfer function H(n), (4.5.2-4), obtained from the deconvolution y(k)(1/*)x(k)) can be expressed as

$$H(n) = H_o(n) \frac{1 + N_y(n)/Y_o(n)}{1 + N_x(n)/X_o(n)} \qquad (4.5.2\text{-}5)$$

where

$$H_o(n) = \frac{Y_o(n)}{X_o(n)} \qquad (4.5.2\text{-}6)$$

Define,

$$M(n) = \frac{1 + N_y(n)/Y_o(n)}{1 + N_x(n)/X_o(n)} \qquad (4.5.2\text{-}7)$$

$H_o(n)$ is the DFT of the stable true impulse response defined by

$$h_o(k) = y_o(k)(1/*)x_o(k) \qquad (4.5.2\text{-}8)$$

Consequently, H(n) can be viewed as the product of the true transfer function $H_o(n)$ and some noisy transfer function, M(n), which may or may not be stable.

$$H(n) = H_o(n)\ M(n) \qquad (4.5.2\text{-}9)$$

If H(n) is filtered by R(n) to deemphasize the effect of the noise, M(n), then D(n) results

$$D(n) = R(n)\ [H_o(n)\ M(n)] \qquad (4.5.2\text{-}10)$$

Assuming that R(n) operates only on M(n), i.e., write

$$D(n) = H_o(n)\ [R(n):M(n)] \qquad (4.5.2\text{-}11)$$

in the limit

$$\lim_{R:M \to 1} D(n) = H_o(n) \qquad (4.5.2\text{-}12)$$

However, it is physically impossible for the filter R(n) to operate solely on M(n). In practice R(n) also operates on $H_o(n)$ with the design of R(n) being such that it affects M(n) more than $H_o(n)$ and makes d(k) at least stable and possibly smooth.

Consequently, filtering attempts to make h(k) stable by remvoing m(k) from the impulse response,

$$h(k) = h_o(k) * m(k) \qquad (4.5.2\text{-}13)$$

in such a way that

$$d(k) = h_o(k) * [r(k) * m(k)] \qquad (4.5.2\text{-}14)$$

that is, d(k) becomes approximately

$$d(k) \doteq h_o(k) \qquad (4.5.2\text{-}15)$$

where the last equation is a judgment by the individual performing the deconvolution. The resultant d(k) does not cause the error, e(k), to go to zero; only h(k) can do so.

Therefore, it can be concluded that a paradoxial situation exists: h(k) makes the error, e(k), vanish; but h(k) is not the desired result, $h_o(k)$, which the individual performing the deconvolution wishes to achieve. The paradox occurs because y(k) is used as the reference waveform or sequence in the error expression, (3.4-1). As pointed out in Section 3.4, $y_o(k)$ is not known; consequently, the known quantity y(k) is used.

Since h(k) and H(n) are given by (4.5.2-3) and (4.5.2-4), respectively, the error, e(k), can be expressed as

$$e(k) = y(k) - x(k)*r(k) * [y(k)(1/*)x(k)] \tag{4.5.2-16}$$

$$= y(k) * [\delta(k) - r(k)] \tag{4.5.2-17}$$

where $\delta(k)$ is a unit impulse. In the discrete frequency domain, the error is

$$E(n) = [Y_o(n) + N_y(n)] - R(n)\left[\frac{Y_o(n) + N_y(n)}{X_o(n) + N_x(n)}\right] [X_o(n) + N_x(n)] \tag{4.5.2-18}$$

which corresponds to (4.5.2-16) while

$$E(n) = [1 - R(n)] [Y_o(n) + N_y(n)] \tag{4.5.2-19}$$

corresponds to (4.5.2-17). The square of the error magnitude is

$$|E(n)|^2 = |1 - R(n)|^2 |Y_o(n) + N_y(n)|^2 \tag{4.5.2-20}$$

and the error energy over the N points comprising the time window or one period of e(k) is, by the Parseval relation [22, 23],

$$\sum_{k=0}^{N-1} e^2(k) = \frac{1}{N} \sum_{n=0}^{N-1} |E(n)|^2 \tag{4.5.2-21}$$

where

$$\sum_{n=0}^{N-1} |E(n)|^2 = \sum_{n=0}^{N-1} |1 - R(n)|^2 |Y_o(n) + N_y(n)|^2 \tag{4.5.2-22}$$

Consequently, it is seen that the error energy per period depends upon the filter characteristics R(n), $Y_o(n)$, and $N_y(n)$. The variance or the square of the standard deviation of the error, i.e., the square of (3.4-4), is given by

$$\sigma^2 = \frac{1}{N} \sum_{k=0}^{N-1} [e(k) - \bar{e}\,]^2 \tag{4.5.2-23}$$

$$\sigma^2 = \frac{1}{N} \sum_{k=0}^{N-1} e^2(k) - \bar{e}^2 \qquad (4.5.2\text{-}24)$$

$$\sigma^2 = \frac{1}{N^2} \sum_{n-0}^{N-1} |E(n)|^2 - \bar{e}^2 \qquad (4.5.2\text{-}25)$$

where $\bar{e}$ is the mean value of the error (3.4-3).

The effects of filtering on the error can be graphically illustrated by considering the magnitude of the error, E(n). From (4.5.2-20) the magnitude is

$$|E(n)| = |1 - R(n)| \; |Y_o(n) + N_y(n)| \qquad (4.5.2\text{-}26)$$

Figure 4.29 shows the error for two choices for R(n) where R(n) is real; the absolute error, $|E(n)|$, approaches zero as R(n) approaches unity.

4.5.3 Error in the Presence of the Filtering Operation and Computation Errors

In the presence of filtering and computation errors the estimated impulse response, d(k), in (3.4-1) is replaced by $d_3(k)$. The error is then given by

$$e(k) = y(k) - x(k) * d_3(k) \qquad (4.5.3\text{-}1)$$

which can be explicitly written as

$$e(k) = y(k) * [\delta(k) - r(k) * m_1(k)] - x(k) * [r(k) * n_{h1}(k) + n_r(k) + n_{d2}(k)] + n_e(k) \qquad (4.5.3\text{-}2)$$

and which will be derived below. e(k) reduces to (4.5.2-13) when the computation errors vanish, that is, when

$$m_1(k) = \delta(k), \text{ unit impulse} \qquad (4.5.3\text{-}3)$$

$$n_{h1}(k) = 0 \qquad (4.5.3\text{-}4)$$

$$n_{d2}(k) = 0 \qquad (4.5.3\text{-}5)$$

$$n_e(k) = 0 \qquad (4.5.3\text{-}6)$$

The derivation of e(k), (4.5.3-2), and the definitions of (4.5.3-3) through (4.5.3-6) are based upon the assumption that the round-off error in calculating a specific quantity, can be represented as an additive term to the desired quantity [16]. For example, if A(n) is being calculated from a(k) by the FFT, then the result is

$$A_1(n) = A(n) + N'_a(n) \qquad (4.5.3\text{-}7)$$

where A(n) is the desired result, and $N'_a(n)$ is the computation error or noise. Similarly, the IFFT

of $A_1(n)$ would also introduce an additive error, $n_{a1}(k)$.

$$a_1(k) = a(k) + n_a(k) + n_{a1}(k) \tag{4.5.3-8}$$

Thus, it is assumed that the cummulative error depends upon the number of computations, and that, a(k) is not exactly recoverable from the numerical calculations.

To derive $e_2(k)$, (4.5.3-2), the starting point is x(k) and y(k) and the procedure follows below.

$$\text{FFT}\,[x(k)] = X_1(n) = X(n) + N'_x(n) \tag{4.5.3-9}$$

$$\text{FFT}\,[y(k)] = Y_1(n) = Y(n) + N'_y(n) \tag{4.5.3-10}$$

Here, $N'_x(n)$ and $N'_y(n)$ are computation errors (noise); note that both X(n) and Y(n) also contain measurement noise corresponding to the noise contained in x(k) and y(k), (4.5.1-5) and (4.5.1-6), respectively.

Next compute the quotient of (4.5.3-10) divided by (4.5.3-9); this result is defined as $H_1(n)$,

$$H_1(n) = \frac{Y_1(n)}{X_1(n)} + N_{h1}(n) \tag{4.5.3-11}$$

where $N_{h1}(n)$ is the additive error due to the calculation of $H_1(n)$.

Filtering $H_1(n)$ yields $D_1(n)$.

$$D_1(n) = R(n)\left[\frac{Y_1(n)}{X_1(n)} + N_{h1}(n)\right] + N_r(n) \tag{4.5.3-12}$$

where $N_r(n)$ is the additive error due to the calculation of $D_1(n)$. In terms of the theoretical transforms Y(n) and X(n), $D_1(n)$ is given by

$$D_1(n) = R(n)\left[\frac{Y(n)}{X(n)}\left\{\frac{1 + N'_y(n)/Y(n)}{1 + N'_x(n)/X(n)}\right\} + N_{h1}(n)\right] + N_r(n) \tag{4.5.3-13}$$

$$= R(n)\left[\frac{Y(n)}{X(n)}\, M_1(n) + N_{h1}(n)\right] + N_r(n) \tag{4.5.3-14}$$

where $M_1(n)$ is a multiplicative error term defined by the { } term in (4.5.3-13).

The IFFT of $D_1(n)$ is given by

$$d_2(k) = r(k) *\{[y(k)(1/*)\, x(k)]* m_1(k) + n_{h1}(k)\} + n_r(k) + n_{d2}(k) \tag{4.5.3-15}$$

where $n_{d2}(k)$ is a complex additive error due to the IFFT calculation of $d_2(k)$. The error e(k) is given by

$$e(k) = y(k) - x(k) * d_2(k) + n_e(k) \tag{4.5.3-16}$$

and

$$d_3(k) = d_2(k) + n_e(k)(1/*)x(k) \tag{4.5.3-17}$$

then

$$e(k) = y(k) - x(k) * d_3(k) \tag{4.5.3-18}$$

$$e(k) = y(k) * [\delta(k) - r(k) * m_1(k)] - x(k) * [r(k) * n_{h1}(k) + n_r(k) + n_{d2}(k)] + n_e(k). \quad (4.5.3\text{-}19)$$

where $n_e(k)$ is the additive error due to the calculation of $e(k)$.

When the filtering operation is deleted

$$r(k) = \delta(k) \quad (4.5.3\text{-}20)$$

$$n_r(k) = 0 \quad (4.5.3\text{-}21)$$

$d_3(k)$ will change to $d_4(k)$ and the error becomes

$$e(k) = y(k) - x(k) * d_4(k) \quad (4.5.3\text{-}22)$$

Explicitly, from (4.5.3-19) the error becomes

$$e(k) = y(k) * [\delta(k) - m_1(k)] - x(k) * [n_{h1}(k) + n_{d2}(k)] + n_e(k). \quad (4.5.3\text{-}23)$$

This is the error obtained when no filtering is employed and computation noise is present. When computation noise is absent, (4.5.3-3) through (4.5.3-6) prevail and $e(k)$ reduces to zero.

As pointed out in connection with (4.5.3-15), $d_2(k)$ is complex due to the complex additive error term, $n_{d2}(k)$, which results from the IFFT calculation. Consequently, the estimated impulse response, (4.5.3-15)

$$d_2(k) = r(k) * [y(k)(1/*)x(k) * m_1(k) + n_{h1}(k)] + n_r(k) + n_{d2}(k)$$

or

$$d_2(k) = r(k) * [h(k) * m_1(k) + n_{h1}(k)] + n_r(k) + n_{d2}(k) \quad (4.5.3\text{-}24)$$

is complex. Also, the estimated impulse response, without filtering, $d_4(k)$, is complex. From (4.5.3-24) with (4.5.3-20) and (4.5.3-21) $d_4(k)$ is given by

$$d_4(k) = [h(k) * m_1(k) + n_{h1}(k)] + n_r(k) + n_{d2}(k) \quad (4.5.3\text{-}25)$$

4.6 The One Parameter Method

4.6.1 The Filter Structure

This filter is an optimum filter in the sense that it minimizes a performance measure with respect to the weighted sum of the error power and a smoothness power [18,19]. It turns out that there is only one adjustable parameter in this method, and the parameter adjusts the weighting between the error and smoothness. The derivation of the filter is given below.

Given the two signals x(k) and y(k),

$$x(k) = x_o(k) + n_x(k) \tag{4.6.1-1}$$

$$y(k) = y_o(k) + n_y(k) \tag{4.6.1-2}$$

computation and filtering lead to the estimated impulse response, $d_2(k)$, (4.5.3-16),

$$d_2(k) = r(k) * [y(k)(1/*)\ x(k) *\ m_1(k) + n_{h1}(k)] + n_r(k) + n_{d2}(k) \tag{4.6.1-3}$$

and the error function.

$$e(k) = y(k) - x(k) * d_2(k) \tag{4.6.1-4}$$

In addition to the error constraint, e(k), define an auxillary constraint, s(k),

$$s(k) = c(k) * d_2(k) \tag{4.6.1-5}$$

where c(k) is the discrete function that defines the constraint. In the present discrete case c(k) is the second-difference operator, which corresponds to the second derivative operator for continuous functions, d^2/dt^2, thus s(k) depends upon the rate-of-change of the first derivative of $d_2(k)$, i.e., its smoothness.

$$\overline{E} = \sum_{k=0}^{N-1} [e(k)]^2, \tag{4.6.1-6}$$

$$\overline{S} = \sum_{k=0}^{N-1} [s(k)]^2 \tag{4.6.1-7}$$

Now, to obtain an optimum solution, adjust $d_2(k)$ until the performance measure, P,

$$P = \overline{E} + \gamma\ \overline{S} \tag{4.6.1-8}$$

is minimized. γ is a weighting function which selects the degree of smoothing. Consequently, the minimization condition is approximately expressed by

$$\frac{\partial}{\partial d_2(k)} \left[\overline{E} + \gamma\ \overline{S} \right] = 0 \tag{4.6.1-9}$$

assuming that N is a very large number so that $d_2(k)$ approximates a continuous function rather than a discrete one.

In the DF domain $d_2(k)$, e(k), and s(k) become

$$D_2(n) = R(n)\left[\frac{Y(n)}{X(n)} M_1(n) + N_{h1}(n)\right] + N_r(n) + N_{d2}(n) \tag{4.6.1-10}$$

$$E(n) = Y(n) - X(n)D_2(n) \tag{4.6.1-11}$$

$$S(n) = C(n)\, D_2(n) \tag{4.6.1-12}$$

Parseval's relation [22,23]

$$\sum_{k=1}^{N} a^2(k) = \frac{1}{N} \sum_{n=0}^{N-1} A^2(n) \tag{4.6.1-13}$$

gives for $\overline{E}$ and $\overline{S}$,

$$\overline{E} = \frac{1}{N} \sum_{n=0}^{N-1} E(n)\, E^*(n) \tag{4.5.1-14}$$

$$\overline{S} = \frac{1}{N} \sum_{n=0}^{N-1} S(n)\, S^*(n) \tag{4.5.1-15}$$

where $E^*(n)$ denotes the complex conjugate of E(n), etc. The sum $\overline{E} + \gamma\, \overline{S}$ is given by

$$\overline{E} + \gamma\, \overline{S} = \frac{1}{N} \sum_{n=0}^{N-1} [E(n)\, E^*(n) + \gamma\, S(n)\, S^*(n)] \tag{4.5.1-16}$$

where

$$E(n)\, E^*(n) = [Y(n) - X(n)\, D_2(n)]\, [Y(n) - X(n)\, D_2(n)]^* \tag{4.6.1-17}$$

$$\begin{aligned} S(n)\, S^*(n) &= C(n)\, D_2(n)\, C^*(n)\, D_2^*(n) \\ &= |C(n)\, D_2(n)|^2 \end{aligned} \tag{4.6.1-18}$$

Consider $E(n)\, E^*(n)$;

$$\begin{aligned} E(n)\, E^*(n) &= [Y(n) - X(n)\, D_2(n)]\, [Y(n) - X(n)\, D_2(n)]^* \\ &= [Y(n) - X(n)\, D_2(n)]\, [Y^*(n) - \{X(n)\, D_2(n)\}^*] \\ &= Y(n)\, Y^*(n) - Y(n)\, \{X(n)D_2(n)\}^* - Y^*(n)\, \{X(n)D_2(n)\} \\ &\quad + \{X(n)\, D_2(n)\}\, \{X(n)\, D_2(n)\}^* \end{aligned}$$

$$E(n)E^*(n) = |Y(n)|^2 - Y(n)\{X(n)\,D_2(n)\}^* - Y^*(n)\{X(n)\,D_2(n)\} + |X(n)\,D(n)|^2 \quad (4.6.1\text{-}19)$$

Hence

$$\overline{E} + \gamma\,\overline{S} = \frac{1}{N}\sum_{n=0}^{N-1}\left\{|Y(n)|^2 + |X(n)\,D_2(n)|^2 + \gamma|C(n)\,D_2(n)|^2 - Y^*(n)\{X(n)\,D_2(n)\} - Y(n)\{X(n)\,D_2(n)\}^*\right\} \quad (4.6.1\text{-}20)$$

Consider $X(n)\,D_2(n)$. From, (4.6.1-10),

$$X(n)\,D_2(n) = R(n)\,[Y(n)\,M_1(n) + X(n)\,N_{h1}(n)] + X(n)\,[N_r(n) + N_{d2}(n)] \quad (4.6.1\text{-}21)$$

First, $|M_1(n)|$ is of the form $1 + J(n)$, where $J(n)$ is complex and $|J(n)| \ll 1$. Second, $|N_{h1}(n)|$, $|N_r(n)|$, and $|N_{d2}(n)|$ are the magnitudes of computation errors; consequently, they too will be much less than unity. When $X(n)\,D_2(n)$ is written in the form

$$X(n)\,D_2(n) = R(n)\,Y(n)\,A(n) \quad (4.6.1\text{-}22)$$

$A(n)$ will be complex and of the form

$$A(n) = 1 + K(n) \quad (4.6.1\text{-}23)$$

where

$$|K(n)| \ll 1\,. \quad (4.6.1\text{-}24)$$

Consequently, $A(n)$ will be very close to the real number unity.

Some complex relations:

$$XD = RYA \quad (4.6.1\text{-}25)$$

$$= |R||Y||A| \angle \phi_r + \phi_y + \phi_a \quad (4.6.1\text{-}26)$$

$$(XD)^* = |R||Y||A| \angle -(\phi_r + \phi_y + \phi_a) \quad (4.6.1\text{-}27)$$

$$Y = |Y| \angle \phi_y \quad (4.6.1\text{-}28)$$

$$Y^* = |Y| \angle -\phi_y \quad (4.6.1\text{-}29)$$

$$Y^*(XD) = |R||Y|^2|A| \angle \phi_r + \phi_a \quad (4.6.1\text{-}30)$$

$$Y\,(XD)^* = |R||Y|^2|A| \angle -(\phi_r + \phi_a) \quad (4.6.1\text{-}31)$$

$$Y^*(XD) + Y(XD)^* = |R||Y|^2|A|\{2\,\mathrm{Cos}(\phi_r - \phi_a)\} \quad (4.6.1\text{-}32)$$

$$= 2R|Y|^2|A|\,\mathrm{Cos}\,\phi \quad (4.6.1\text{-}33)$$

where

$$\phi = \phi_r - \phi_a \quad (4.6.1\text{-}34)$$

Putting (4.6.1-33) into (4.6.1-20) yields

$$\bar{E} + \gamma \bar{S} = \frac{1}{N} \sum_{n=0}^{N-1} \Big\{ |Y(n)|^2 + |R(n)\ Y(n)\ A(n)|^2 + \gamma |C(n)\ D_2(n)|^2 - 2R(n)|Y(n)|^2\ A(n)\ \cos\phi(n) \Big\} \qquad (4.5.1\text{-}35)$$

$$= \frac{1}{N} \sum_{n=o}^{N-1} |Y(n)|^2 + |R(n)\ Y(n)\ A(n)|^2 + \gamma \left| \frac{C(n)\ R(n)\ Y(n)\ A(n)}{X(n)} \right|^2 - 2R(n)|Y(n)|^2 |A(n)| \cos\phi(n) \Big\} \qquad (4.6.1\text{-}36)$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} |Y(n)|^2 \Big\{ 1 + R^2(n)\ |A(n)|^2 + \frac{\gamma\ |C(n)|^2\ R^2(n)\ |A(n)|^2}{|X(n)|^2} - 2\ R(n)\ |A(n)|\ \cos\phi(n) \Big\} \qquad (4.6.1\text{-}37)$$

In principle, the performance measure (4.6.1-8)

$$P = \bar{E} + \gamma \bar{S}$$

can be expressed in terms of discrete time domain quantities or in terms of discrete frequency domain quantities by virtue of the Parseval relation (4.5.2-21). However, in practice, given the time domain quantities x(k) and y(k), subsequent numerical computations introduce errors into the frequency domain quantities. Consequently, in the frequency domain the result is P',

$$P' = \bar{E}' + \gamma \bar{S}'$$

which is not equal to P. To minimize P, $d_2(k)$ is adjusted. Varying $d_2(k)$ corresponds to adjusting the frequency domain filter R(n). Consequently, P and P' are minimized through the relations,

$$\frac{\partial P}{\partial d_2(k)} = 0$$

$$\frac{\partial P'}{\partial R(n)} = 0$$

respectively. Because the computation errors are very small, a first order approximation is to assume that P is equal to P'. With that assumption being made a change in $d_2(k)$ corresponds to a change in R(n), i.e.,

$$\frac{\partial}{\partial d_2(k)} \left[\bar{E} + \gamma \bar{S} \right] = \frac{\partial}{\partial R(n)} \left[\bar{E} + \gamma \bar{S} \right] \qquad (4.6.1\text{-}38)$$

and

$$\frac{\partial}{\partial R(n)} \left[\bar{E} + \gamma \bar{S} \right] = \frac{\partial}{\partial |R(n)|} \left[\bar{E} + \gamma \bar{S} \right] + \frac{\partial}{\partial \phi_r} \left[\bar{E} + \gamma \bar{S} \right] \qquad (4.5.1\text{-}39)$$

Consequently,

$$\frac{\partial}{\partial R(n)}\left[\overline{E} + \gamma \overline{S}\right] = \frac{1}{N}\sum_{n=0}^{N-1} |Y(n)|^2 \left\{2|R(n)||A(n)|^2 + \frac{2\gamma|C(n)A(n)|^2|R(n)|}{|X(n)|^2}\right.$$

$$\left. - 2|A(n)|\mathrm{Cos}[\phi_r(n) + \phi_a(n)] + 2|A(n)R(n)|\mathrm{Sin}[\phi_r(n) + \phi_a(n)]\right\} \qquad (4.6.1\text{-}40)$$

To minimize $[\overline{E} + \gamma \overline{S}]$ with respect to R(n), $|R(n)|$ and $\phi_r(n)$ must be such that (4.6.1-40) is zero. Since $|Y(n)|^2$ is not generally zero, the $\{\ \}$ -term in (4.6.1-40) must be made equal to zero to force (4.6.1-40) to zero. Choosing the filter phase angle to be the negative of the computation error phase angle $\phi_a(n)$,

$$\phi_{ro}(n) = -\phi_a(n) \qquad (4.6.1\text{-}41)$$

eliminates the phase angle dependence, and gives for $|R_o(n)|$,

$$|R_o(n)| = \frac{|X(n)|^2}{|A(n)||X(n)|^2 + \gamma|A(n)||C(n)|^2} \qquad (4.6.1\text{-}42)$$

Consequently, the optimum filter $|R_o(n)|\ \underline{/\phi_{ro}(n)}$ is given by (4.6.1-41) and (4.6.1-42) in the presence of computation noise.

If smoothing is not employed, γ is set equal to zero; and the filter becomes

$$|R_o(n)|_{\gamma=o} = \frac{1}{|A(n)|} \qquad (4.6.1\text{-}43)$$

$$[\phi_{ro}(n)]_{\gamma=0} = \phi_a(n) \qquad (4.6.1\text{-}44)$$

which says that the filter magnitude and phase are such that they just cancel out the computation errors so as to force the error function e(k) to zero. Furthermore, when computation errors are not present A(n) is unity, i.e.,

$$|A(n)| = 1 \qquad (4.6.1\text{-}45)$$

$$\phi_a = 0 \qquad (4.6.1\text{-}46)$$

which, in turn, means that (4.6.1-43) and (4.6.1-44) reduce to

$$\underset{\substack{\gamma=0\\A(n)=1}}{[R(n)]} = 1 \qquad (4.6.1\text{-}47)$$

$$\underset{\substack{\gamma=0\\A(n)=1}}{[\phi_R(n)]} = 0 \qquad (4.6.1\text{-}48)$$

which is the case of no filtering, i.e., R(n) = 1. As R(n) approaches unity, the filtering 1-R(n) approaches zero, figure 4.29.

In the presence of smoothing, $\gamma \neq 0$, and in the absence of the computation errors, A(n) = 1, the optimum filter is real and is given by

$$R(n) = \frac{|X(n)|^2}{|X(n)|^2 + \gamma |C(n)|^2} \qquad (4.6.1\text{-}49)$$

Before discussing the form of C(n), some summary comments are in order to clarify the meaning of the optimal result in the presence of computation errors, (4.6.1-41) and (4.6.1-42). It must be kept in mind that in any given problem A(n) is not known, i.e., $|A(n)|$ and $\phi_a(n)$ are unknown; consequently, the optimal filter cannot be calculated. Equations (4.6.1-41) and (4.6.1-42) are only of theoretical interest in that they show that computation errors require that the optimum filter, $R_o(n)$, be different from that in the absence of computation errors, R(n). Specifically, $R_o(n)$ equals A(n) R(n). In practice, since A(n) is not known, it is convenient to let R(n) be real and given by (4.6.1-49), which is the case in the absence of computation errors. Again, because A(n) is very close to the real value of unity, the use of (4.6.1-49) is the best that can be done.

c(k) is defined as the second difference sequence operator and corresponds to the second derivative operator in the continuous case. The first backward-difference sequence, ∇, is defined by the operation $\{x(kT) - x[k-1)T]\}/T$ where the interval, T, between the k-th and (k-1)-th values is explicitly shown. Implicitly, the operation is denoted as x(k) - x(k-1). $\nabla * x(k)$ is given by

$$\nabla * x(k) = [1,-1,0,\ldots, N-1] * x(k) \qquad (4.6.1\text{-}50)$$

The second backward-difference sequence, ∇^2, is defined by the implicit operation x(k)-2x(k-1)+x(k-2). $\nabla^2 * x(k)$ is given by

$$\nabla^2 * x(k) = [1,-2,1,0,\ldots, N-1] \qquad (4.6.1\text{-}51)$$

Note that there are N-3 and N-4 zeros in the ∇ and ∇^2 sequences, respectively. The DFT of

$$\nabla^2 = [1, -2, 1, 0, 0, 0, \ldots .N-1.] \qquad (4.6.1\text{-}52)$$

is given by the complex sequence

$$\Delta^2 = \sum_{k=0}^{2} \nabla^2(k) e^{-j\Omega Tkn}, \; n = 0,1,2,0,0,0,\ldots N-1 \qquad (4.6.1\text{-}53)$$

$$= 1 - 2e^{-j\Omega Tn} + e^{-j2\,\Omega Tn} \; ; \; n = 0,1,2,0,0,0,\ldots N-1 \qquad (4.6.1\text{-}54)$$

which also contains N-4 zeros, and where

$$\Omega T = \frac{2\pi}{N} \qquad (4.6.1\text{-}55)$$

Since c(k) is defined as $\nabla^2(k)$, then C(n) is given by

$$C(n) = 1 - 2e^{-j\frac{2\pi n}{N}} + e^{-j\frac{4\pi n}{N}}, \; n = 0,1,2,0,0,0,\ldots N-1 \qquad (4.5.1\text{-}56)$$

and

$$|C(n)|^2 = C(n)C(n)^* = \left[1-2e^{-j\frac{2\pi n}{N}} + e^{-j\frac{4\pi n}{N}}\right]\left[1-2e^{j\frac{2\pi n}{N}} + e^{j\frac{4\pi n}{N}}\right];$$

$$n = 1,1,2,0,0,0,\ldots N-1 \qquad (4.5.1\text{-}57)$$

$$|C(n)|^2 = 6 - 8\,\mathrm{Cos}\,\frac{2\pi n}{N} + 2\,\mathrm{Cos}\,\frac{4\pi n}{N}\,;\quad n = 0,1,2,0,0,0,\ldots N-1 \qquad (4.6.1\text{-}58)$$

which can be reduced to [21]

$$|C(n)|^2 = \mathrm{Sin}^4(2\pi n/N);\quad n = 0,1,2,0,0,0,\ldots,N-1 \qquad (4.6.1\text{-}59)$$

When N becomes very large, the argument, $2\pi n/N$, approaches a continuous function of f; also, the argument becomes very small, i.e.,

$$\overset{\text{large}}{N}\quad \left[\frac{2\pi n}{N}\right] \rightarrow 2\pi f,\quad 0 < f < 1. \qquad (4.6.1\text{-}60)$$

Since

$$\mathrm{Sin}\,\omega = \omega - \frac{\omega^3}{3!} + \frac{\omega^5}{5!} - \frac{\omega^7}{7!} + \cdots;\ \omega^2 < \infty\ , \qquad (4.6.1\text{-}61)$$

then for

$$\overset{\text{large}}{N}\quad |C(n)|^2 = |C(j\omega)|^2 = \omega^4 \qquad (4.6.1\text{-}62)$$

which corresponds to the square of the magnitude of the continuous function second derivative operator.

$$\frac{d^2}{dt^2} \longleftrightarrow D^2 = (j\,\omega)^2 = -\omega^2 \qquad (4.6.1\text{-}63)$$

$$|D|^4 = \omega^4 \qquad (4.6.1\text{-}64)$$

4.6.2 Conversion of a Step-Like Waveform into a Duration Limited One

Given a waveform or sequence, f(k), comprised of N values, successive application of the DFT and IDFT yield an F(jω) and f(k) which are periodic of period, N. Two examples are shown in figure 4-30, where the given sequence of N points are at the left-hand side of the figure. The resultant IDFT's are periodic in which one cycle is identical to the original f(k).

For the sequence $f_a(k)$, the spectrum $F_a(n)$ corresponds to the intended-time sequence $f_a(k)$. In the case of $f_b(k)$ the spectrum $G_b(n)$ does not correspond to the intended-time sequence. Consider the following physical situation, the step-like waveform $f_b(k)$ is N points of a sequence that begins at zero and continues to n = ∞, remaining at the constant value B for $k \geq k_1$ where $k_1 \leq N-1$. Because the DFT will use only N points, the $f_b(k)$ is abruptly truncated at k = N-1. The resultant spectrum $G_b(n)$ is not the intended-spectrum, $F_b(n)$. The IDFT shows $g_b(k)$, a cycle of which is not the intended $f_b(k)$, i.e., the constant step-like waveform. The abrupt truncation $f_b(k)$ introduces the spectral components of an abrupt transition (step function).

A step-like sequence, f(k), which has achieved a constant value, B, for $k_1 \leq N-1$, can be converted to a duration-limited one, of length M = 2N, figure 4.31. The conversion is accomplished by extending and delaying f(k) by N units to obtain f(k-N), figure 4.31a and b. f(k) and f(k-N) both consist of M = 2N points. f(k-N) is subtracted from f(k) to yield $f_2(k)$.

$$f_2(k) = f(k) - f(k-N), \quad k = 0,1,2,\cdots M-1 \tag{4.6.2-1}$$

The DFT of $f_2(k)$ is given by

$$F_2(m) = F(m) - F(m)e^{-j\Omega TmN}, \quad m = 0,1,2,\cdots, M-1 \tag{4.6.2-2}$$

where F(m) is the DFT of f(k) and $\Omega = \frac{2\pi}{MT}$. $F_2(m)$ is the product

$$F_2(m) = (1 - e^{-j\Omega TmN}) F(m) \qquad m = 0,1,2,\cdots,M-1 \tag{4.6.2-3}$$

$$F_2(m) = F_1(m) F(m) \tag{4.6.2-4}$$

where

$$F_1(m) = 1 - e^{-j\Omega TmN} = 1 - e^{-jm\pi} \quad m = 0,1,2,\cdots, M-1 \tag{4.6.2-5}$$

$$|F_1(m)| = 2\,|\mathrm{Sin}\, m\pi/2| \tag{4.6.2-6}$$

$$\phi_1(m) = \mathrm{Tan}^{-1} \frac{\mathrm{Sin}\, m\pi}{1-\mathrm{Cos}\, m\pi} \tag{4.6.2-7}$$

Because

$$|F_1(m)| = \begin{cases} 0, & m = 0,2,4\cdots \\ 2, & m = 1,3,5\cdots \end{cases} \tag{4.6.2-8}$$

$$\phi_1(m) = 0, \quad m = 0,1,2,\cdots \tag{4.6.2-9}$$

the DFT's of F(n) and $F_2(m)$ are simply related. Figure 4-32 shows F(n) and $F_2(m)$ where m has been used in the argument of F_2. Note that $|F_2(m)|$ is twice the value of $|F(n)|$ for m = 2n-1. Explicitly, F(n) and $F_2(m)$ are $F(n\Omega)$ and $F_2[m(\Omega/2)]$. The discrete frequency $n\Omega$ corresponding to n = N-1 is equal to the discrete frequency $m\Omega/2$ with m = M-2 because M = 2N. Thus, the spectrum of $F_2(m)$ has (a) the same number of non-zero values, N, as that for F(n), but are graphically spaced twice as far apart, and (b) the magnitude of $|F_2(m)|$ is twice that of $|F(n\Omega)|$ at the non-zero values.

In deconvolution studies the effect of the time domain conversion of step-like waveforms for x(k) and y(k) yields an estimated impulse response due to a unit impulse applied at k=0, followed by a negative unit impulse applied at k=N-1, figure 4.33.

4.6.3 Application of the One Parameter Method

4.6.3.1 Procedural Summary

In the discussion presented here, the notation will be that employed in Section 4.5; consequently, d(k) will be more precisely specified by $d_2(k)$; similarly $D_1(n)$ will replace D(n), and so on for other quantities as the occasion arises. The reason for doing so, is to emphasize that the quantities contain errors. In the sections before 4.5, the errors were not delineated.

The procedure starting with the time domain sequences x(k) and y(k) and ending with the deconvolved time domain impulse sequence $d_2(k)$ progresses as follows:

1. Offset correction for x(k) and y(k). Eliminate the offset, i.e., insure that x(o) = x(N-1) = 0 and y(o) = y(N-1) = 0. The offset error is systematically introduced by the data acquisition device. If x(k) and y(k) are step-like waveforms so that x(N-1) and y(N-1) are not zero, they must be transformed to impulsive waveforms. See subsection 4.6.2.
2. DF transform x(k) and y(k) to obtain $X_1(n)$ and $Y_1(n)$; display $|X_1(n)|_{dB}$ and $|Y_1(n)|_{dB}$, (4.5.3-9), (4.5.3-10).
3. Calculate the complex ratio $H_1(n) = Y_1(n)/X_1(n)$, (4.5.3-11); display $|H_1(n)|_{dB}$.
4. Select the regularization filter coefficient γ; then filter $H_1(n)$ to yield $D_1(n)$, (4.5.3-14) display $|D_1(n)|_{dB}$.
5. IDF transform $D_1(n)$ to obtain the time domain deconvolved result, $d_2(k)$. (4.5.3-14); display $d_2(k)$.
6. Calculate the error criteria (4.5.13-19); display e(k) and the values for $e_{max.}$, $e_{min.}$, $\bar{e}$, and σ.
7. Judge the quality of the deconvolved impulse response estimate, $d_2(k)$, in terms of the error criteria; if unsatisfactory, return to step 4 and change the filter parameter.
8. If the quality is satisfactory, print the $d_2(k)$ graphic display, related error criteria, and the numerical values.

Three examples of the one-parameter deconvolution method will now be presented; they are:

1. The estimation of the insertion impulse response of 300 meters of RG 58 C/U coaxial cable; this example uses the same insertion waveform data as was used in the two-parameter method example, Section 4.4.1.
2. The insertion impulse response of a filter for a step-like, 100 ps transition duration, reference waveform generator.

The first example allows the reader to compare the one and two parameter methods; the second example, in addition to illustrating the one parameter method, illustrates the procedure for using step-like waveform data rather than duration limited waveforms.

4.6.3.2 Estimation of the Insertion Impulse Response of 300 Meters of RG 58 C/U Coaxial Cable

The insertion waveforms x(k) and y(k) are the same ones used in the two parameter method experiments in Section 4.4.2, figures 4.9 and 4.10, respectively. Three degrees of regularization were used, γ = 100, 500, and 10,000. The estimated insertion impulse responses and their corresponding error functions, e(k), and filtered (regularized) spectrum magnitudes are shown in figures 4.34 through 4.36. The regularization filter was that of (4.6.1-49),

$$R(n) = \frac{|X(n)|^2}{|X(n)|^2 + \gamma|C(n)|^2}$$

which is a real filter.

When γ = 100, the regularization is moderate. Compare figure 4.34 with figures 4.7b and 4.12 of the two parameter method. The results are comparable. Increasing γ to 500 provides strong regularization. Compare figure 4.35 with figures 4.7c and 4.13 of the two parameter method; note that the Gibbs phenomena has appeared in the results of the two parameter method, figure 4.7c, while the one parameter method produces no Gibbs oscillations. Increasing γ to 10,000 provides still stronger regularization with no evidence of Gibbs oscillations (compare figures 4.36 to 4.7c and 4.13). Furthermore, figure 4.36 very clearly shows the small variations (reflections) discussed in the sixth paragraph of section 4.4.2. Thus, it is seen that the one parameter method uses a smooth filtering that does not introduce the Gibbs effect (as was pointed out earlier in section 4.4.2).

4.6.3.3 The Insertion Impulse Response of a Filter for a Step-like, 100 ps Transition Duration, Reference Waveform Generator

In this example, the insertion impulse response of a liquid filled coaxial transmission line is estimated. The coaxial line is about 20 cm long and is filled with a dilute binary solution in which heptance is the solvent and heptanone is the solute. Because heptanone is a polar molecule (heptane is non-polar), the solution comprises a dilute polar dielectric and as such, exhibits a dielectric relaxation law. Consequently, the frequency dependence of the dielectric constant, e(s) causes the transmission line to behave as a low pass filter. The filter is used to shape a 20 ps transition duration (10%-90%) step-like waveform into a 100 ps transition duration waveform [25].

The insertion step-like waveforms x(k) and y(k) are converted to duration limited waveforms, $\bar{x}_1(k)$ and $\bar{y}_1(k)$, and the DFT's $\bar{X}_1(n)$ and $\bar{Y}_1(n)$ are obtained, respectively, figure 4.37. The effects of filtering on the DFT's are shown in figure 4.38; here, the real filter, R(n), of (4.6.1-49),

$$R(n) = \frac{|\bar{X}(n)|^2}{|X(n)|^2 + \gamma|C(n)|^2}$$

has been used for $\gamma = 0, 10^2, 10^4$, and 10^6. R(n) operates on $\bar{H}_1(n)$ to yield $\bar{D}_1(n)$

$$\bar{D}_1(n) = R(n)\ \bar{H}_1(n)$$

and the IDFT of $\bar{D}_1(n)$ is $\bar{d}_2(k)$. The estimated impulse response for very strong regularization, $\gamma = 10^6$ is shown in figure 4.39. Note that $\bar{d}_2(k)$

$$\bar{d}_2(k) = \bar{y}(k)(1/*)\ \bar{x}(k)$$

is the response due to a pair of impulses as shown in figure 4.33, and $d_2(k)$ is the desired estimated impulse response, the first pulse of the pair in $\bar{d}_2(k)$.

This example completes the discussion of frequency domain deconvolution. In the next chapter, some time domain deconvolution methods will be presented.

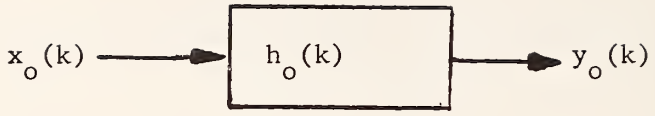


Figure 4.1 $x_o(k)$ and $y_o(k)$ are the exact sequences which are uniquely related through the system impulse response sequence $h_o(k)$.

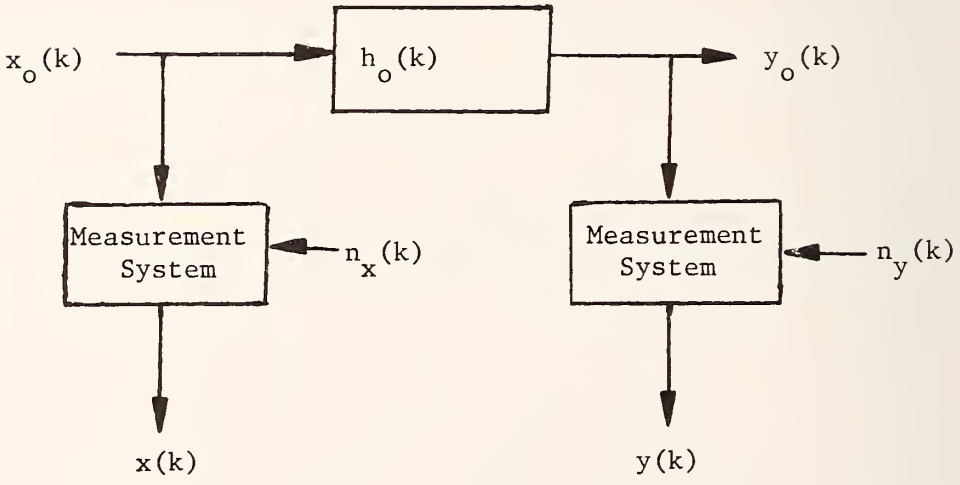


Figure 4.2 x(k) and y(k) are corrupted versions of $x_o(k)$ and $y_o(k)$. The noise sources $n_x(k)$ and $n_y(k)$ are independent of each other.

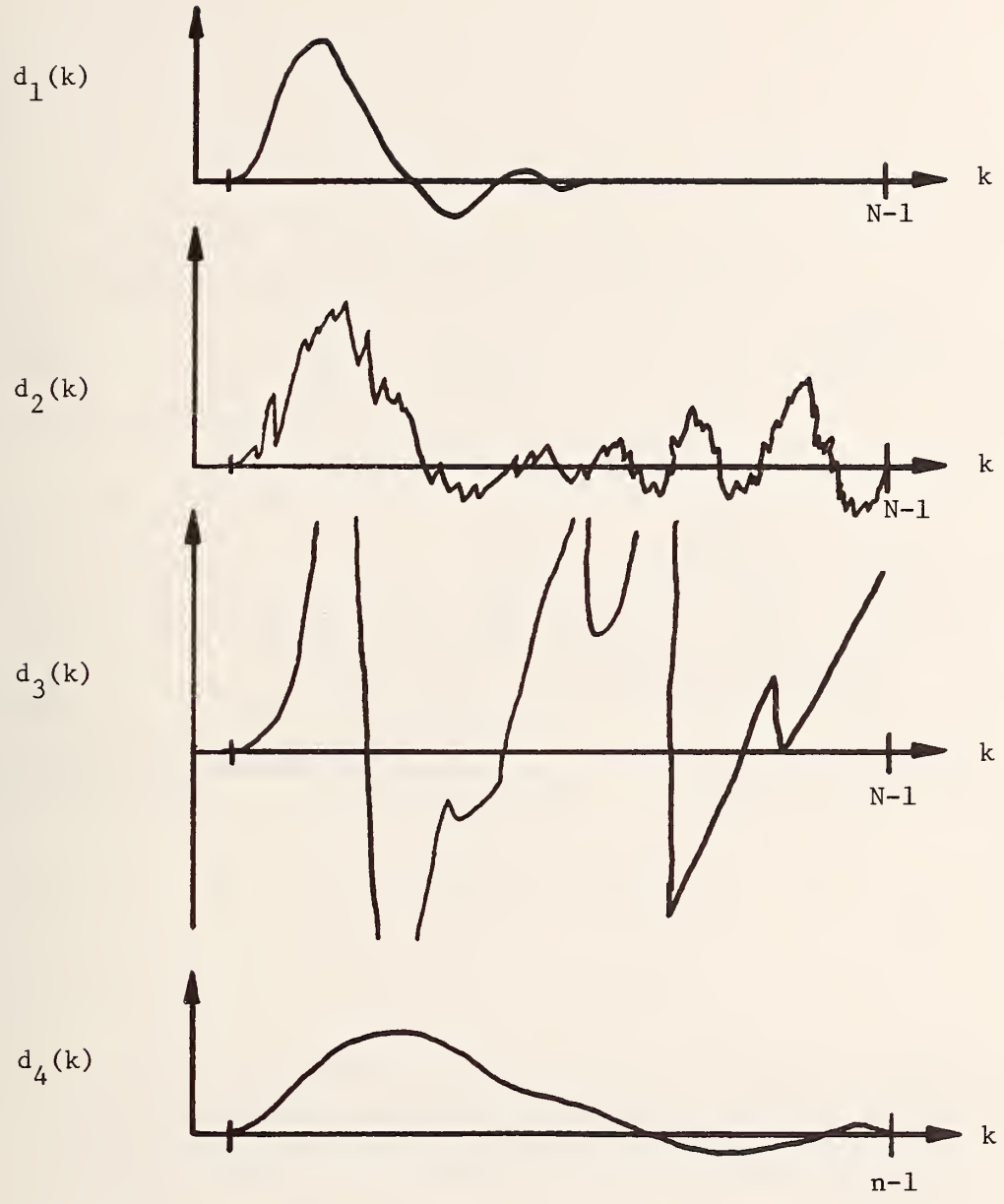


Figure 4.3 The effects of filtering. $d_1(k)$, the stable response resulting from the "proper" amount of filtering, $h_o(k)$ in the operator's judgment. $d_2(k)$, almost stable, needs more filtering. $d_3(k)$, very unstable, very little filtering, if any. $d_4(k)$, stable, too much filtering, response is slowed down.

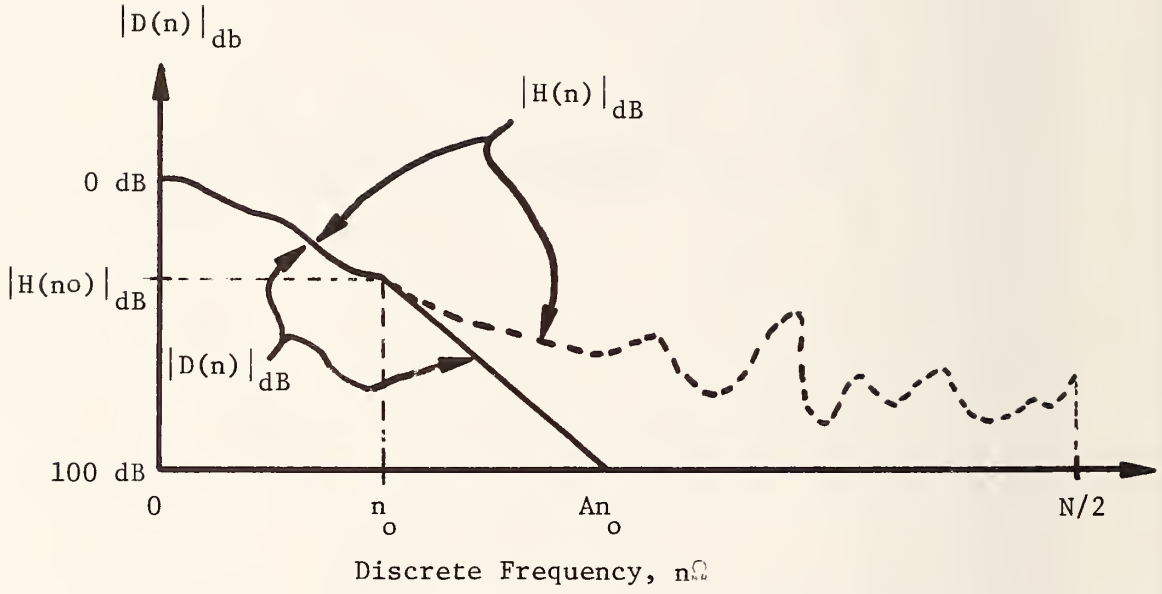


Figure 4.4 The magnitude, $|D(n)|_{dB}$, of the filtered H(n) versus n Ω in dB, 20 log $|D(n)|$, and the magnitude of the unfiltered $|H(n)|_{dB}$.

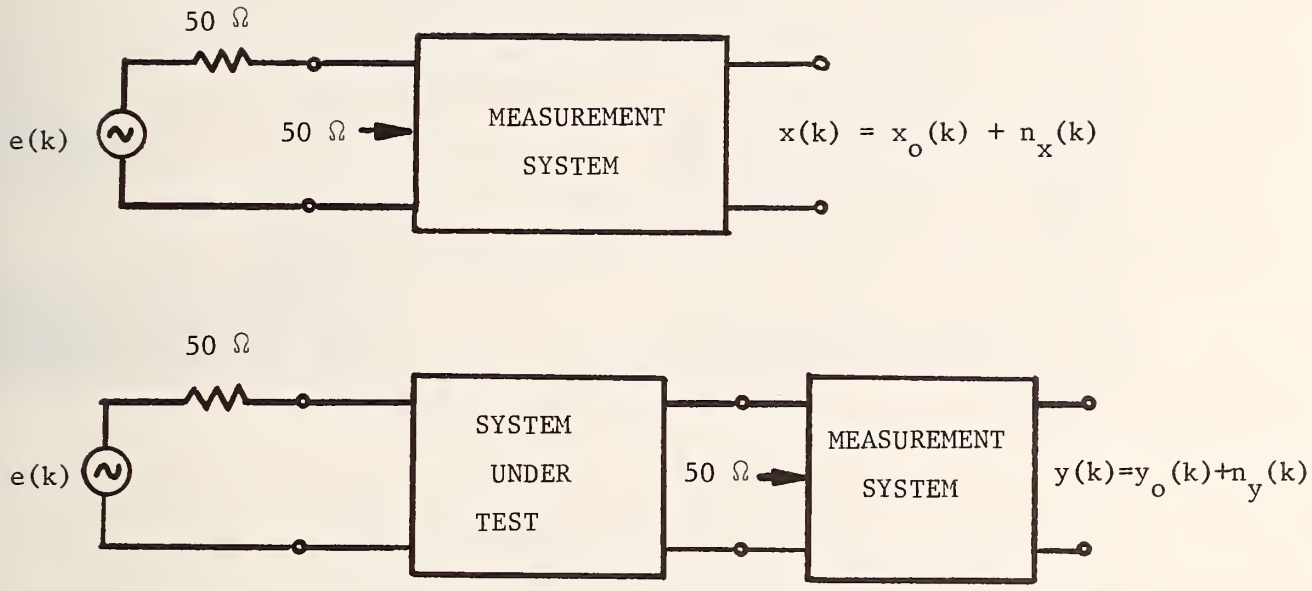
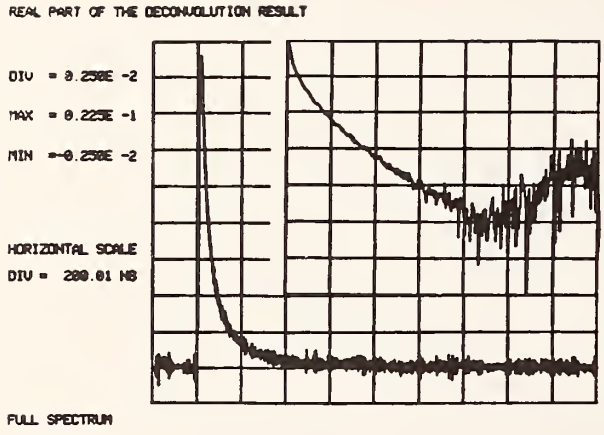
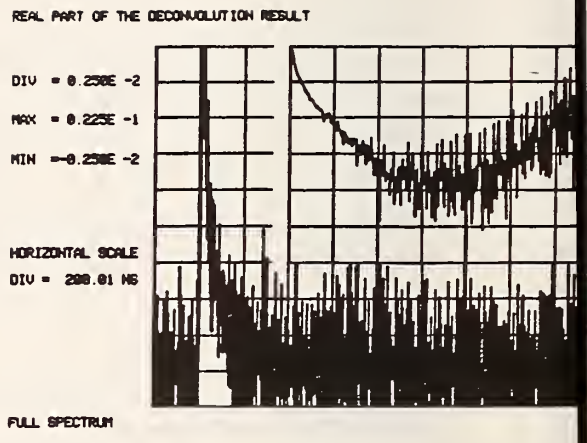


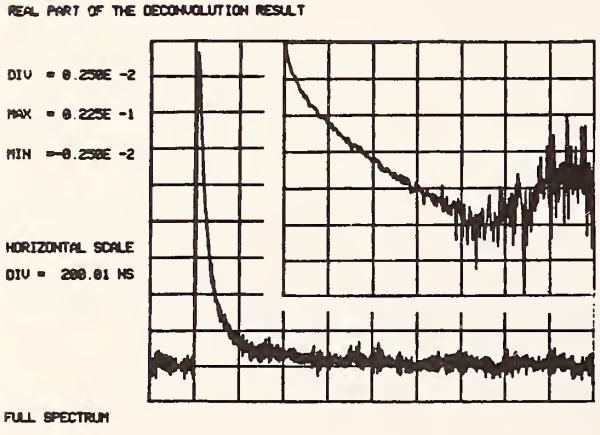
Figure 4.5 Insertion response measurement method. The observables are the waveforms x(k) and y(k) obtained before and after the system under test is inserted between the 50 Ω impedance generator and measurement system, respectively. $H(n) = Y(n)/X(n) = S_{21}$, the insertion scattering parameter [17].



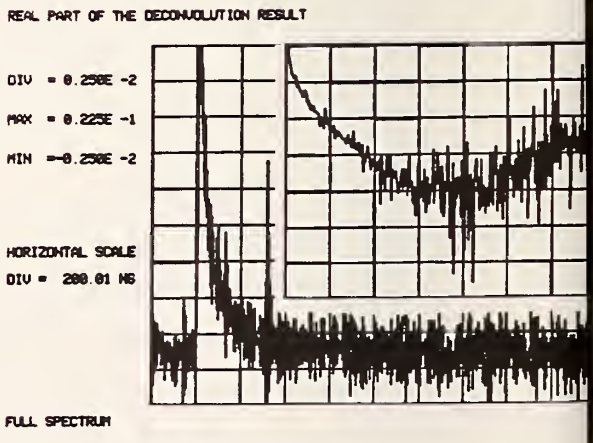
(a)



(b)



(C)



(d)

Figure 4.6 Deconvolution without any regularization, i.e., full spectrum of H(n) is used; h(k) is shown for various input and output signal to noise ratios. (a) SNR-x 75 dB, SNR-y 46 dB. (b) SNR-x 75 dB, SNR-y 36 dB. (c) SNR-x 45 dB, SNR-y 46 dB. (d) SNR-x 40 dB, SNR-y 40 dB. $|H(n)|_{db}$ is shown in the upper right hand portion of each figure with the vertical scale of -10 dB/division commencing at 0.0 dB and the horizontal scale of 25 MHz/division commencing at 0.0 MHz.

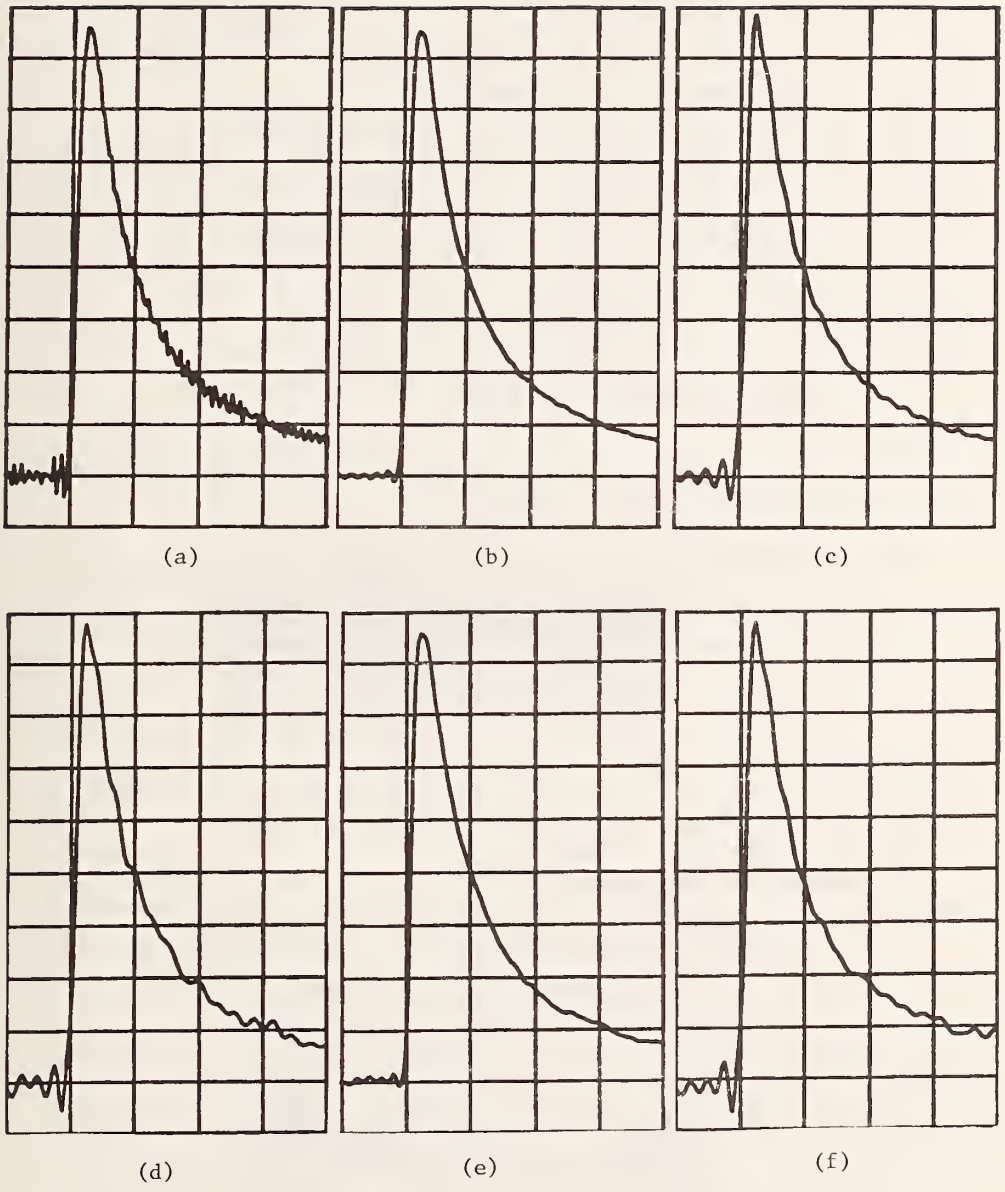


Figure 4.7 Deconvolution with regularization; d(k) is shown, (b-f), for various input and output signal to noise ratios and degrees of filtering. SNR-x 75 dB, SNR-y 46 dB: (a) h(k), no regularization, $n_o\Omega$ = 250 MHz, A = 0; (b) moderate regularization, $n_o\Omega$ = 100 MHz, A = 5; (c) strong regularization, $n_o\Omega$ = 74 MHz, A = 6. (d) SNR-x 75 dB, SNR-y 36 dB, strong regularization, $n_o\Omega$ = 74 MHz, A = 6. (e) SNR-x 45 dB, SNR-y 46 dB, moderate regularization, $n_o\Omega$ = 100 MHz, A = 5. (f) SNR-x 40 dB, SNR-y 40 dB, strong regularization, $n_o\Omega$ = 74 MHz, A = 6. Vertical scale 2.5 x 10^{-3}/Div., Horizontal scale 50 ns/Div.

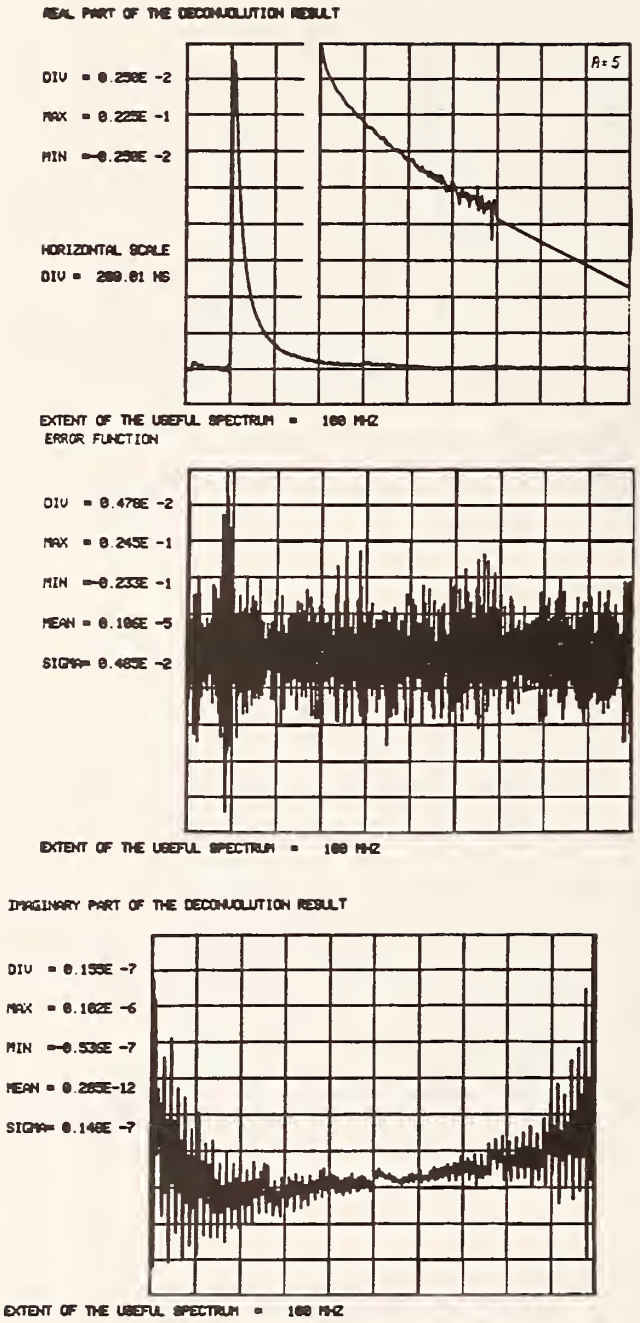


Figure 4.8 An example of the imaginary part of the deconvolution result. If the computation procedures of the DFT (FFT) were exact, the imaginary part would vanish.

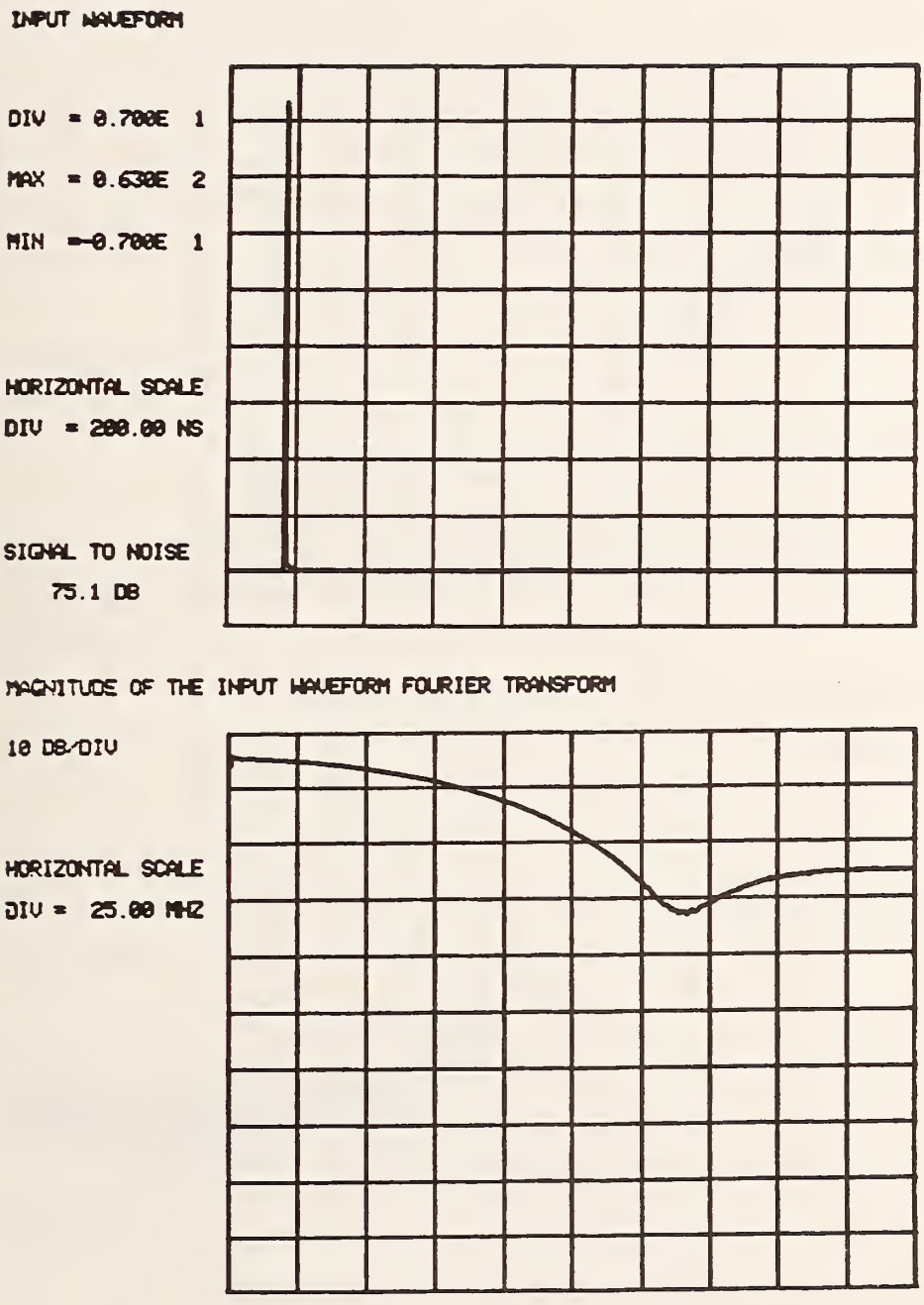


Figure 4.9 The input waveform x(k) and the magnitude of its DFT, $|X(n)|$. SNR-x 75 dB. x(k) is shown with an abscissa of 50 ns/div. in figure 5.15.

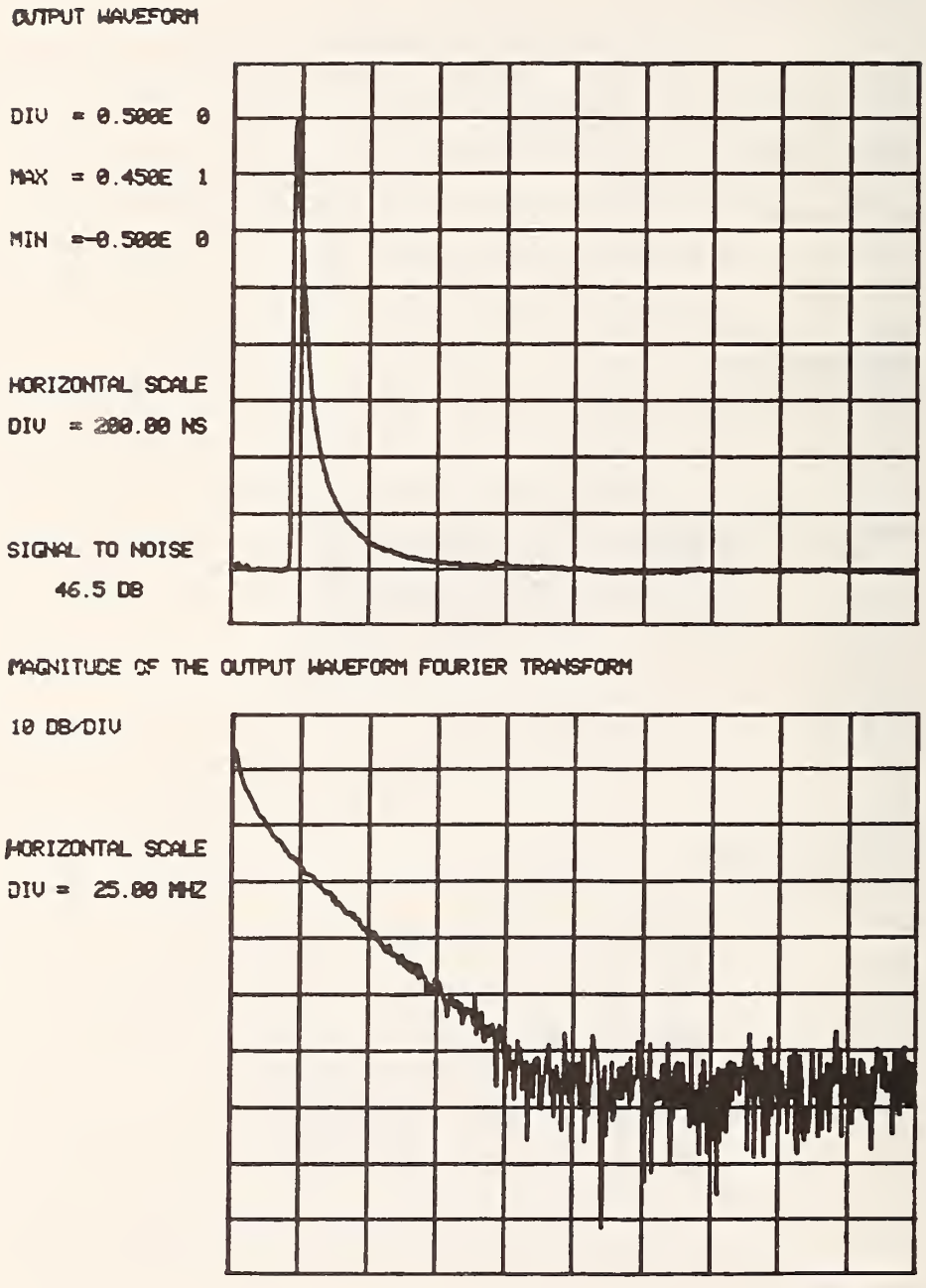


Figure 4.10 The output waveform, y(k), and the magnitude of its DFT, $|Y(n)|$. SNR-y 46 dB. y(k) is shown with an abscissa of 50 ns/div. in figure 5.15.

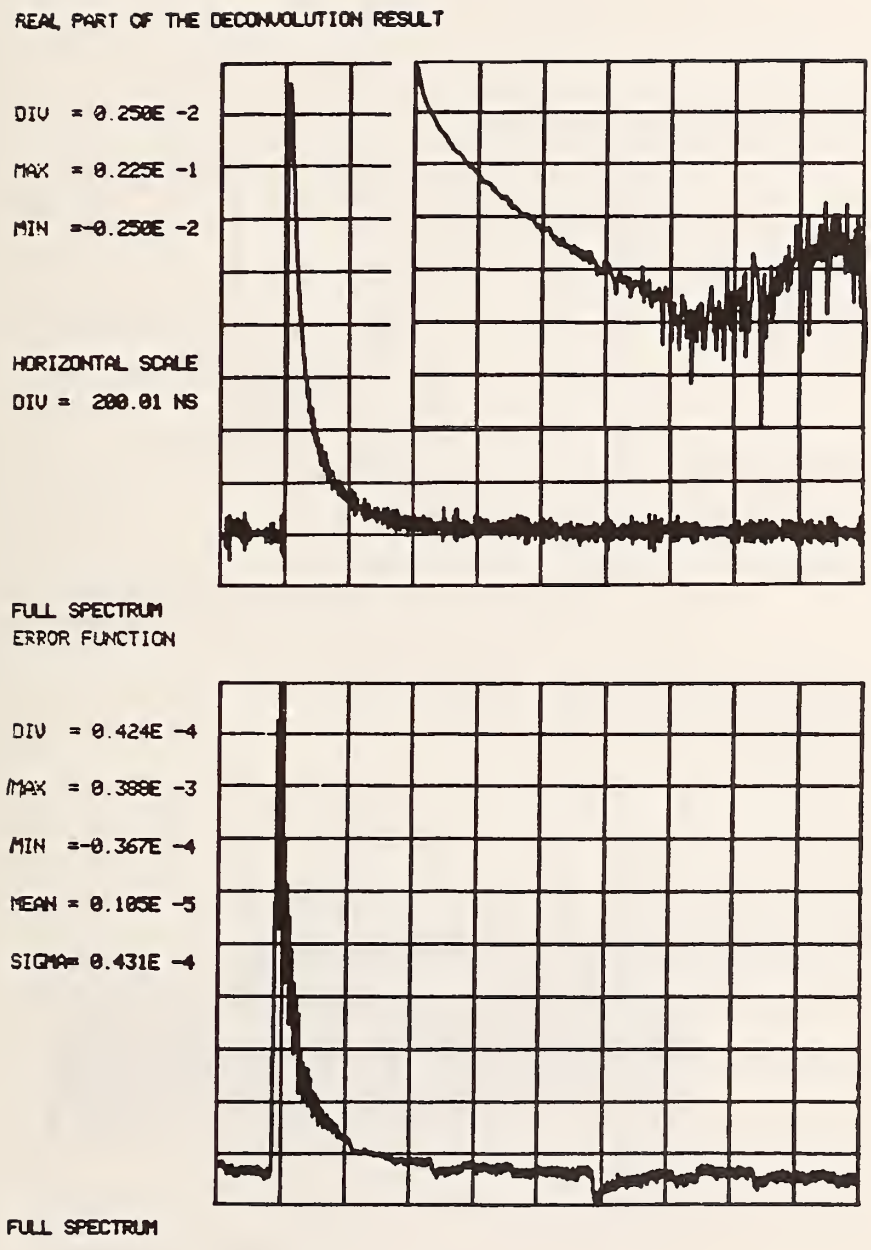


Figure 4.11 SNR-x 75 dB, SNR-y 46 dB. Deconvolution without regularization, $|H(n)|_{dB}$, h(k), and e(k). $|H(n)|_{dB}$ is in the upper right hand portion of the upper figure; vertical scale -10 dB/Div. commencing at 0.0 dB, horizontal scale 25 MHz/Div. commencing at 0.0 MHz. No regularization used; this relatively smooth result for h(k) is possible only because of the large values of SNR-x and SNR-y (in particular SNR-y). Note the small value of the error function e(k) and that its shape is comparable to y(k).

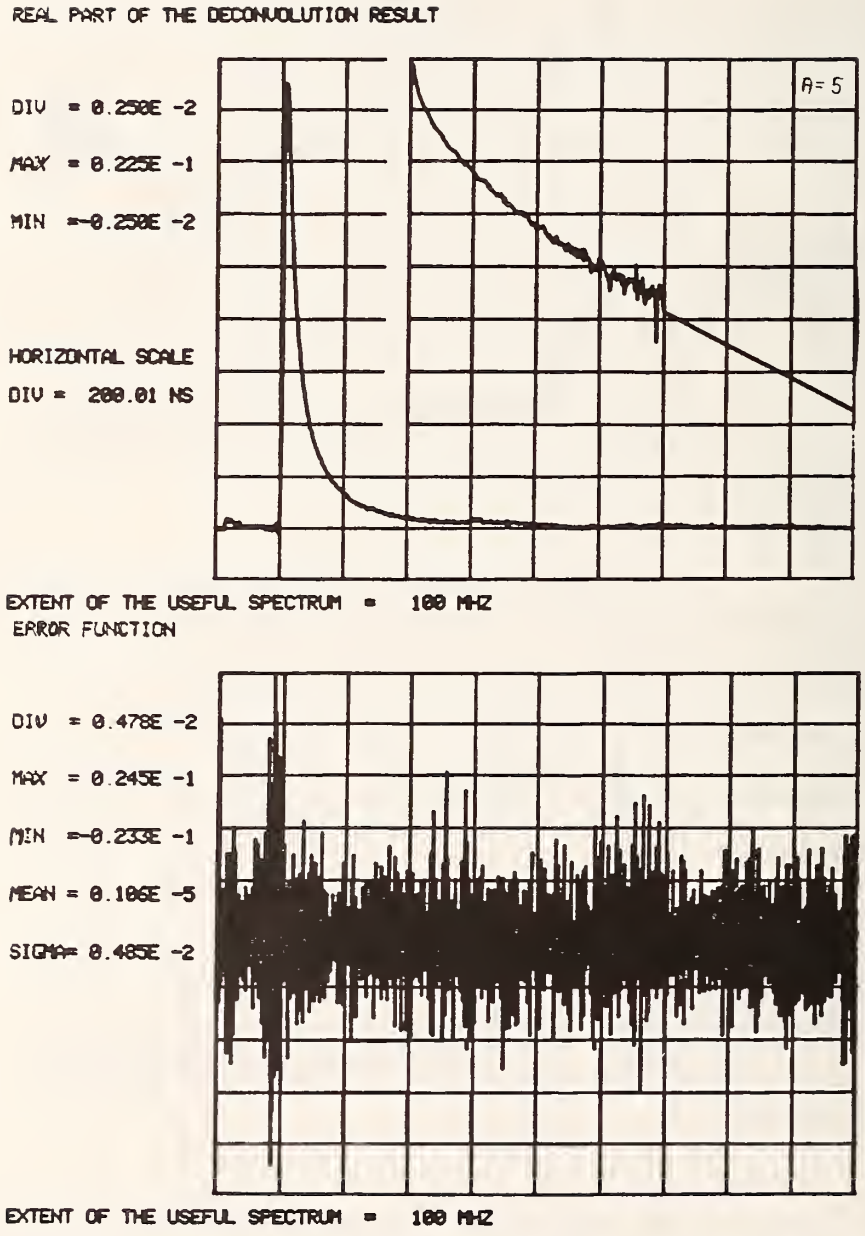


Figure 4.12 SNR-x 75 dB, SNR-y 46 dB. Deconvolution with moderate regularization $n_o\Omega$ = 100 MHz, A = 5. $|D(n)|_{dB}$, d(k), and e(k). Note the filtering of H(n) to obtain D(n). $|D(n)|_{dB}$ is shown in the upper right hand portion of the upper figure, Vertical - 10 dB/Div. commencing at 0.0 dB, horizontal 25 MHz/Div. commencing at 0.0 MHz.

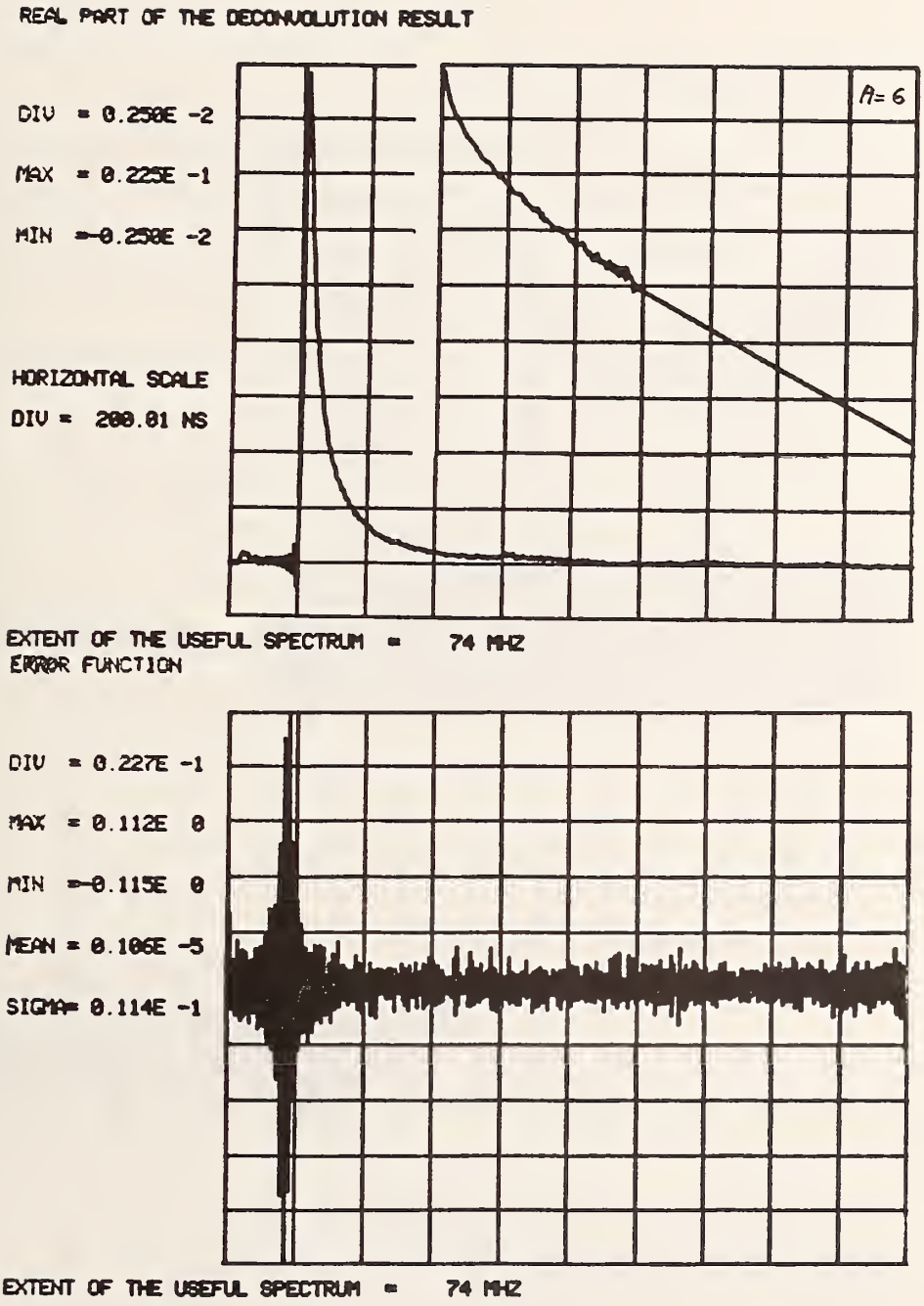


Figure 4.13 SNR-x 75 dB, SNR-y 46 dB. Deconvolution with strong regularization $n_o\Omega$ = 74 MHz, A = 6. $|D(n)|$, d(k), and e(k). Note the filtering of H(n) to obtain D(n). $|D(n)|_{dB}$ is shown in the upper right hand portion of the upper figure, vertical -10 dB/Div. commencing at 0.0 dB, horizontal 25 MHz/Div. commencing at 0.0 MHz.

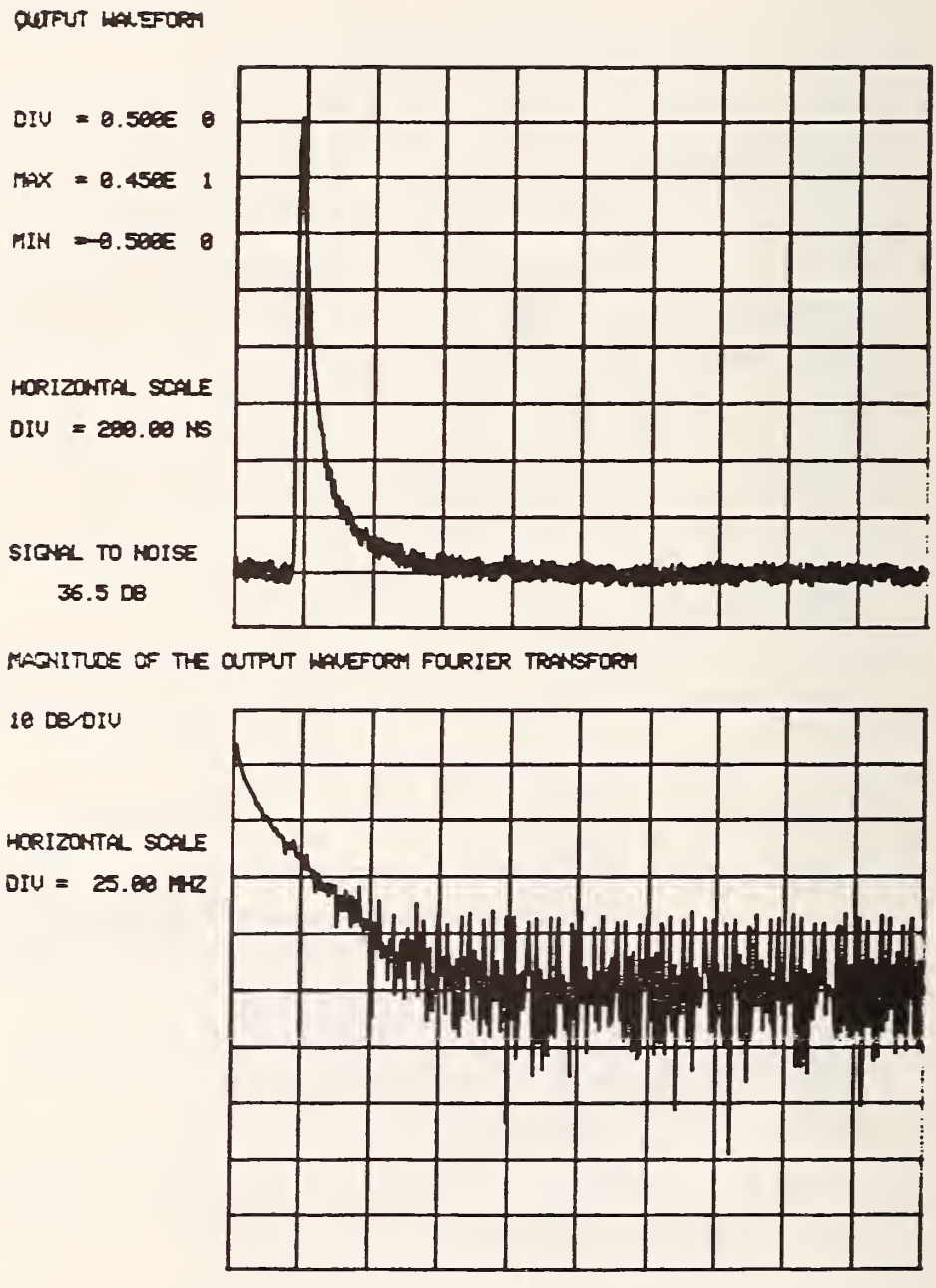


Figure 4.14 The modified output waveform, y(k), and the magnitude of its DFT, $|Y(n)|_{dB}$, after the addition of 10 dB of noise to yield a SNR-y of 36 dB.

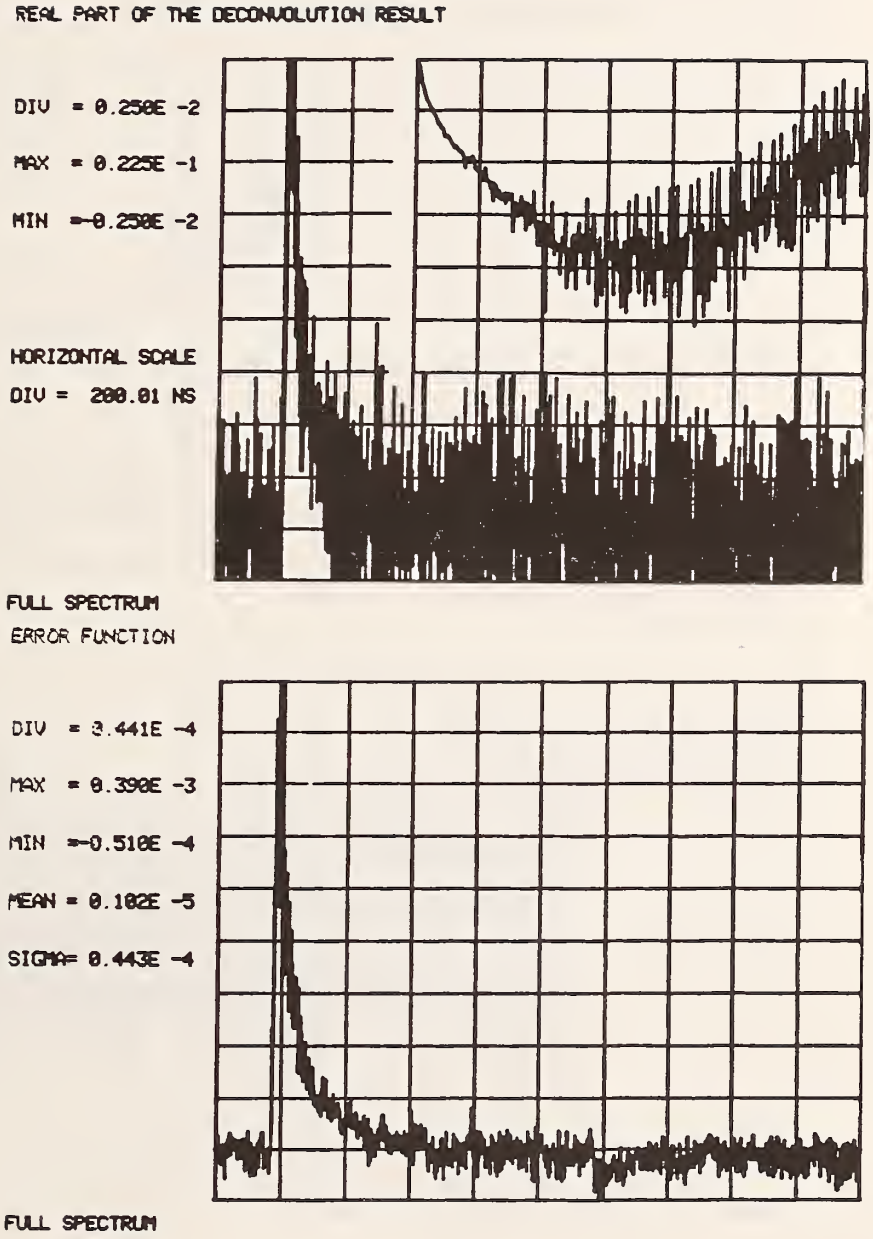


Figure 4.15 SNR-x dB, SNR-y 36 dB. Deconvolution without regularization $|H(n)|_{dB}$, h(k) and e(k) $|H(n)|_{dB}$ is in the upper right hand portion of the upper figure, vertical scale -10 dB/Div. commencing at 0.0 dB, horizontal scale 25 MHz/Div. commencing at 0.0 MHz.

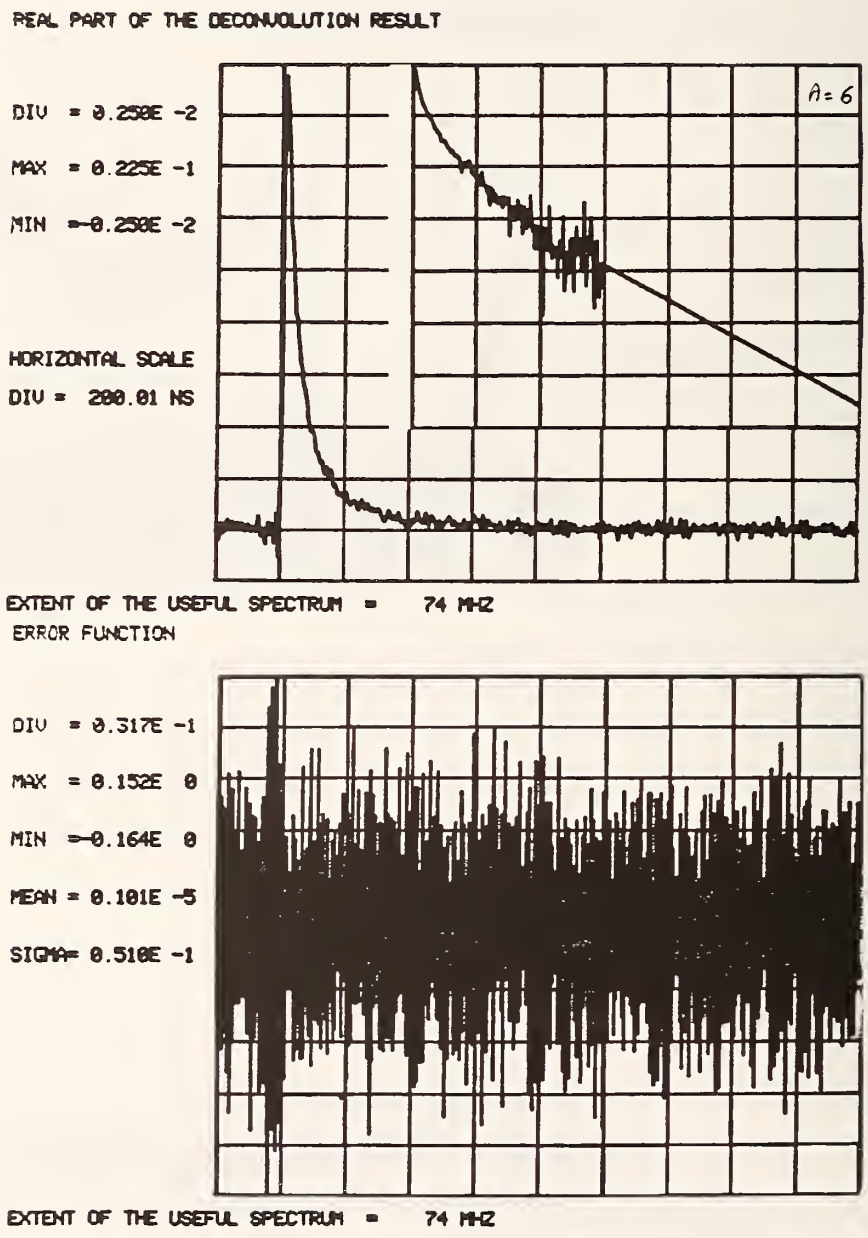


Figure 4.16 SNR-x 75 dB, SNR-y 36 dB. Deconvolution with strong regularization, $n_o\Omega$ = 74 MHz, A = 6. $|D(n)|_{dB}$, d(k) and e(k). Note the filtering of H(n) to obtain D(n). $|D(n)|_{dB}$ is shown in the upper right hand portion of the upper figure, Vertical -10 dB/Div. commencing at 0.0 dB, horizontal 25 MHz/Div. commencing at 0.0 MHz.

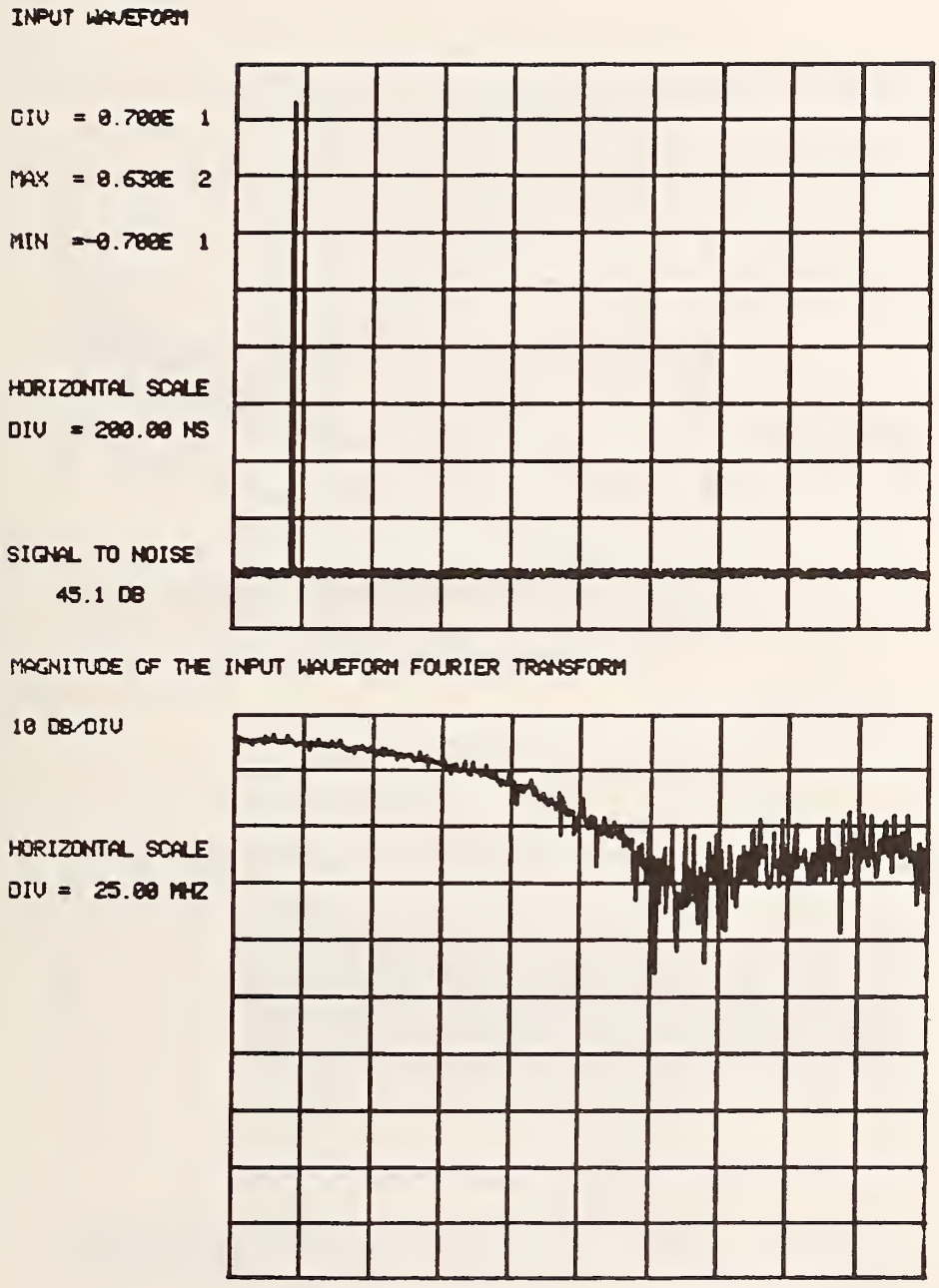


Figure 4.17 The modified input waveform, x(k) and the magnitude of its DFT, $|X(n)|_{dB}$, after the addition of 30 dB of noise to yield a SNR-x of 45 dB.

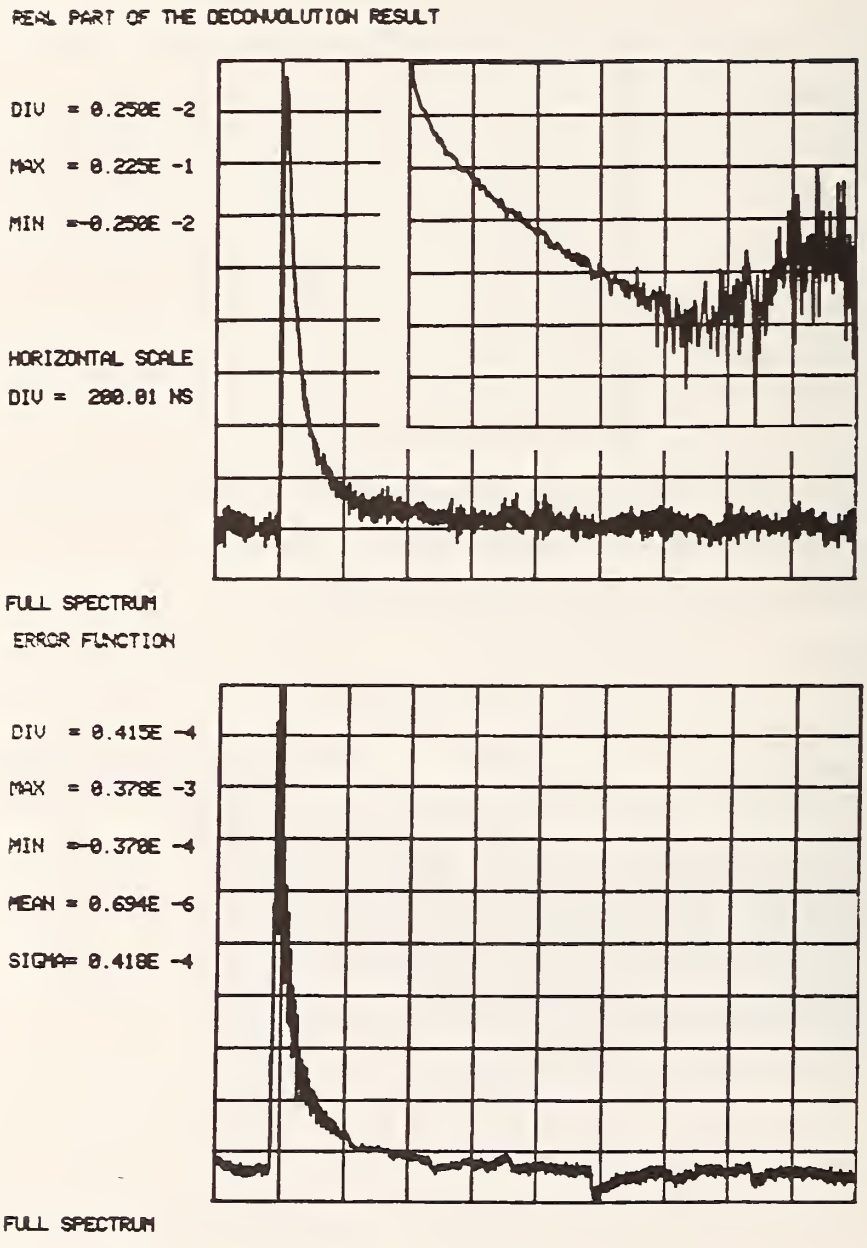


Figure 4.18 SNR-x, 45 dB, SNR-y, 46 dB. Deconvolution without regularization, $|H(n)|_{dB}$, h(k), and e(k). $|H(n)|_{dB}$ is in the upper right hand portion of the upper figure, vertical scale -10 dB/Div. commencing at 0.0 dB, horizontal scale 25 MHz/Div. commencing at 0.0 MHz.

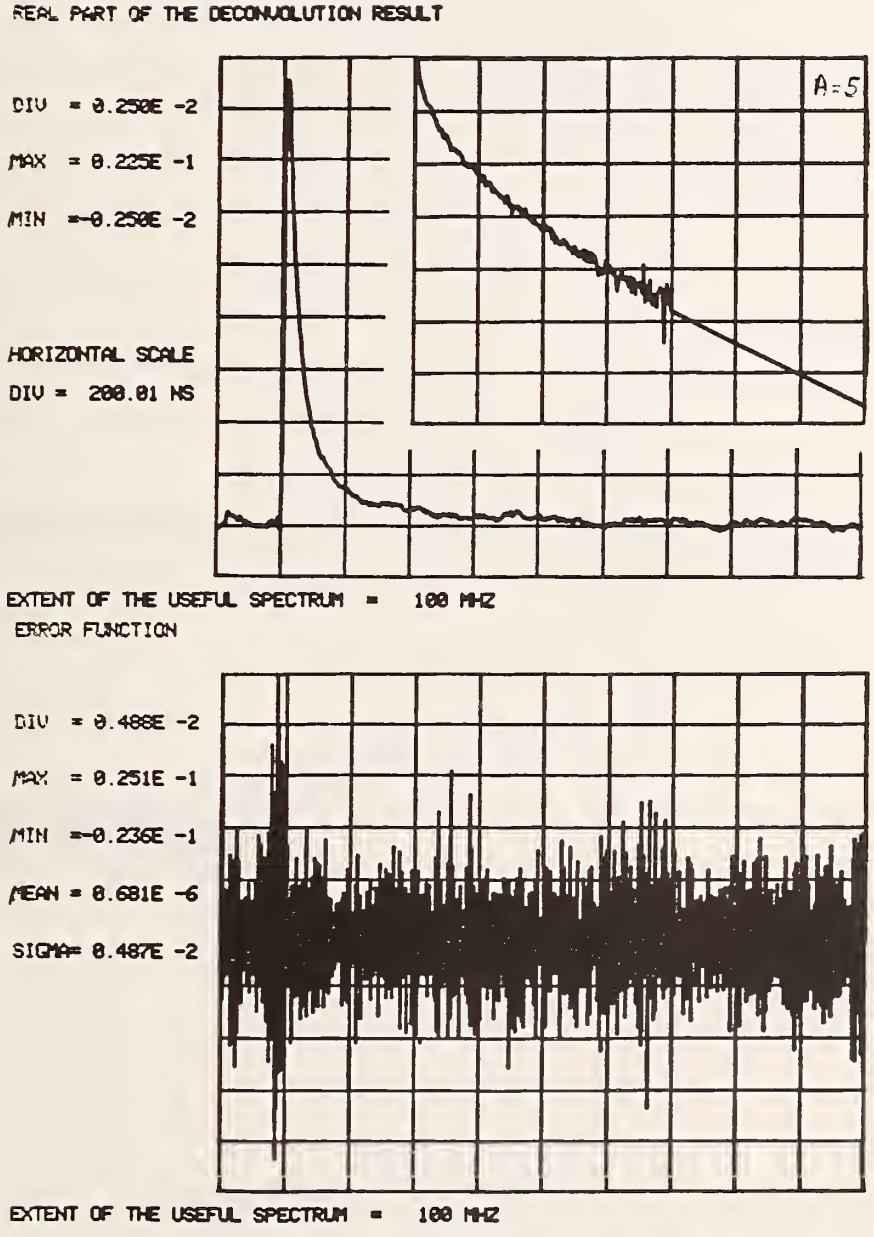


Figure 4.19 SNR-x 45 dB, SNR-y 46 dB. Deconvolution with moderate regularization, $n_o\Omega$ = 100 MHz, A = 5. $|D(n)|$dB, d(k), and e(k). Note filtering of H(n) to obtain D(n). $|D(n)|_{dB}$ is shown in the upper right hand portion of the upper figure, vertical scale - 10 dB/div. commencing at 0.0 dB, horizontal scale, 25 MHz/Div., commencing at 0.0 MHz.

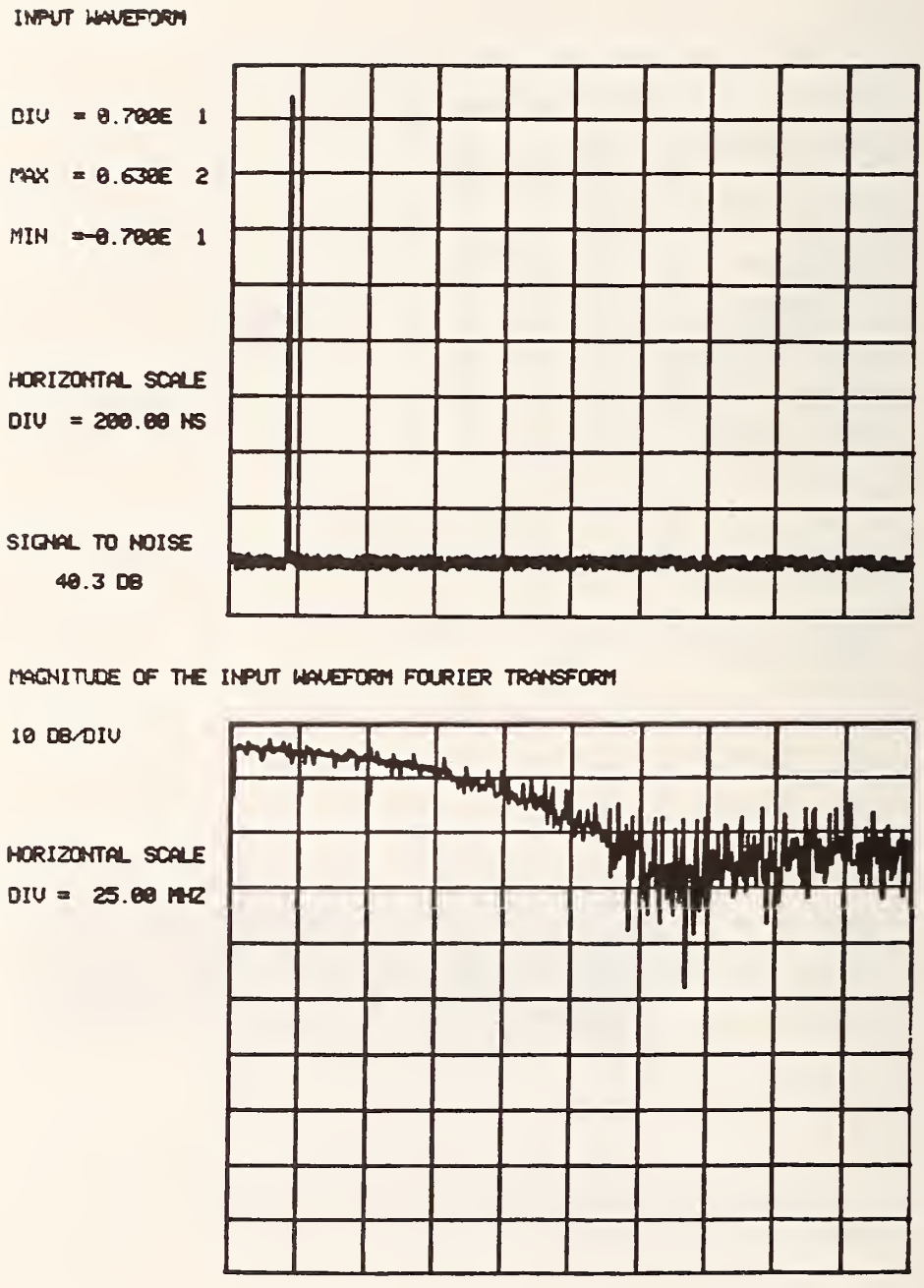


Figure 4.20 The modified input waveform, x(k), and the magnitude of its DFT, $|X(n)|_{dB}$, after the addition of 35 dB of noise to yield a SNR-x, 40 dB.

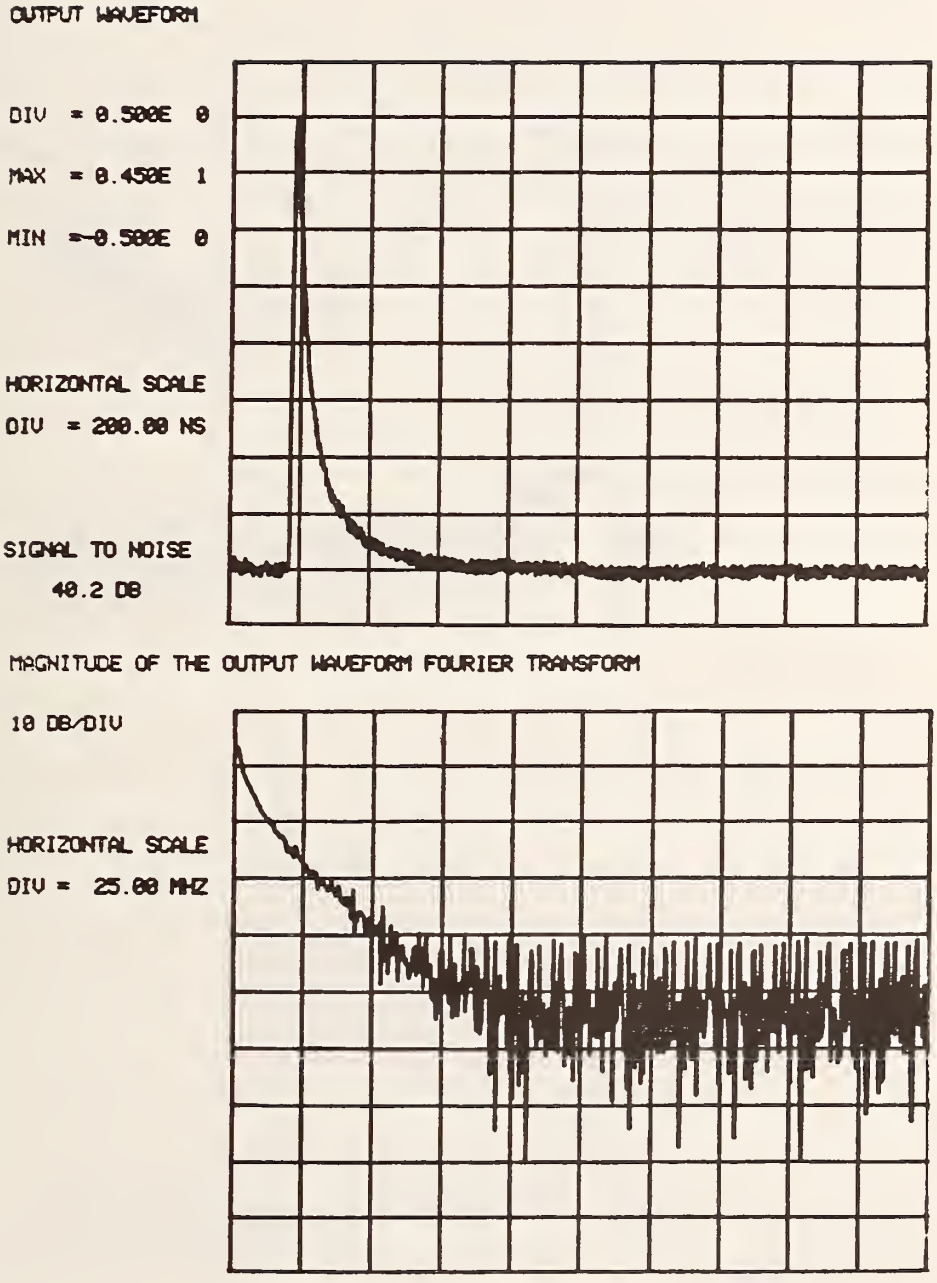


Figure 4.21 The modified output waveform, y(k), and the magnitude of its DFT, $|Y(n)|_{dB}$, after the addition of 6 dB of noise to yield a SNR-y of 40 dB.

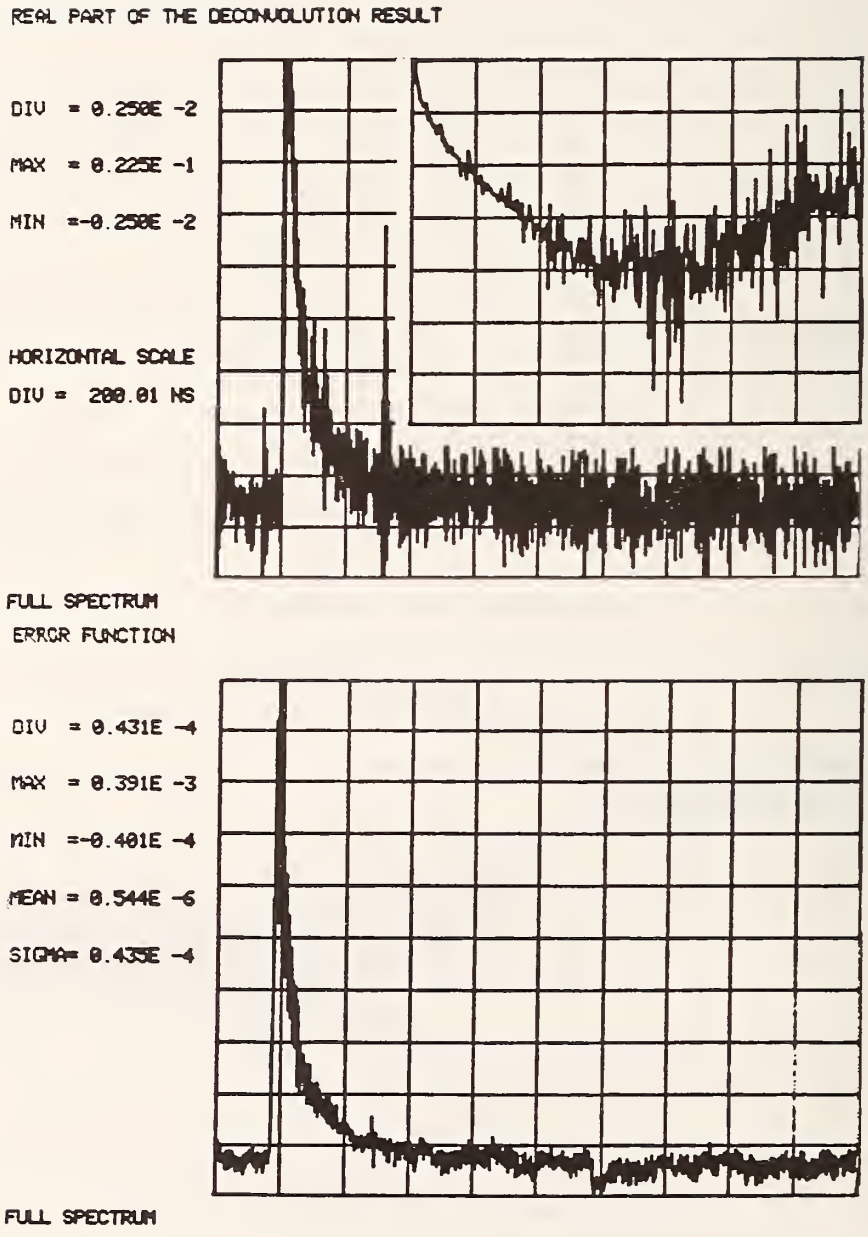


Figure 4.22 SNR-x 40 dB, SNR-y 40 dB. Deconvolution without regularization, $|H(n)|_{dB}$, h(k), and e(k). $|H(n)|_{dB}$ is in the upper right hand portion of the upper figure, vertical scale -10 dB/Div. commencing at 0.0 dB, horizontal scale 25 MHz/Div. commencing at 0.0 MHz.

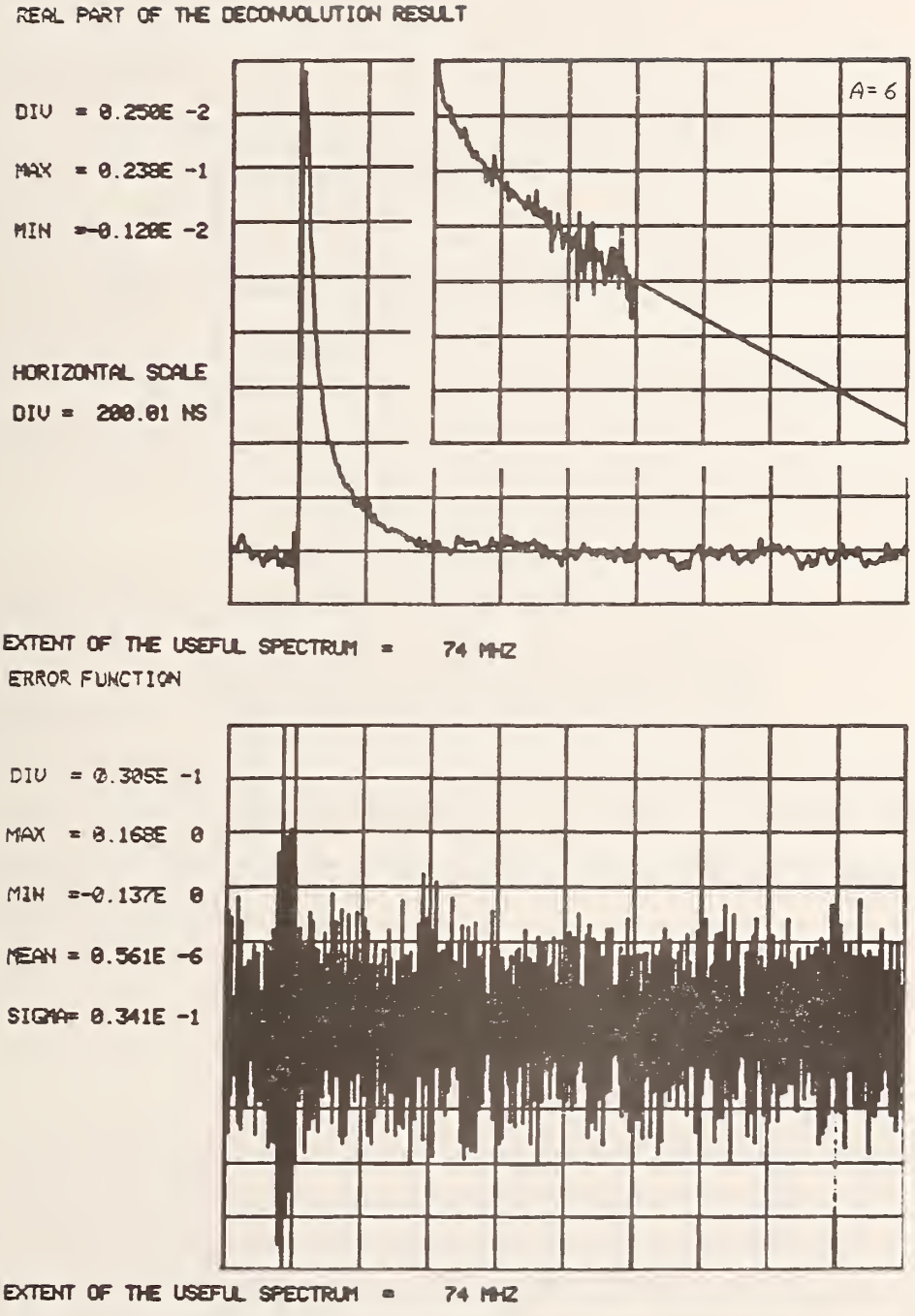


Figure 4.23 SNR-x 40 dB, SNR-y 40 dB. Deconvolution with strong regularization, $n_o\Omega$ = 74 MHz, A = 6. $|D(n)|_{dB}$, d(k), and e(k). Note filtering of H(n) to obtain D(n). $|D(n)|_{dB}$ is shown in the upper right hand portion of the upper figure, vertical scale -10 dB/Div. commencing at 0.0 dB, horizontal scale 25 MHz/Div. commencing at 0.0 MHz.

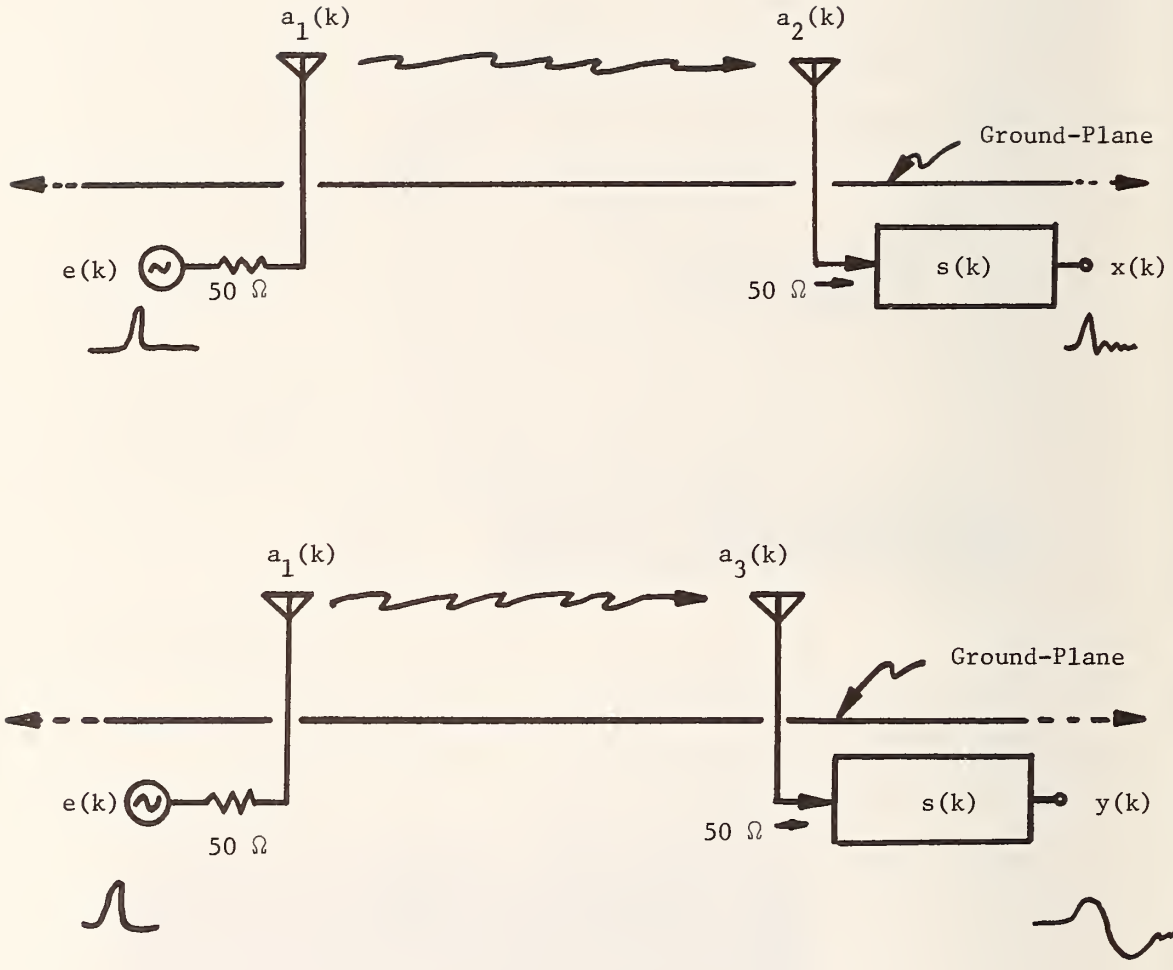


Figure 4.24 The time domain antenna range insertion measurement method. The reference input waveform x(k) is produced by the system comprised of the 50 ohm source e(k), the source antenna $a_1(k)$, the receiving antenna $a_2(k)$, and the waveform measurement system s(k). The measurement output voltage y(k) is produced when $a_3(k)$ replaces $a_2(k)$.

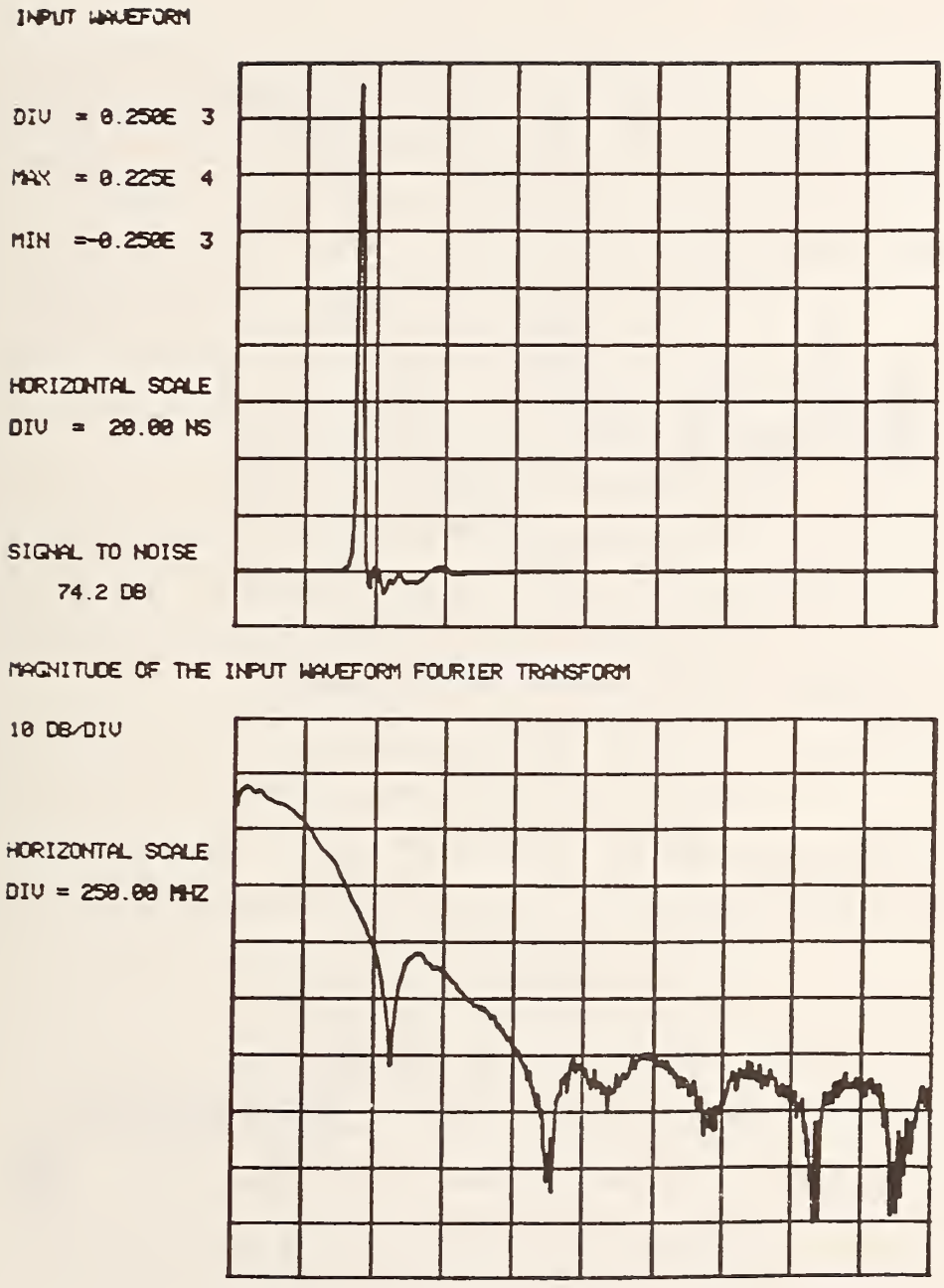


Figure 4.25 The input waveform x(k) and the magnitude of its DFT, $|X(n)|$.

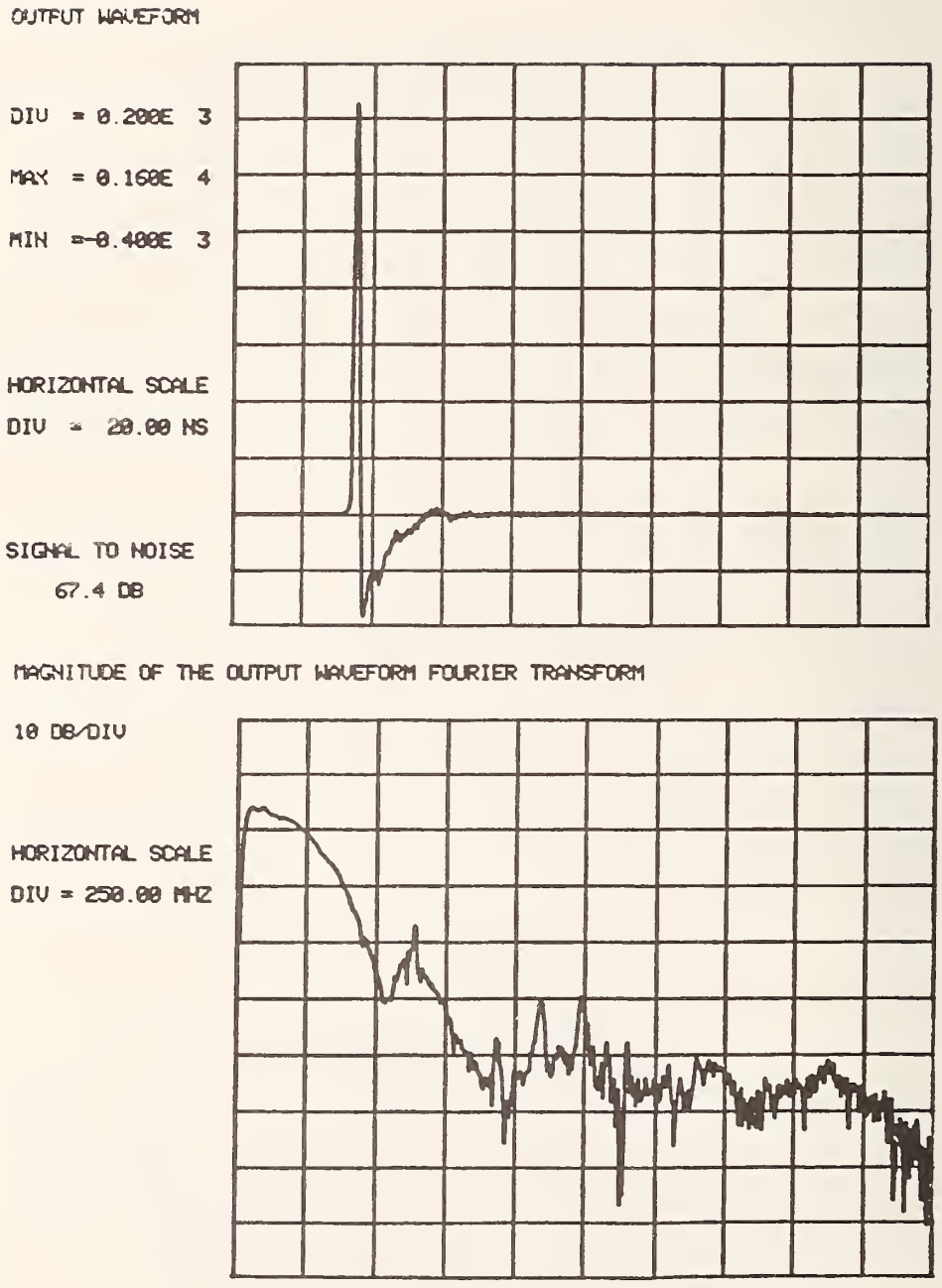


Figure 4.26 The output waveform y(k) and the magnitude of its DFT, $|Y(n)|$.

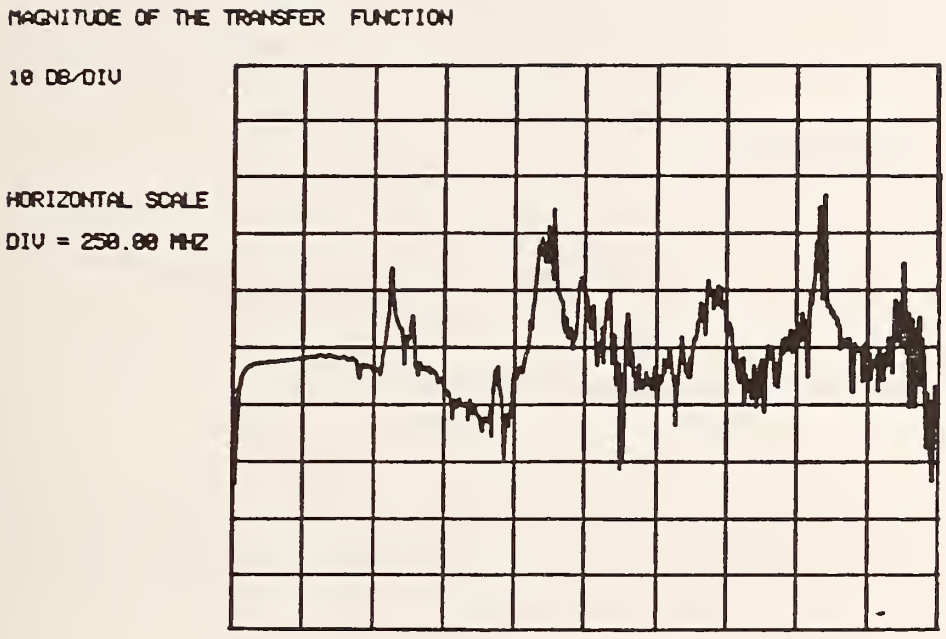
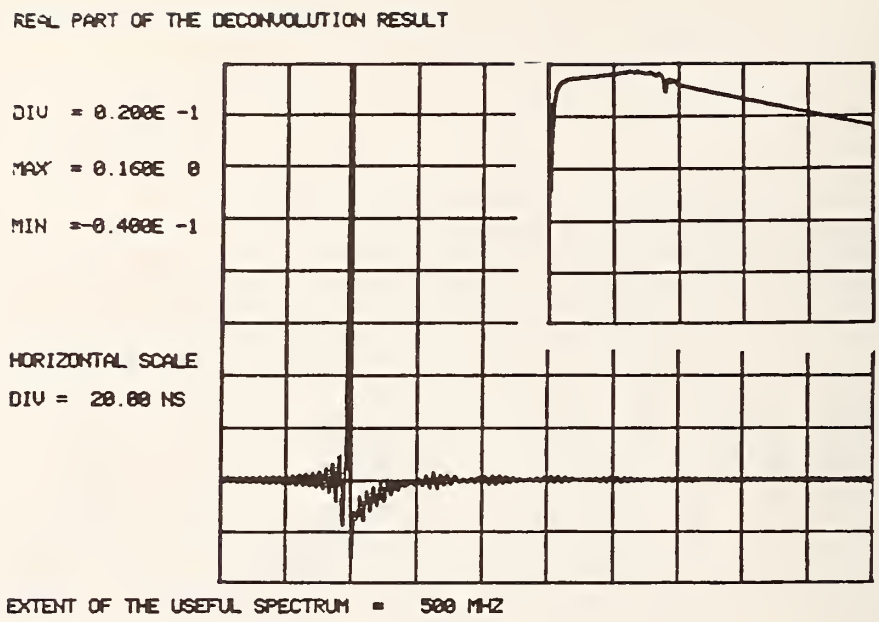
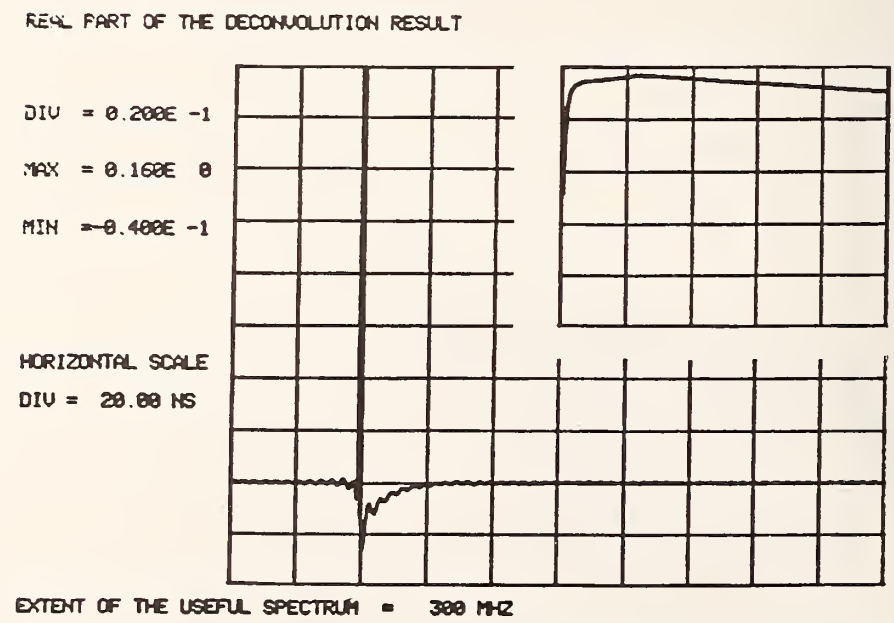


Figure 4.27 The magnitude $|H(n)| = |Y(n)| \div |X(n)|$



(a) Filter Parameters: $n_o\ \Omega$ = 500 MHz, A = 20 $|D(n)|$ is shown in upper right hand corner with scales as in Figure 4.27. Error Parameters: $|e(k)|_{max}$ = 46, σ = 43, $[|e(k)|_{max} \div |y(k)|_{max}] \times 100 = [46/1600] \times 100 = 2.9\%$



(b) Filter Parameters: $n_o\ \Omega$ = 300 MHz, A = 100, $|D(n)|$ is shown in upper right hand corner with scales as in Figure 4.27. Error parameters: $|e(k)|_{max}$ = 80, σ = 7, $[|e(k)|_{max} \div |y(k)|_{max}] \times 100 = [80/1600] \times 100 = 5\%$.

Figure 4.28

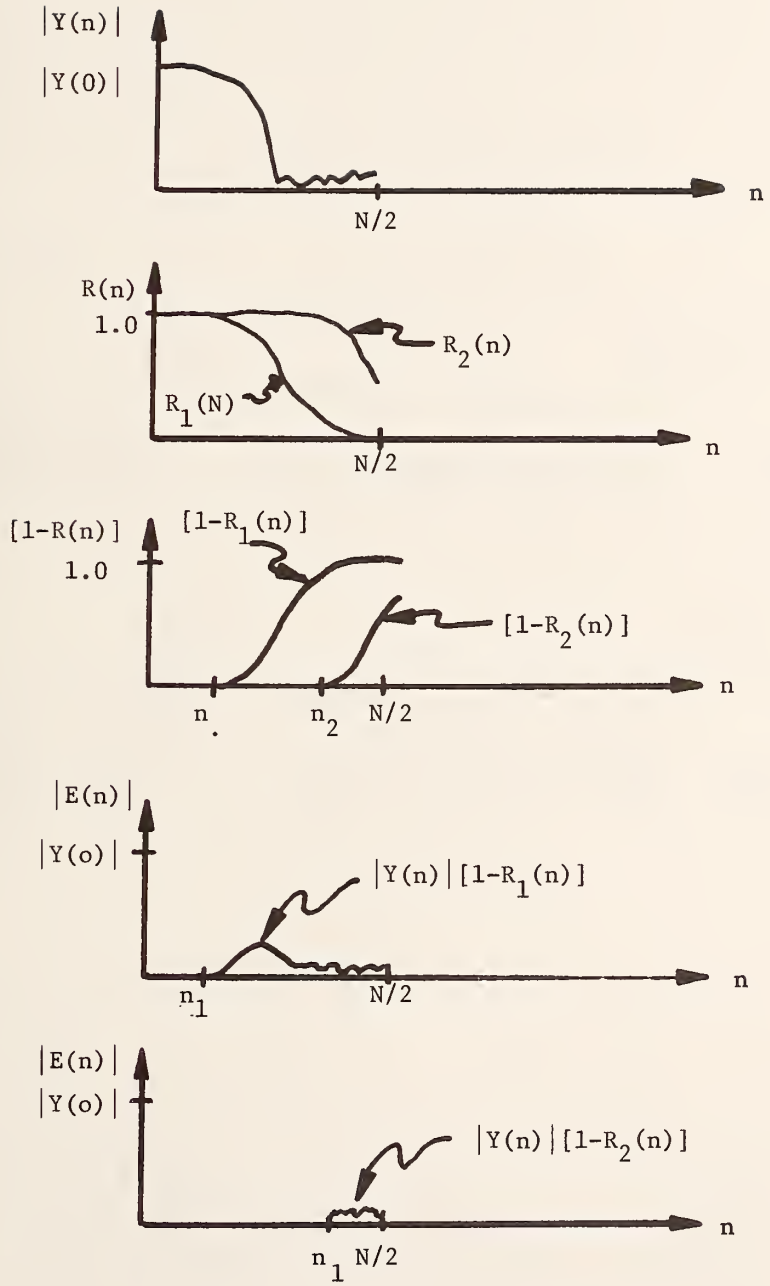


Figure 4.29 Filtering increases the absolute error $|E(n)|$.

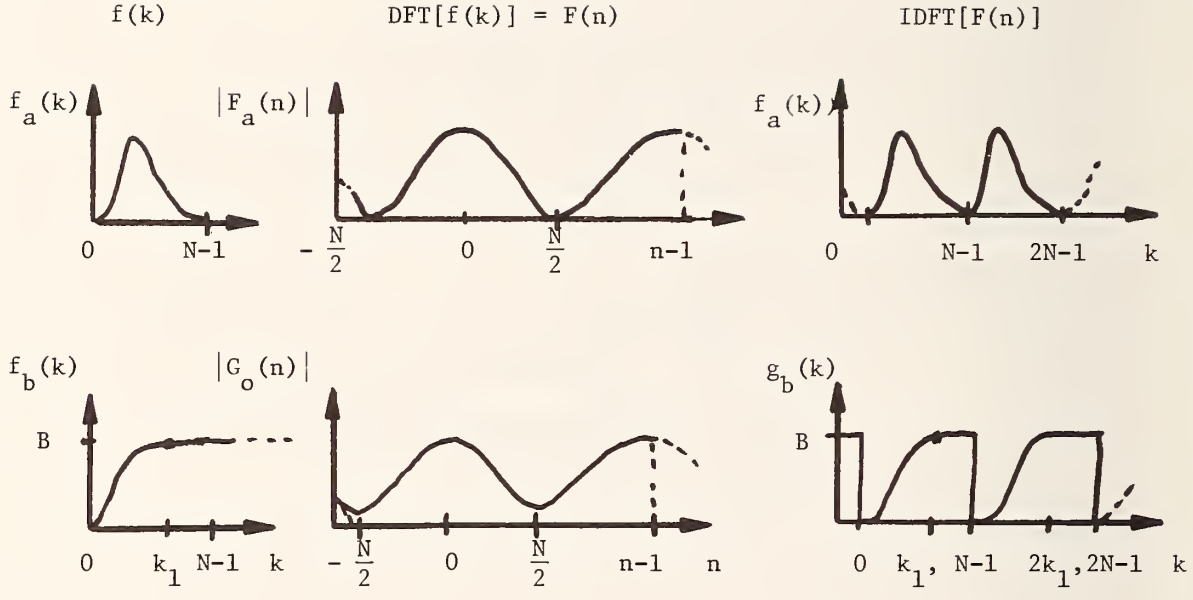


Figure 4-30 Moving from left to right, successive application of the DFT and IDFT yield periodic functions for $|F(n)|$ and $|f(k)|$. A cycle of the resultant $g_b(k)$ may not be the intended $f_b(k)$.

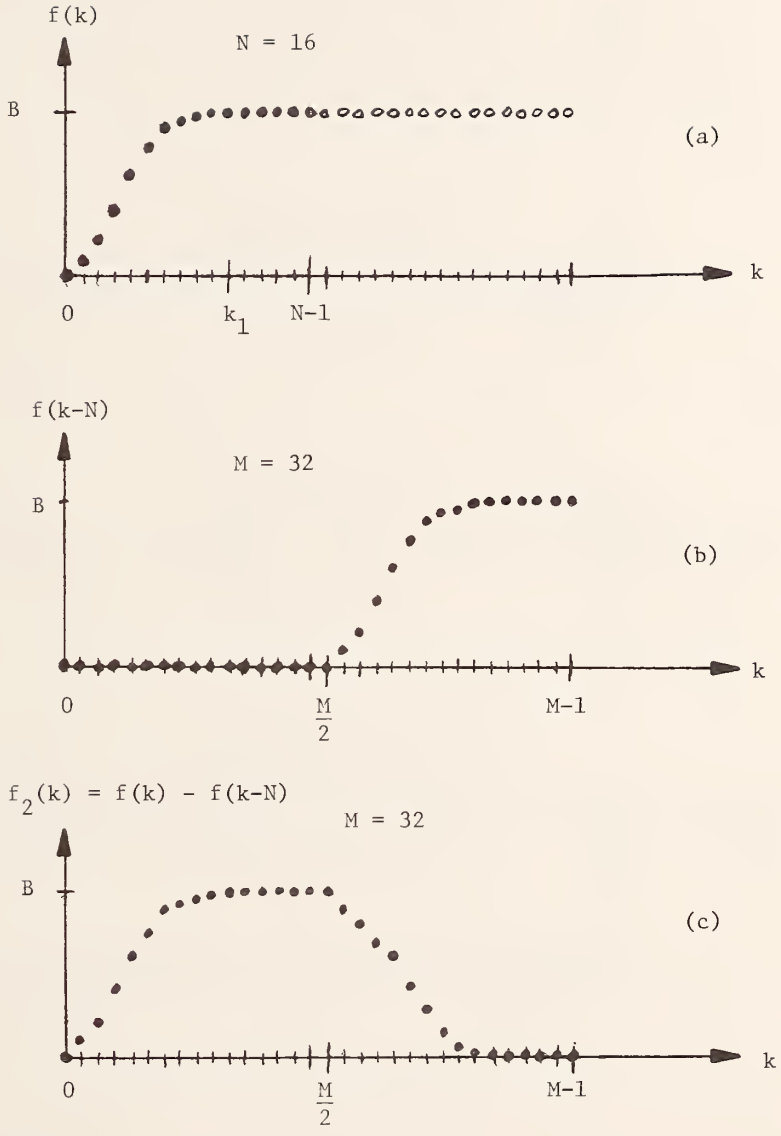


Figure 4-31. (a) The original sequence, f(k), (dark points) is N units in duration and is constant for $k \geq k_1$. f(k) is extended by N units (light points).

(b) The extended f(k) is delayed by N units to yield f(k-N). f(k) and f(k-N) both consist of M = 2N points.

(c) f(k-N) is subtracted from f(k) to yield $f_2(k)$.

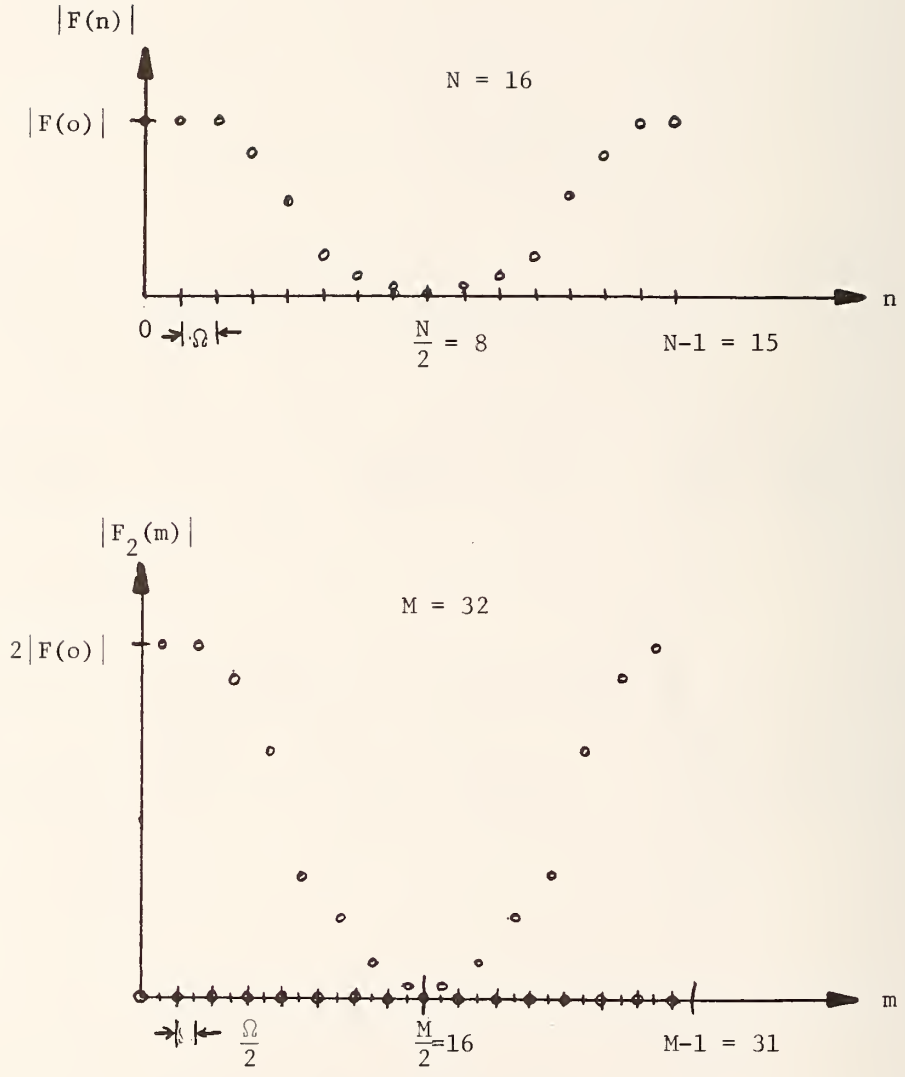


Figure 4-32. The magnitude of the spectra for F(n) and $F_2(m)$. The discrete frequency spacings for F(n) and $F_2(m)$ are Ω and $\Omega/2$ due to the time windows NT and 2NT, respectively. M=2N, thus the total number of non-zero values in $F_2(m)$ is the same as in F(n).

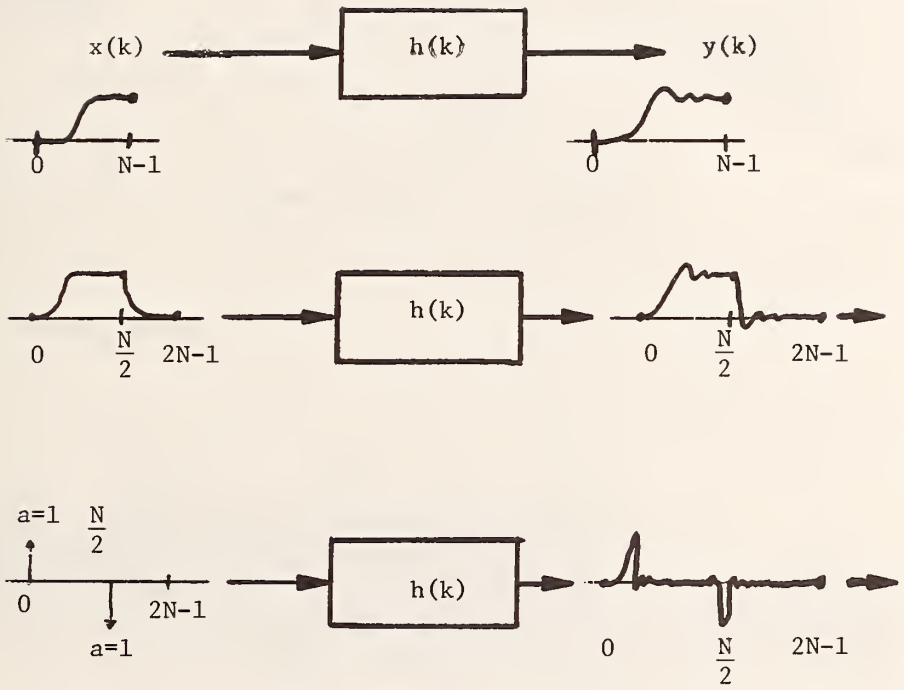


Figure 4-33 Deconvolution using converted step-like waveforms yields a pair of impulse responses.

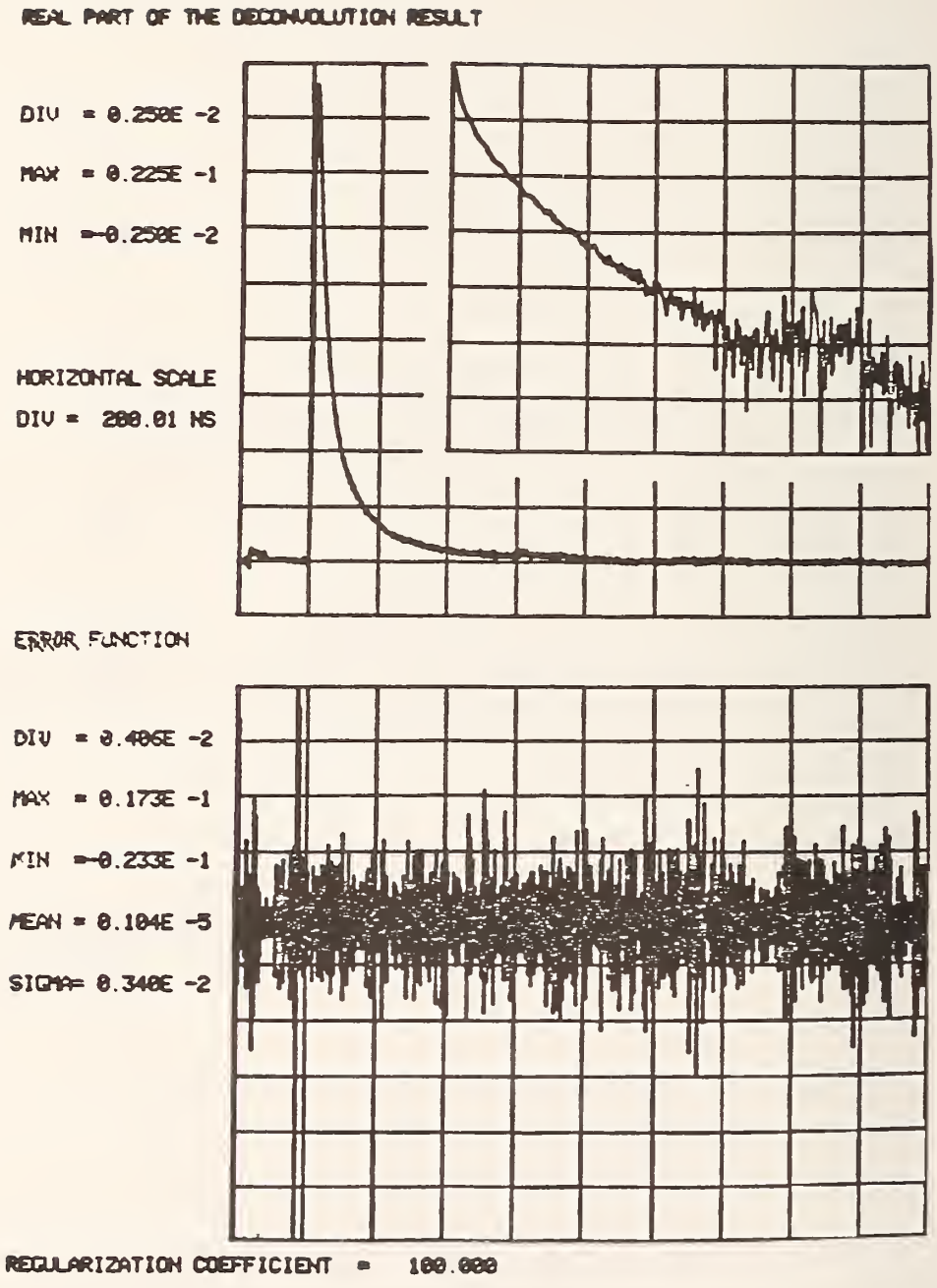


Figure 4.34 SNR-x 75 dB, SNR-y 46 dB. Deconvolution with moderate regularization, $\gamma = 100$. $|D_1(n)|$, $d_2(k)$, and $e(k)$. Note filtering of $H_1(n)$ to obtain $D_1(n)$. $|D_1(n)|$dB is shown in the upper right hand portion of the upper figure, vertical -10 dB/Div. commencing at 0.0 dB, horizontal 25 MHz/Div. commencing at 0.0 MHz.

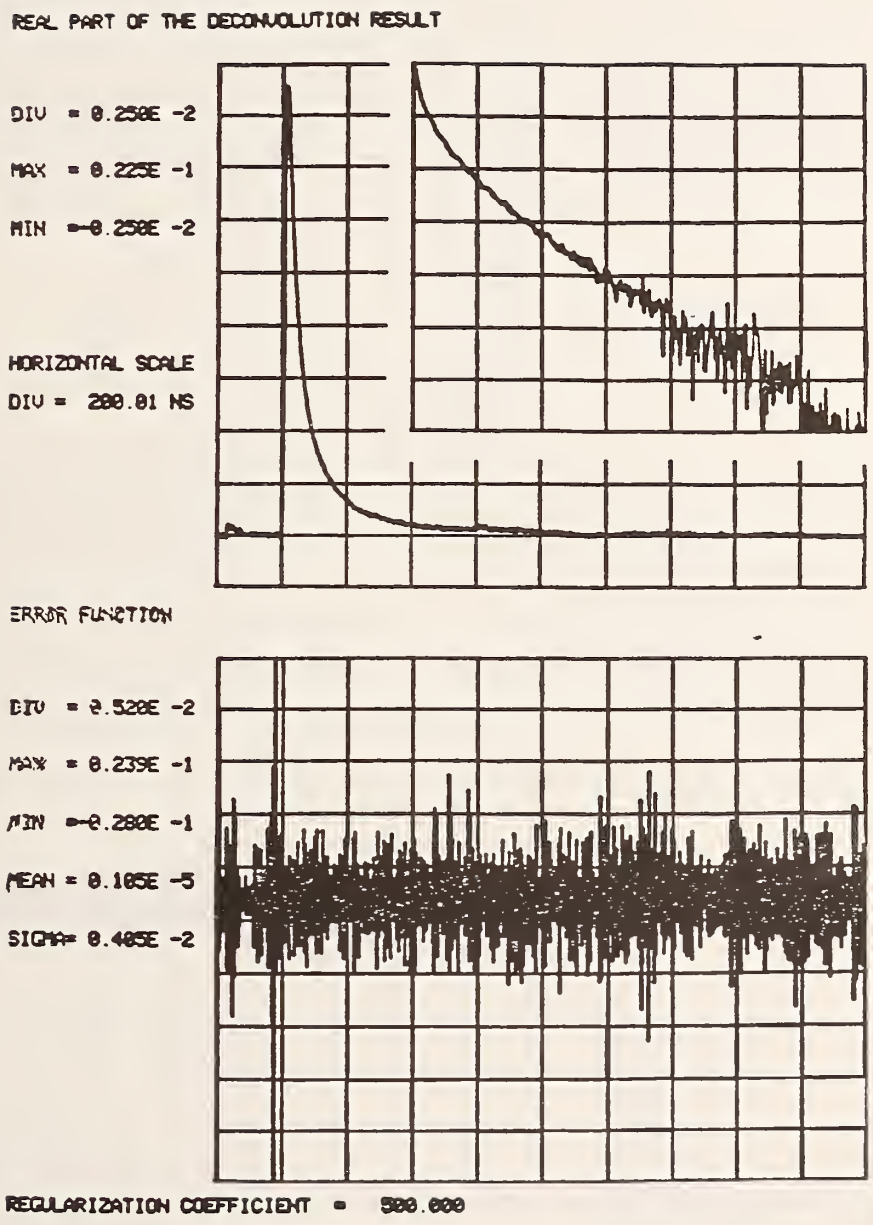


Figure 4.35 SNR-x 75 dB, SNR-y 46 dB. Deconvolution with strong regularization, $\gamma = 500$. $|D_1(n)|$, $d_2(k)$, $e(k)$. Note filtering of $H_1(n)$ to obtain $|D_1(n)|$. $|D_1(n)|$dB is shown in the upper right hand portion of the upper figure, vertical -10 dB/Div. commencing at 0.0 dB, horizontal 25 MHz/Div. commencing at 0.0 MHz.

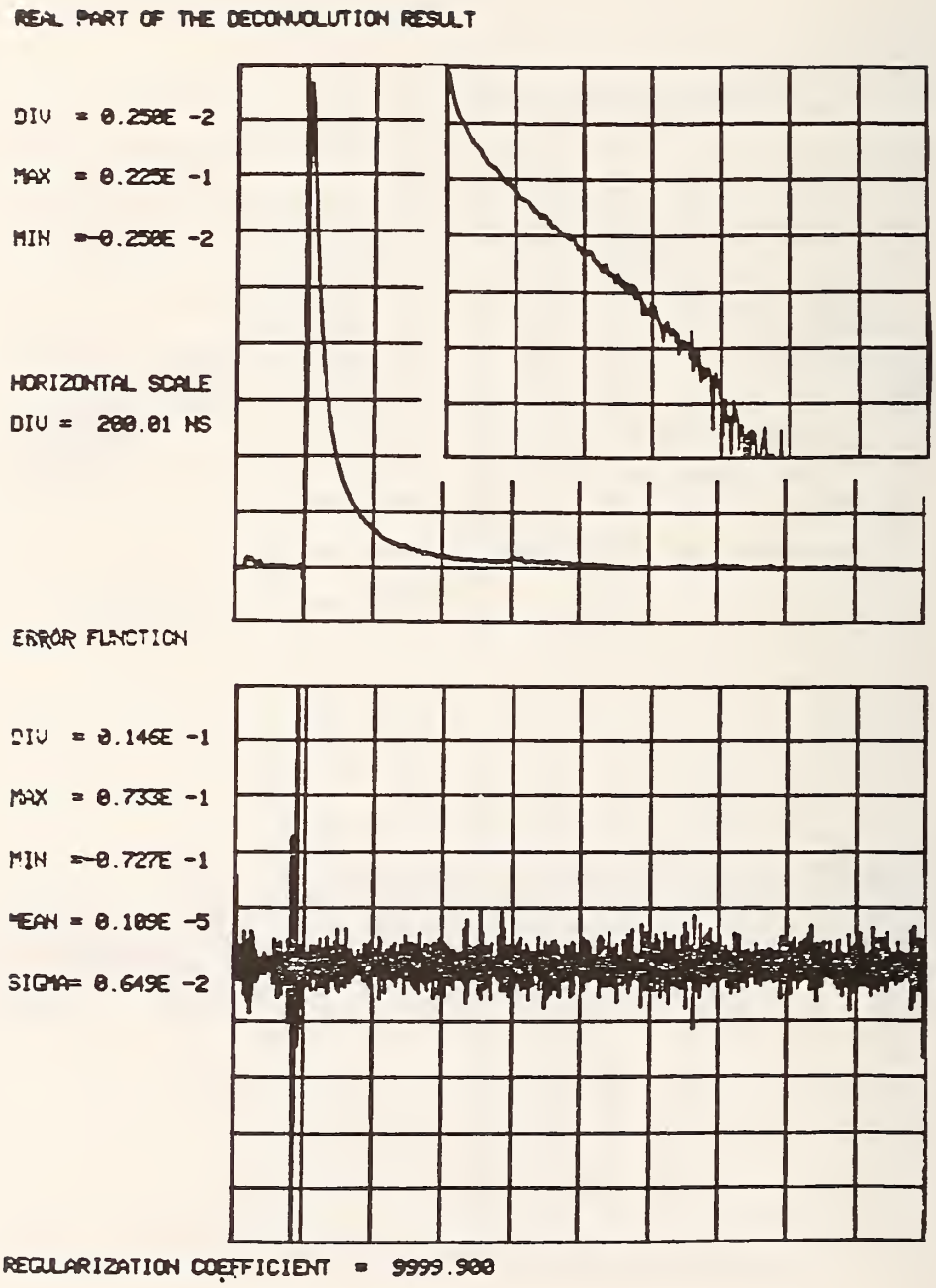


Figure 4.36. SNR-x 75 dB, SNR-y 46 dB. Deconvolution with very strong regularization, γ= 10,000. $|D_1(n)|$, $d_2(k)$, $e(k)$. Note filtering of $H_1(n)$ to obtain $D_1(n)$. $|D_1(n)|$dB is shown in the upper right hand portion of the upper figure, vertical, -10 dB/div. commencing at 0.0 dB, horizontal, 25 MHz/div. commencing at 0.0 MHz.

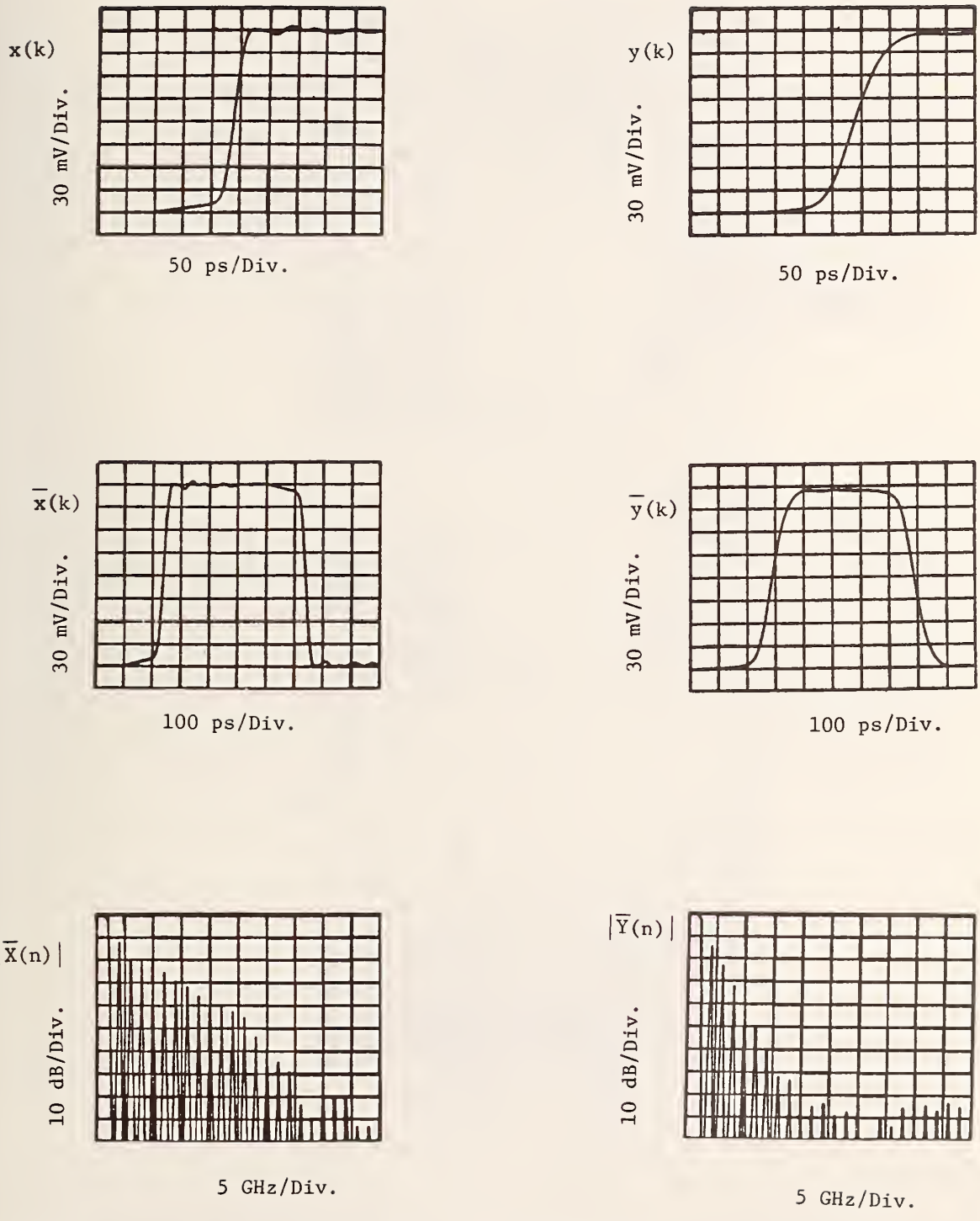


Figure 4.37 The step-like insertion waveforms x(k) and y(k), and their duration limited counterparts $\bar{x}(k)$ and $\bar{y}(k)$, respectively. $|X_1(n)|$ and $|\bar{Y}_1(n)|$ are the magnitudes of the DFT's of $\bar{x}(k)$ and $\bar{y}(k)$, respectively.

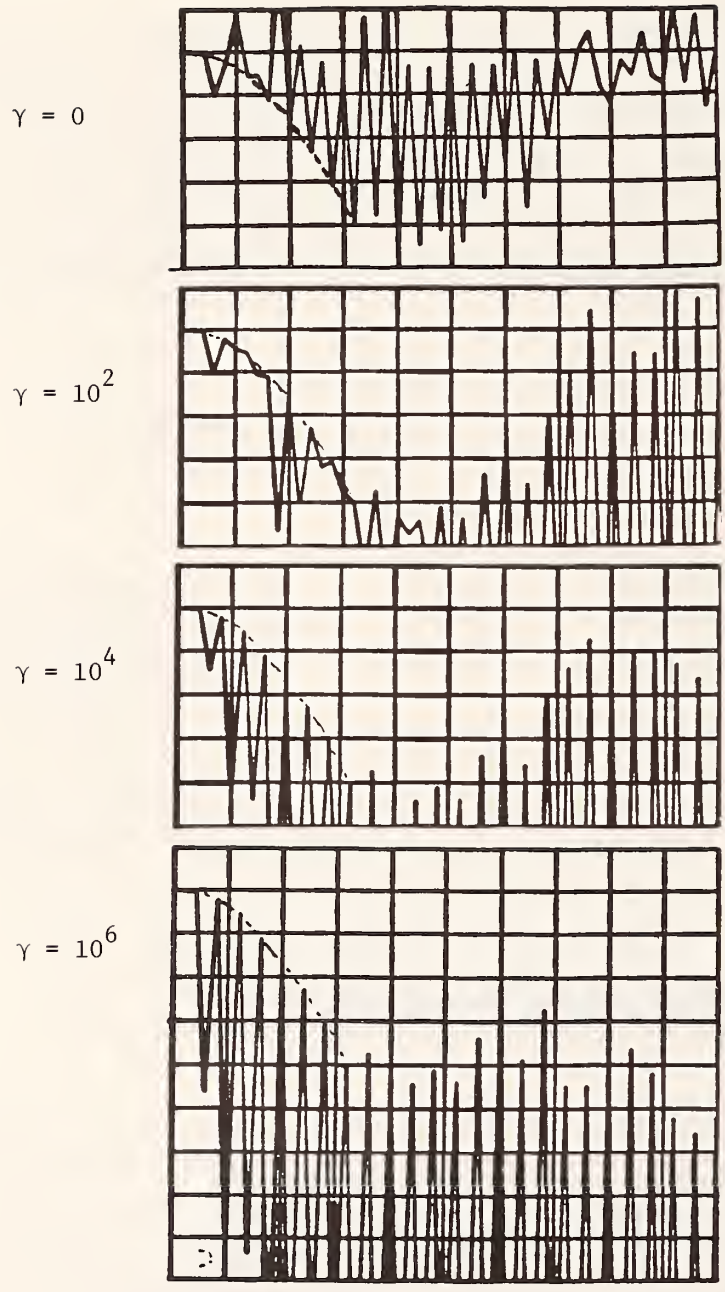


Figure 4.38 The effect of filtering, $|\overline{D}_1(n)| = |\overline{Y}_1(n)/\overline{X}_1(n)|$ R(n). $\gamma = 0$ no filtering, $\gamma = 10^6$ very strong filtering. Vertical scale 10 dB/Div., horizontal scale, 5 GHz/Div. The regularization filter, R(n) is real. When $\gamma = 0$, $|\overline{D}_1(n)| = |\overline{H}_1(n)|$.

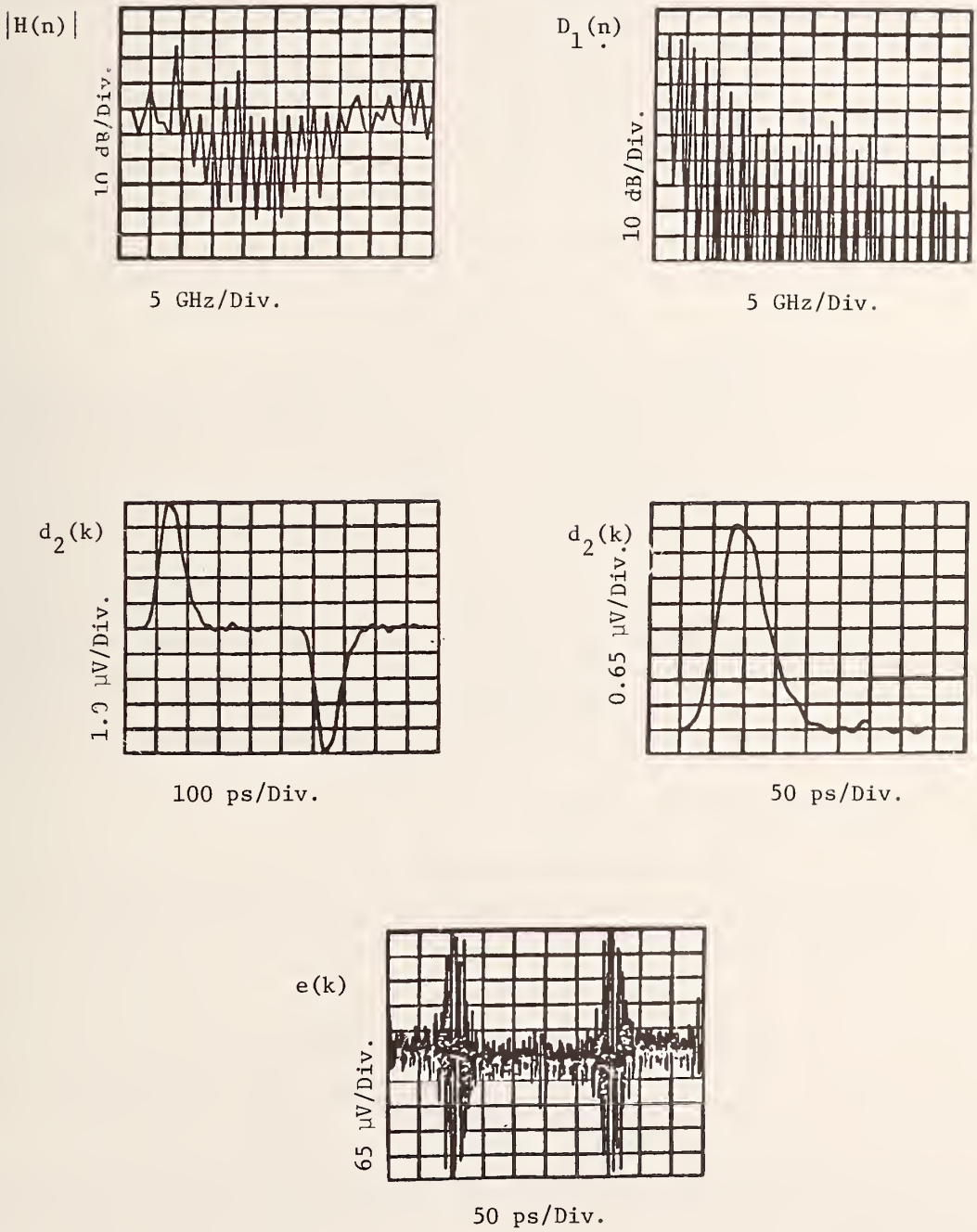


Figure 4.39 The waveforms and spectral magnitudes for very strong regularization, $\gamma = 10^6$.

5. TIME DOMAIN DECONVOLUTION

5.1 Introduction

Time domain deconvolution methods operate directly on the time domain data rather than on transform domain data. There are certain advantages in time domain operations. Because the physical manifestation of the signal (pulse, transient, etc.) occurs with the evolution of time, the time domain is the natural or physical domain of the signal. Direct operations on a physical signal involve operations on real quantities. The causal and stable nature of a signal is readily recognized in the time domain and can be directly preserved under approximations to, and operations on the signal. Also, convolution and deconvolution processes implemented in the time domain only employ data corresponding to times such that $k \leq k_1$ for the computation of $y(k_1)$ and $h(k_1)$, respectively, in

$$y(k) = x(k) * h(k) \tag{5.1-1}$$

$$h(k) = y(k)(1/*) x(k) \tag{5.1-2}$$

The presence of noise will ultimately yield undesirable results (probably unstable) in a time domain deconvolution calculation; this fact is illustrated later in the discussion. However, filtering or smoothing of the time domain data by direct operations on the data can produce stable deconvolution results, i.e., estimates of h(k). The apparent promise of determining h(k) from the time domain quantities x(k) and y(k) by direct operations in the time domain is a sufficient stimulus to try and possibly not fail.

The time domain methods presented here are not as easy to apply in comparison to the frequency domain methods presented in Chapter 4, nor are the results as smooth. In the authors' view, the time domain methods are still in the experimental stages of development while the frequency domain methods can be used with a significantly greater degree of confidence on the part of the user. However, the promise of the time domain method remains to spur further development studies.

In the remaining sections of this chapter the following topics will be discussed:

1. Classical time domain deconvolution
2. Iterative time domain deconvolution
3. Regularization or filtering in the time domain
4. Experimental results of classical deconvolution with regularization applied to the RG 58 C/U transmission line problem of Chapter 4.

5.2 Discrete Classical Time Domain Deconvolution

5.2.1 Convolution

The discrete classical time domain deconvolution method is an algebraic process which is the inverse of the convolution algebraic process [26]. For the moment consider the multiplication of the two polynomials $f_1(x)$ and $f_2(x)$.

$$f_1(u) = a_o + a_1 u + a_2 u^2 + \cdots + a_{N-1} u^{N-1} \tag{5.2.1-1}$$

$$f_2(u) = b_o + b_1 u + b_2 u^2 + \cdots + b_{N-1} u^{N-1} \tag{5.2.1-2}$$

The expression for the coefficients of the product

$$f_3(x) = f_1(x)\, f_2(x) \tag{5.2.1-3}$$

is identical to the convolution of the N term sequences

$$h(k) = a_o, a_1, a_2, \cdots, a_{N-1} \tag{5.2.1-4}$$

$$x(k) = b_o, b_1, b_2, \cdots, b_{N-1} \tag{5.2.1-5}$$

and is given by

$$y(k) = h(k) * x(k) = \text{Coefficients of the product, } f_1(u)f_2(u)$$

$$= \sum_{i=0}^{k-1} a_i b_{k-i} \tag{5.2.1-6}$$

In the notation used in the earlier chapters, the convolution result for y(k) is given by

$$y(k) = \sum_{i=0}^{k-1} h(i)x(k-i) \tag{5.2.1-7}$$

Because convolution corresponds to multiplication of two polynomials, deconvolution corresponds to division of two polynomials. These views are clearly seen in the Z-transform representation which is discussed in section 5.2.2.

For the causal signals h(k) and x(k), both are equal to zero for $k < 0$. Expansion of (5.2-7) for causal signals leads to a result in which the $k = 0$ term is zero; hence the resultant sequence for y(k) commences with $k = 1$, figure 5-1a. Accordingly, deleting the initial zero values h(0) and x(0), (5.2-7) may be written as

$$y(k) = \sum_{i=1}^{k} h(i)\, x(k+1-i) \tag{5.2.1-8}$$

which yields the set of equations for the y sequence values.

$$y(1) = h(1)\, x(1) \tag{5.2.1-9}$$

$$y(2) = h(1)\, x(1) + h(2)\, x(1) \tag{5.2.1-10}$$

$$y(3) = h(1)\, x(3) + h(2)\, x(2) + h(3)\, x(1) \tag{5.2.1-11}$$

and so on. Notice that $y(k_1)$ is computed from earlier values of h(k) and x(k), i.e., $k \leq k_1$.

5.2.2 Z-transform Representation

Some of the properties of convolution and deconvolution are more easily developed in the Z-transform domain [29]. Before proceeding with the discussion, some of the properties of the Z-transform will be briefly described. Given a pair of consistent (exact) discrete waveforms x(k) and y(k), each one of them can be transformed to the complex z domain by the Z-transformation

$$F(z) = \sum_{k=0}^{\infty} f(k)\ z^{-k} \qquad (5.2.2\text{-}1)$$

that is, f(k) is transformed to F(z); hence

$$X(z) = \sum_{k=0}^{\infty} x(k)\ z^{-k} \qquad (5.2.2\text{-}2)$$

$$Y(z) = \sum_{k=0}^{\infty} y(k)\ z^{-k} \qquad (5.2.2\text{-}3)$$

Furthermore, the discrete convolution

$$y(k) = x(k) * h(k)$$

transforms to

$$Y(z) = X(z)\ H(z) \qquad (5.2.2\text{-}4)$$

The ratio of the two polynomials for Y(z) and X(z) yields H(z),

$$H(z) = \frac{\sum_{k=0}^{\infty} y(k)\ z^{-k}}{\sum_{k=0}^{\infty} x(k)\ z^{-k}} \qquad (5.2.2\text{-}5)$$

which is another polynomial in z,

$$H(z) = \sum_{k=0}^{\infty} h(k)\ z^{-k} \qquad (5.2.2\text{-}6)$$

Consequently, as it was pointed out in Section 5.2.1 the coefficients of the convolution and deconvolution sequences are generated according to the rule for polynomial products and quotients, respectively.

5.2.3 Deconvolution

A set of equations for generating the h sequence can be obtained from eqs. (5.2.1-9) through (5.2.1-11) (and so on) by first rearranging them to

$$h(1)\ x(1) = y(1)$$
$$h(2)\ x(1) = y(2) - h(1)\ x(1)$$
$$h(3)\ x(1) = y(3) - h(1)\ x(3) - h(2)\ x(2)$$

and so on. In summation form,

$$h(k)\ x(1) = y(k) - \sum_{i=1}^{k-1} h(i)\ x(k + 1 - i) \qquad (5.2.3\text{-}1)$$

Dividing x(k) and y(k) by x(1) gives

$$\bar{x}(k) = \frac{x(k)}{x(1)} \qquad (5.2.3\text{-}2)$$

$$\bar{y}(k) = \frac{y(k)}{x(1)} \tag{5.2.3-3}$$

where

$$x(1) \neq 0 \tag{5.2.3-4}$$

$$\bar{x}(1) = 1 \tag{5.2.3-5}$$

Consequently, h(k) can be written as

$$h(k) = \bar{y}(k) - \sum_{i=1}^{k-1} h(i)\ \bar{x}(k+1-i) \tag{5.2.3-6}$$

Briefly, (5.2.3-6) generates the following sequence for h(k);

$$h(1) = \bar{y}(1) \tag{5.2.3-7}$$

$$h(2) = \bar{y}(2) - h(1)\ \bar{x}(2) \tag{5.2.3-8}$$

$$h(3) = \bar{y}(3) - h(1)\ \bar{x}(3) - h(2)\ \bar{x}(2) \tag{5.2.3-9}$$

$$h(4) = \bar{y}(4) - h(1)\ \bar{x}(4) - h(2)\ \bar{x}(3) - h(3)\ \bar{x}(2) \tag{5.2.3-10}$$

$$h(5) = \bar{y}(5) - h(1)\ \bar{x}(5) - h(2)\ \bar{x}(4) - h(3)\ \bar{x}(3) - h(4)\ \bar{x}(2) \tag{5.2.3-11}$$

.
.
.

Inspection of the values of h(k) on a line to line basis shows that the earlier values of the impulse response sequence, e.g., h(1), h(2), $\cdots$, $h(k_1-1)$, are used to calculate the specific value $h(k_1)$. Because the h-values can be expressed in terms of the input and output sequence values, $\bar{x}$ and $\bar{y}$, the h values are actually being computed from the $\bar{x}$ and $\bar{y}$ values. Note that h(1) can be eliminated from the expression for h(2) by using the first relation for h(1) and h(2).

$$h(2) = \bar{y}(2) - [\bar{y}(1)]\ \bar{x}(2)$$

Similarly, for h(3) there results,

$$h(3) = \bar{y}(3) - \bar{y}(1)\ \bar{x}(3) - [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(2)$$

continuing the process obtains the partial set of equations for h(k),

$$h(1) = \bar{y}(1) \tag{5.2.3-12}$$

$$h(2) = [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)] \tag{5.2.3-13}$$

$$h(3) = [\bar{y}(3) - \bar{y}(1)\ \bar{x}(3)] - [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(12) \tag{5.2.3-14}$$

$$h(4) = [\bar{y}(4) - \bar{y}(1)\ \bar{x}(4)] - [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(3) - [\bar{y}(3) - \bar{y}(1)\ \bar{x}(3)]\ \bar{x}(2) + [y(2) - y(1)\ x(2)]\ x^2(2) \tag{5.2.3-15}$$

$$\begin{aligned} h(5) = {} & [\bar{y}(5) - \bar{y}(1)\ \bar{x}(5)] - [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(4) \\ & - [\bar{y}(3) - \bar{y}(1)\ \bar{x}(3)[\ \bar{x}(3) + [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(2)\ \bar{x}(3) \\ & - [\bar{y}(4) - \bar{y}(1)\ \bar{x}(4)]\ \bar{x}(2) + [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}(3)\ \bar{x}(2) \\ & + [\bar{y}(3) - \bar{y}(1)\ \bar{x}(3)]\ \bar{x}^2(2) - [\bar{y}(2) - \bar{y}(1)\ \bar{x}(2)]\ \bar{x}^3(2) \end{aligned} \tag{5.2.3-16}$$

and so on. Defining the quantities,

$$A(i) = \bar{y}(i) - \bar{y}(1)\,\bar{x}(i) \tag{5.2.3-17}$$

$$\alpha(1) = 1 \tag{5.2.3-18}$$

$$\alpha(r) = -\sum_{i=1}^{r-1} \alpha(i)\,\bar{x}(r+1-i) \tag{5.2.3-19}$$

enable h(k) to be written as

$$h(1) = 0 \tag{5.2.3-20}$$

$$h(2) = \alpha(1)\,A(2) \tag{5.2.3-21}$$

$$h(3) = \alpha(1)\,A(3) + \alpha(3)\,A(2) \tag{5.2.3-22}$$

$$h(4) = \alpha(1)\,A(4) + \alpha(2)\,A(3) + \alpha(3)\,A(2) \tag{5.2.3-23}$$

$$h(5) = \alpha(1)\,A(5) + \alpha(2)\,A(4) + \alpha(3)\,A(3) + \alpha(4)\,A(2) \tag{5.2.3-24}$$

and so on. Explicitly, in terms of (5.2.3-16) through (5.2.3-19), h(k) is given by

$$h(k) = \bar{y}(1) + \sum_{i=1}^{k-1} \alpha(i)\,A(k+1-i) \tag{5.2.3-25}$$

$$= \bar{y}(1) + \sum_{i=1}^{k-1} \alpha(i)\,[\bar{y}(k+1-i) - \bar{y}(1)\,\bar{x}(k+1-i)] \tag{5.2.3-26}$$

which is the general result for h(k) in terms of $\bar{x}(k)$ and $\bar{y}(k)$.

To gain further insight into this result for h(k), (5.2.3-26), consider the signals shown in figure 5.1. x(k) and h(k) are given signals while y(k) is calculated from x(k) * h(k), figure 5.1a. Subsequently, the deconvolution result, $h_d(k)$, is calculated from y(k)(1/*)x(k). Note that both y(k) and $h_d(k)$ commence at k = 1 while x(k) and h(k) commence at k = 0. From (5.2.3-17) through (5.2.3-19) using the data of figure 5.1, there results

$$A(2) = \bar{y}(2) \tag{5.2.3-27}$$

$$A(3) = \bar{y}(3) \tag{5.2.3-28}$$

$$A(4) = \bar{y}(4) \tag{5.2.2-39}$$

$$A(5) = \bar{y}(5) \tag{5.2.3-30}$$

and so on. Next,

$$\alpha(1) = 1 \tag{5.2.3-31}$$

$$\alpha(2) = \bar{x}(2) \tag{5.2.3-32}$$

$$\alpha(3) = -\bar{x}(3) + \bar{x}^2(2) \tag{5.2.3-33}$$

$$\alpha(4) = \bar{x}(2)\,\bar{x}(3) - [-\bar{x}(3) + \bar{x}^2(2)]\,\bar{x}(2) = 2\bar{x}(2) - \bar{x}^3(2) \tag{5.2.3-34}$$

and so on. Using the above values, calculate the deconvolution of y(k) and x(k) to obtain $h_d(k)$. To insure clarity, the deconvolution result for h(k) is denoted by $h_d(k)$. The terms in $h_d(k)$ will be uniquely related to those of the given sequence, h(k). From (5.2.3-20) through (5.2.3-24),

$$h_d(k) = \sum_{i=1}^{k-1} \alpha(i)A(k+1-i) \tag{5.2.3-35}$$

The first four terms are

$$h_d(1) = 0 \tag{5.2.3-36}$$

$$h_d(2) = \alpha(1)A(2) = \bar{y}(2) \tag{5.2.3-37}$$

$$h_d(3) = \alpha(1)A(3) + \alpha(2)A(2) = \bar{y}(3) - \bar{x}(2)\,\bar{y}(2) \tag{5.2.3-38}$$

$$h_d(4) = \alpha(1)A(4) + \alpha(2)A(3) + \alpha(3)A(2)$$

$$= \bar{y}(4) - \bar{x}(2)\,\bar{y}(3) + [-\bar{x}(3) + \bar{x}^2(2)]\,\bar{y}(2) \tag{5.2.3-39}$$

Upon substituting the values of y(k) from the convolution (5.2.1-8) into (5.2.3-36) through (5.2.3-39) and recalling that $\bar{x}(k) = x(k)/x(1)$ and $\bar{y}(k) = y(k)/x(1)$, the following values for $h_1(k)$ are obtained.

$$h_d(1) = 0 \tag{5.2.3-40}$$

$$h_d(2) = h(1) \tag{5.2.3-41}$$

$$h_d(3) = h(2) \tag{5.2.3-42}$$

$$h_d(4) = h(3) \tag{5.2.3-43}$$

Inspection of these results shows that $h_d(k)$ is equal to h(k-1), figure 5.1b.

Figure 5.2 shows the result of classical deconvolution using two noise free signals, x(k) and y(k). In this case x(k) and h(k) were specified as

$$x(k) = e^{-ck}(1-e^{-ck}); \quad k = 0,1,\cdots,49 \tag{5.2.3-44}$$

$$h(k) = e^{-3ck}(1-e^{ck}); \quad k = 0,1,\cdots,49 \tag{5.2.3-45}$$

which, in turn, specified y(k) as

$$y(k) = x(k) * h(k) \qquad k = 0,1,\cdots,49. \tag{5.2.3-46}$$

The data for y(k) was calculated using (5.2.1-8). The resultant y(k) was used with x(k), (5.2.3-44), in (5.2.3-26) to calculate the deconvolution.

$$h(k) = y(k)(1/*)\ x(k) \quad k = 0,1,\cdots,49 \tag{5.2.3-47}$$

In applying (5.2.3-6) or (5.2.3-26) it should be kept in mind that for k equal to zero, h(k) is always zero (causal system); consequently, the first term, zero, must be added to the sequence for h(k). In the present example, (5.2.3-6) or (5.2.3-26) yields a sequence of 49 terms, and including zero as the first term in the sequence gives a total of 50 terms in the sequence.

5.2.4 The Effect of Errors in Classical Deconvolution

The effect of errors or noise on the deconvolution process can be seen by considering (5.2.3-6) and its partial expansion (5.2.3-7) through (5.2.3-11) in the presence of errors. For the moment, recall the deconvolution example given above in Section 5.2.3. The noise free signals x(k) and h(k), (5.2.3-44) and (5.2.3-45), respectively, were used to calculate a noise free y(k), (5.2.3-46); this y(k) was consistent with x(k) and h(k) through the convolution equation (5.2.1-8). Now imagine that one value of the y(k) sequence is changed to an incorrect value; let y(1) be altered to $\tilde{\bar{y}}(1)$,

$$\tilde{\bar{y}}(1) = \bar{y}(1) + e_y(1) \tag{5.2.4-1}$$

where $e_y(1)$ is the error added to the correct value, $y(1)$. Putting $\bar{y}(1)$ into (5.2.2-7) through (5.2.2-11) yields a partial set of the erroneous $h(k)$'s, $\tilde{h}(k)$,

$$\tilde{h}(1) = h(1) + e_y(1) \tag{5.2.4-2}$$

$$\tilde{h}(2) = h(2) + e_y(1)\,[-\bar{x}(2)] \tag{5.2.4-3}$$

$$\tilde{h}(3) = h(3) + e_y(1)\,[-\bar{x}(3) + \bar{x}^2(2)] \tag{5.2.4-4}$$

$$\tilde{h}(4) = h(4) + e_y(1)\,[\bar{x}(4) - 2\bar{x}(2)\,\bar{x}(3) - \bar{x}^3(2)] \tag{5.2.4-5}$$

$$\tilde{h}(5) = h(5) + e_y(1)\,[\bar{x}(5) + \bar{x}^2(3) + \bar{x}^2(2) + \bar{x}^4(2)] \tag{5.2.4-6}$$

and so on. The effect of the error, $e_y(1)$, is increased with increasing k; the error in $h(k)$ is simply

$$e_h(k) = \tilde{h}(k) - h(k) \tag{5.2.4-7}$$

Here, the set of errors, $e_h(k)$, in $h(k)$ is due to the single error in $y(k)$, $e_y(1)$. In practice, the data values for $x(k)$ and $y(k)$ will each be corrupted by noise; consequently, the errors in $h(k)$ will be dependent upon the sets of error for both $x(k)$ and $y(k)$. Consequently, one can imagine that $\tilde{h}(k)$ will be very sensitive to the errors and most probably will be an unstable sequence.

The analysis can be extended to a set of errors in both $\bar{y}(k)$ and $\bar{x}(k)$. First, let $\bar{x}(k)$ be error free. Let the erroneous values of $\bar{y}(k)$ be given by $\tilde{\bar{y}}(k)$,

$$\tilde{\bar{y}}(k) = \bar{y}(k) + e_y(k) \tag{5.2.4-8}$$

Putting $\tilde{\bar{y}}(k)$ into the general result for $h[\bar{x}(k), \bar{y}(k)]$, (5.2.3-26) obtains

$$\tilde{h}(k) = h(k) + e_y(1) + \sum_{i=1}^{k-1} \alpha(i)\,[e_y(k+1-i) - e_y(1)\,\bar{x}(k+1-i)], \tag{5.2.4-9}$$

The error in $h(k)$ is $e_h(k)$ and is given by

$$e_h(k) = \tilde{h}(k) - h(k) \tag{5.2.4-10}$$

$$= e_y(1) + \sum_{i=1}^{k-1} \alpha(i)\,[e_y(k+1-i) - e_y(1)\,\bar{x}(k+1-i)] \tag{5.2.4-11}$$

Next, include the errors in $\bar{x}(k)$ by replacing $\bar{x}(k)$ and $\alpha(k)$ by

$$\tilde{\bar{x}}(k) = \bar{x}(k) + e_x(k) \tag{5.2.4-12}$$

$$\tilde{\alpha}(k) = \alpha(k) + e_\alpha(k) \tag{5.2.4-13}$$

The final result for the erroneous h values is given by

$$\tilde{h}(k) = \bar{y}(1) + \sum_{i=1}^{k-1} \alpha(i)\,[\bar{y}(k+1-i) - \bar{y}(1)\,\bar{x}(k+1-i)] - \sum_{i=1}^{k-1} \alpha(i)\,\bar{y}(1)\,e_x(k+1-i)$$

$$+ \sum_{i=1}^{k-1} e_\alpha(i)\{\bar{y}(k+1-i) - \bar{y}(1)\,[\bar{x}(k+1-i) + e_x(k+1-i)]\} + \text{(con't, next page)}$$

$$+ e_y(k) + \sum_{i=1}^{k-1} [\alpha(i) + e_q(i)]\{e_y(k+1-i) - e_y(1)[\bar{x}(k+1-i) + e_x(k+1-i)]\} \qquad (5.2.4\text{-}14)$$

The error in h(k) is then

$$e_h(k) = \tilde{h}(k) - h(k) \qquad (5.2.4\text{-}15)$$

$$= e_y(k) - \sum_{i=1}^{k-1} \alpha(i)\, \bar{y}(1)\, e_x(k+1-i) + \sum_{i=1}^{k-1} e_\alpha(i)\, \{\bar{y}(k+1=i) - \bar{y}(1)[\bar{x}(k+1-i) + e_x(1+1-i)]\}$$

$$+ \sum_{i=1}^{K-1} [\alpha(i) + e_q(i)]\, \{e_y(k+1-i) - e_y(1)[\bar{x}(k+1-i) + e_x(k+1-i)]\} \qquad (5.2.4\text{-}16)$$

To illustrate the effect of noise on the deconvolution process, noise was added to the noise-free signals (5.2.3-44) and (5.2.3-46) so that each signal possessed a 40 dB signal to noise ratio (as defined in Section 4.1); the result is $\tilde{h}(k)$, (5.2.4-14), and is shown in figure 5.3. It is clear that the errors (noise) lead to an unstable result which bears no resemblance to h(k), figure 5.2. Similar results are also obtained for the iterative deconvolution method which will be discussed in section 5.3

5.2.5 The Critical Dependence of Classical Deconvolution on the First Point in x(k).

In the practical application of the classical time domain method to the deconvolution problem, the choice of the first term in the input sequence, x(1), is critical. If x(1) is ill-chosen the resultant deconvolved result for h(k) will not converge. To gain an understanding of the origin of this feature, consider the Z-transforms of the two known sequences x(m) and y(n) having their durations denoted as M' and N', respectively. Thus,

$$X(z) = x(1)\, z^{-1} + x(2) z^{-2} + \cdots + x(M')$$

$$= z^{-M}[x(1)\, z^{M} + x(2)\, z^{M-1} + \cdots + x(M+1)]; \quad M'-1 = M \qquad (5.2.5\text{-}1)$$

and

$$Y(z) = y(1)\, z^{-1} + y(2)\, z^{-2} + \cdots + y(N')$$

$$= z^{-N}[y(1)\, z^{N} + y(2)\, z^{N-1} + \cdots + y(N+1)] \quad N'-1 = N \qquad (5.2.5\text{-}2)$$

respectively, where

$$N \geq M \qquad (5.2.5\text{-}3)$$

The Z-transform of the impulse response corresponding to x(k) and y(k) is

$$H(\ell) = Y(n)/X(m) \qquad (5.2.5\text{-}4)$$

When N equals M, H(ℓ) becomes H(n) and is a constant.

$$H(n) = \frac{Y(n)}{X(n)} = \text{constant} = C; \quad N = M \qquad (5.2.5.5)$$

The inverse Z-transform of H(n) is then the impulse sequence C, 0, 0, ··· 0 in which the first term is C and is followed by N-zeroes.

By synthetic division H(z) is given by

$$H(z) = \frac{1}{z^{N-M}}\left\{\frac{y(1)}{x(1)}\, z^{N-M} + \left[y(2) - \frac{y(1)\,x(2)}{x(1)}\right]\frac{z^{N-M-1}}{x(1)}\right.$$

$$\left. + \left\{\left[y(3) - \frac{y(1)x(3)}{x(1)}\right] - \left[y(2) - \frac{y(1)x(2)}{x(1)}\right]\frac{x(2)}{x(1)}\right\}\frac{z^{N-M-2}}{x(1)} + \cdots\right\} \tag{5.2.5-6}$$

Using the variables $\bar{x}(k)$ and $\bar{y}(k)$ which are the x(k) and y(k) normalized to x(1), H(z) can be written as

$$H(z) = \bar{y}(1) + [\bar{y}(2) - \bar{y}_1(1)\bar{x}(2)]z^{-1} + \{\bar{y}(3) - \bar{y}(1)\bar{x}(3) - [\bar{y}(2) - \bar{y}(1)\bar{x}(2)]\bar{x}(2)\}z^{-2} + \cdots \tag{5.2.5-7}$$

Now consider that the point x(1) is corrupted by an error, e_x, and thus changed to $\tilde{\bar{x}}(1)$,

$$\tilde{\bar{x}}(1) = x(1) + e_x \tag{5.2.5-8}$$

while all other values of x(k) and y(k) are correct. The normalized x(m) and y(n) would be given by

$$\tilde{\bar{x}}(m) = 1, \frac{x(2)}{x(1) + e_x}, \cdots, \frac{x(M')}{x(1) + e_x}$$

$$= 1, \bar{x}(2) + e_x(2), \cdots, \bar{x}(M') + e_x(M') \tag{5.2.5-9}$$

$$\tilde{\bar{y}}(n) = \frac{y(1)}{x(1) + e_x}, \frac{y(2)}{x(1) + e_x}, \cdots, \frac{y(N')}{x(1) + e_x}$$

$$= y(1) + e_y(1), y(2) + e_y(2), \cdots, y(N') + e_y(N') \tag{5.2.5-10}$$

respectively. Then the only correct term in $\tilde{\bar{x}}(m)$ would be $\tilde{\bar{x}}(1)$ which equals $\bar{x}(1) = 1$. The y(n) sequence would be scaled to the erroneous value $x(1) + e_x$. Consequently, the ratio for H(z) would also be erroneous,

$$\tilde{H}(z) = \frac{\tilde{\bar{Y}}(z)}{\tilde{\bar{X}}(z)} = \left\{\tilde{\bar{y}}(1) + [\tilde{\bar{y}}(2) - \tilde{\bar{y}}(1)\,\tilde{\bar{x}}(2)]\,\tilde{\bar{z}}^{-1} + \left\{\tilde{\bar{y}}(3) - \tilde{\bar{y}}(1)\,x(3)\right.\right.$$

$$\left.\left. - [\tilde{\bar{y}}(2) - \tilde{\bar{y}}(1)\,\tilde{\bar{x}}(2)]\,\tilde{\bar{x}}(2)\right\}\tilde{\bar{z}}^{-2} + \cdots\right\} \tag{5.2.5-11}$$

The inverse Z-transform of $\tilde{H}(z)$ gives the erroneous sequence $\tilde{h}(k)$, (5.2.4-14), and an error in h(k) of $e_h(k)$, (5.2.4-16).

Consequently, it has been shown that the error terms, i.e.,

$e_x(k)$: $e_x(2), e_x(3), \cdots$

$e_\alpha(k)$: $e_\alpha(2), e_\alpha(3), \cdots$

$e_y(k)$: $e_y(1), e_y(2), \cdots$

$$e_h(k) : e_h(2), e_h(3), \cdots$$

have been generated by the single error in x(1) through the normalization processes (5.2.5-9) and (5.2.5-10). Thus the selection of x(1) is critical to the deconvolution process.

5.3 Discrete Iterative Time Domain Deconvolution

5.3.1 Iterative Method

By using an iteration process it is also possible to solve the deconvolution equation by an iterative convolution process [4,5,27,28] proposed by P. H. van Cittert. [28]. The principle of this method is to form successive approximations, $h_i(k)$, to the unknown system impulse response, h(k), using the convolution equation,

$$y(k) = h(k) * x(k) \tag{5.3.1-1}$$

For physically realizable signals the duration of x(k) is less than y(k). As a first approximation to h(k), $h_1(k)$, use y(k),

$$h_1(k) = y(k) \tag{5.3.1-2}$$

Using $h_1(k)$ in the convolution equation (5.3.1-1) gives the result $y_1(k)$.

$$y_1(k) = h_1(k) * x(k) \tag{5.3.1-3}$$

$y_1(k)$ differs from the true result y(k) by the error, $[\Delta y(k)]_1$

$$[\Delta y(k)]_1 = y(k) - y_1(k) \tag{5.3.1-4}$$

which corresponds to an error in h(k) of

$$[\Delta h(k)]_1 = [\Delta y(k)]_1 \ (1/*) \ x(k) \tag{5.3.1-5}$$

correcting $h_1(k)$ by adding $[\Delta h(k)]_1$ yields a second approximation to h(k), $h_2(k)$.

$$h_2(k) = h_1(k) + [\Delta h(k)]_1 \tag{5.3.1-6}$$

$$= h_1(k) + [y(k) - h_1(k) * x(k)] \tag{5.3.1-7}$$

Consequently, the i-th approximation is related to (i-1)th approximation by

$$h_i(k) = h_{i-1}(k) + [y(k) - h_{i-1}(k) * x(k)] \tag{5.3.1-8}$$

When

$$h_i(k) * x(k) = y(k), \tag{5.3.1-9}$$

then

$$h_i(k) = h_{i+1}(k) = h(k) \tag{5.3.1-10}$$

Here, the y(k) and x(k) are considered to be noise free. Of course, they are not in practice, and the noise does limit the direct application of the iterative method.

Continuing the discussion on the iterative method, apply the Z-transformation to (5.3.1-8). The result is

$$Y(z) - H_i(z)\, X(3) + H_i(z) = H_{i+1}(z) \tag{5.3.1-10}$$

The form of (5.3.2-7) parallels that of (5.3.1-8) with the convolution operation replaced by a product. The iteration process starts with the first estimate for H(z) being chosen as Y(z),

$$H_1(z) = Y(z) \tag{5.3.1-11}$$

Putting $H_1(z)$ into (5.3.2-4) yields for $H_2(z)$

$$H_2(z) = Y(z)[2-X(z)] \tag{5.3.1-12}$$

Repeating the process for the next values yields

$$H_3(z) = Y(z) + H_2(z)[1 - X(z)] = Y(z)\,\{1 + [2 - X(z)][1 - x(z)]\} \tag{5.3.1-13}$$

$$H_4(z) = Y(z)\,\{1 + [1-X(z)] + [2-X(z)][1-Y(z)]^2\} \tag{5.3.1-14}$$

$$H_5(z) = Y(z)\,\{1 + [1-X(z)] + [1-X(z)]^2 + [2-X(z)][1-X(z)]^3\} \tag{5.3.1-15}$$

$$H_6(z) = Y(z)\,\{1 + [1-X(z)] + [1-X(z)]^2 + [1-X(z)]^3 + [2-X(z)][1-X(z)]^4\} \tag{5.3.1-16}$$

and so on. The result for $H_i(z)$ is a power series of the form

$$H_i(z) = Y(z)\left[i + B_1\, X(z) + B_2\, X(z)^2 + \cdots + B_{i-1}(z)^{i-1}\right] \tag{5.3.1-17}$$

whose inverse Z-transform is

$$h_i(k) = iy(k) + B_1 y(k) * X(k) + B_2 y(k) * X(k) * X(k) + \cdots + B_{i-1}\, y(k) * x(k) * x(k) * \cdots x(k) \tag{5.3.1-18}$$

Hence, it is seen that the direct discrete iterative method requires a large number of convolution operations which in turn will require a large amount of computation time.

Using the x(k) and y(k) of (5.2.3-44, 5.2.3-46), the direct discrete iterative deconvolution method was employed to calculate h(k); the results for 500 and 1000 iterations are shown in figures 5.4 and 5.5 where it is seen that the latter result is identical to the result obtained by the direct discrete classical deconvolution method. However, the discrete iterative method required 45 minutes of computation time as compared to 10 seconds for the classical method.

5.3.2 The Effect of Errors in Iterative Deconvolution

If y(k) contains an error, say for k = 0, then the erroneous output sequence $\tilde{y}(k)$ is

$$\tilde{y}(k) = e_y(0) + y(0), y(1), y(2), \cdots \tag{5.3.2-1}$$

where $e_y(o)$ is the error in y(0). The Z-transform of $\tilde{y}(k)$ is

$$\tilde{Y}(z) = e_y(0) + y(0) + y(1)z^{-1} + y(2)\ z^{-2} + \cdots \tag{5.3.2-2}$$

$$= e_y(0) + \sum_{k=0}^{\infty} y(k)\ z^{-k} \tag{5.3.2-3}$$

$$= e_y(0) + Y(z) \tag{5.3.2-4}$$

Using the erroneous output $\tilde{Y}(z)$ in (5.3.1-17) yields the erroneous i-th approximation to the deconvolution result

$$\tilde{H}_i(z) = [e_y(0) + Y(z)]\ [i + B_1X(z) + B_2X^2(z) + \cdots + B_{i-1}X^{i-1}(z)] \tag{5.3.2-5}$$

$$= +\ H_i(z) + e_y(0)\ F_i[X(z)] \tag{5.3.2-6}$$

where the function of X(z) is defined by

$$F_i[X(z)] = [i + B_1X(z) + B_2X^2(z) + \cdots + B_{i-1}X^{i-1}(z)] \tag{5.3.2-7}$$

Consequently, the effect of the error $e_y(0)$ increases with increasing k. This result is similar to the error result for the classical deconvolution technique with a single error (5.2.4-6).

For a set of errors, $e_y(k)$, in y(k), i.e., for an erroneous output sequence of the form

$$\tilde{y}(k) = e_y(k) + y(k) \tag{5.3.2-8}$$

the Z-transform of the sequence is

$$\tilde{Y}(z) = E_y(z) + Y(z) \tag{5.3.2-9}$$

Using (5.3.2-9) in (5.3.1-17) yields the erroneous i-th approximation to the deconvolution result

$$\tilde{H}_i(z) = [E_y(z) + Y(z)]\ F_i[X(z)] \tag{5.3.2-10}$$

$$= H_i(z) + E_y(z)\ F_i[X(z)] \tag{5.3.2-11}$$

whose inverse Z-transform yields the erroneous time domain sequence, $\tilde{h}_i(k)$,

$$\tilde{h}_i(k) = h_i(k) + e_y(k) * f_i(k) \tag{5.3.2-12}$$

where $f_i(k)$ is the inverse Z-transform of $F_i[X(z)]$. $\tilde{h}_i(k)$ is the y-noise or error corrupted result for the noise free i-th iteration approximation, $h_i(k)$, to the system impulse response, h(k).

If x-noise or errors, $e_x(k)$, are also present they transform to $E_x(z)$; and then X(z) becomes $\tilde{X}(z)$,

$$\tilde{X}(z) = X(z) + E_x(z) \tag{5.3.2-13}$$

and the corrupted i-th iteration,is given by

$$\tilde{H}_i(z) = H_i(z) + E_y(z)\, G_i[X(z) + E_x(z)] \tag{5.3.2-14}$$

$G_i[X(z) + E_x(z)]$ is shown in expanded form in (5.3.2-16). The error in $H_i(z)$ is then

$$E_{hi}(z) = \tilde{H}_i(z) - H_i(z)$$

$$= E_y(z) G_i[X(z)+E_x(z)] \tag{5.3.2-15}$$

$$= E_y(z) \left\{ i + B_1[X(z)+E_x(z)] + B_2[X(z)+E_x(z)]^2 + \cdots + B_{i-1}[X(z)+E_x(z)]^{i-1} \right\} \tag{5.3.2-16}$$

In the time domain, the i-th corrupted iteration, $\tilde{h}_i(k)$, for h(k) is given by the inverse Z-transform of (5.3.2-14).

$$\tilde{h}_i(k) = h_i(k) + E_y(k) * g_i[x(k), e_x(k)] \tag{5.3.2-17}$$

where $g_i(k)$ is the inverse Z-transform of $G_i[X(z) + E_x(z)]$.

Figure 5.5 shows the result for $\tilde{h}_i(k)$ with i=750 (iterations) using noise corrupted values of (5.2.3-44) and (5.2.3-46) for $\tilde{x}(k)$ and $\tilde{y}(k)$, respectively,

$$\tilde{x}(k) = x(k) + e_x(k) \tag{5.3.2-18}$$

$$\tilde{y}(k) = y(k) + e_y(k) \tag{5.3.2-19}$$

where $e_x(k)$ and $e_y(k)$ were such that the SNR of $\tilde{x}(k)$ and SNR of $\tilde{y}(k)$ were both 40 dB. Note that the result is unstable and bears no resemblance to the system response h(k), (5.2.3-45). Figure 5.5 is very similar to figure 5.2, the result for noise corrupted classical deconvolution under the same x and y signal to noise ratios (40 dB).

5.4 Time Domain Regularization

5.4.1 Averaging Filter

In sections 5.2 and 5.3 two direct methods of deconvolution were presented. By direct it is meant that the mathematical method is directly applied to the data x(k) and y(k). As we demonstrated in sections 5.2 and 5.3 the presence of noise or errors in the data destroys the consistency between x(k) and y(k); that is to say, the computed data $\tilde{x}(k)$ and $\tilde{y}(k)$ are no longer mathematically related to each other through the impulse response of the system, h(k). Thus, the direct application of the mathematical methods to the computed data $\tilde{x}(k)$ and $\tilde{y}(k)$ leads to an erroneous and divergent or unstable result, $\tilde{h}(k)$.

Both the classical and iterative methods applied to noisy signals lead to the same type of incorrect and divergent results.

By filtering the noisy waveforms $\tilde{x}(k)$ and $\tilde{y}(k)$ both in the same prescribed way, the functions $\tilde{x}(k)$ and $\tilde{y}(k)$ can be smoothed or regularized so that when the mathematical deconvolution method is applied to them, the process yields a stable convergent result, d(k), which approximates h(k). The smoothing is accomplished by a regularization filter. Figures 5.7 and 5.8 show the resultant estimates for h(k) obtained by first filtering the corrupted signals $\tilde{x}(k)$ and $\tilde{y}(k)$ which were used in earlier examples. Figure 5.7 was obtained by using the filtered signals in the discrete classical deconvolution method, (5.2.3-6) or (5.2.3-26). Simarily, figure 5.8 was produced by the regularized discrete iterative deconvolution method, i.e., filtered $\tilde{x}(k)$ and $\tilde{y}(k)$ in equation (5.3.1-8). Comparison of figures 5.7 and 5.8 shows that the two methods yield very similar results, but again, the computation time for the iterative method was 45 minutes compared to 10 seconds for the classical deconvolution method. The notation d(k) denotes an estimate of the true impulse response h(k) which was estimated from the filtered or regularized data $\tilde{x}(k)$ and $\tilde{y}(k)$.

The regularizzation filter applied to $\tilde{x}(k)$ and $\tilde{y}(k)$ was a simple one, a four-point averaging filter: the value of, say $\tilde{x}(k)$, at the point k was set equal to the average or mean value of the values of $\tilde{x}(k)$, $\tilde{x}(k+1)$, $\tilde{x}(k+2)$, and $\tilde{x}(k+3)$; i.e., $\tilde{x}(k)$ was set equal to $\overline{\tilde{x}(k)}$.

$$\overline{\tilde{x}(k)} = \frac{1}{N} \sum_{j=k}^{N} \tilde{x}(j) \qquad (5.4.1\text{-}1)$$

where N = k + 3. Without the filter (5.4.1-1), the S/N ratios for $\tilde{x}$ and $\tilde{y}$ of 40 dB led to failure of the direct deconvolution methods. With the filter (5.4.1-1), the S/N ratios of 20 dB led to failure; consequently, it could be said that there was a 20 dB enhancement of the deconvolution process.

A major task in time domain deconvolution is the synthesis of a suitable regularization filter. The example described above employed the simple averaging filter (5.4.1-1); however, to achieve improved estimates of the impulse response, d(k), more sophisticated regularization filters need to be synthesized. In the next section a more sophisticated and flexible filter is described.

5.4.2 Composite Filtering

5.4.2.1 Introduction

In the last section an averaging filter was briefly discussed and applied to two deconvolution problems. The averaging filter is equivalent to a system having a rectangular impulse response, $h_a(k)$,

$$h_a(k) = \begin{cases} 1, & 0 \leq k \leq k_1 \\ 0, & k_1 < k < N-1 \end{cases} \qquad (5.4.2.1\text{-}1)$$

In other words, the impulse response $h_a(k)$ consists of N discrete values, $k_1 + 1$ of which are equal to unity, while the remaining values are zero.

If a given waveform sequence $\tilde{x}(k)$ is convolved with $h_a(k)$ using (5.2.1-7) or (5.2.1-8), the result is the same as that produced by the averaging operation, (5.4.1-1).

Presented below is a filter consisting of three operations which are (1) averaging, (2) least-square first-order polynomial fitting, and (3) Gaussian weighting or smoothing. Each one of these operations is represented by an adjustable parameter. Consequently, the overall composite filter possesses three

parameters, L, S, and A corresponding to averaging, fitting, and weighting, respectively.

5.4.2.2 Filter Description

The main characteristic of this filter is that it is possible to adjust its transfer characteristic over a relatively wide range by adjusting the three parameters L, S, and A. The transfer properties will be considered in the next section while here the filter structure will be presented. L and S are integers while A is a positive real number. Let W(k) be the discrete waveform to be filtered.

L is the running averaging parameter and is equal to the time window (or number of points) of the averaging operation. Each value of W(k), say $W(k_1)$, is replaced by the mean value of $W(k_1)$ and the L-1 values of W(k) following $W(k_1)$, i.e.,

$$\text{New } W(k) \equiv \frac{1}{L} \sum_{i=0}^{L-1} W(k+i) \tag{5.4.2.2-1}$$

S is the polynomial fitting parameter and is equal to the time window (or number of points) over which the least square, first order polynomial fitting [30] is performed. The window is shifted along the waveform in such a way that each point of the waveform W(k), say $W(k_1)$, comes S times in the time window. For each position of the window the S selected points are replaced by a new set of S values deduced from the fitting. The first time the point $W(k_1)$ appears in the window it is transformed to a new value $W(k_1,1)$; the second time the value $W(k_1,2)$ is obtained and so on until the value $W(k_1,S)$ is obtained.

A is the characteristic constant of a Gaussian weighting function which is applied to the S values of $W(k_1, i)$,

$$\text{New } W(k) \equiv W_{S,A}(k) = \sum_{i=1}^{S} a(i)\, W(k,i) \tag{5.4.2.2-2}$$

where

$$a(i) = a\, e^{-i^2/A^2} \tag{5.4.2.2-3}$$

$$\sum_{i=1}^{S} a(i) = 1 \tag{5.4.2.2-4}$$

The expanded form of (5.4.2.2-2) is then

$$W_{S,A}(k) = a\left[e^{-A^{-2}} W(k,1) + e^{-4A^{-2}} W(k,2) + \cdots + e^{-S^2A^{-2}} W(k,S)\right] \tag{5.4.2.2-5}$$

where

$$a = \left[e^{-A^{-2}} + e^{-4A^{-2}} + \cdots + e^{-S^2A^{-2}}\right]^{-1} \tag{5.4.2.2-6}$$

Summarizing, the composite filter employs three filtering or smoothing techniques: averaging, least squares-first order fitting, and Gaussian weighting, with the corresponding adjustable filter parameters L, S, and A, respectively. The filter is divided into two filters, the L-averaging filter and the S, A-fitting/weighting filter.

5.4.2.3 Frequency Domain Analysis of the Composite Filter

Frequency domain analysis of the composite time domain filter (or any other time domain filter) is not necessary for the application of time domain filtering. It is presented here to provide the reader with additional insight to the filtering processes by establishing a connection to the frequency domain and thus to some conventional frequency domain concepts and terminology.

Upon applying each filter, (5.4.2.2-1) and (5.4.2.2-2), to an impulse (a single data point) the impulse response of each filter can be obtained. The resultant impulse response sequences can then be Fourier transformed to the frequency domain.

The transfer function of the averaging (L-parameter) filter (5.4.2.2-1) is characterized by the cut-off frequency fco whose value decreases as L increases. For example, for N = 1024 points and a total time window of 200 ns the corresponding values for L and fco are as follows: L = 2, 3, 4, 5, 6, 7, 8, 9 and 10, and fco (MHz) = 500, 340, 255, 205, 170, 145, 130, 115 and 105, respectively.

The transfer function for an L-filter with L = 3 is shown in figure 5.9A. Also shown is the transfer function for an S, A filter with S = 5, A = 1 and the cascade (logarithmic sum) of the two filters. Note that the L filter cut-off frequency was chosen to occur at the first maximum of the S, A filter to yield a sharper cut-off characteristic and increased attenuation above fco.

The transfer properties of the S, A filter depends upon the two parameters S and A. Figure 5.10 shows how the value of A affects the transfer properties with S being held constant and equal to 5. Similar dependencies on A appear in figures 5.11 and 5.12; however, figure 5.10 provides a systematic view of the effects of A.

In figure 5.10, A is varied over the range of 0.5 through 5.0. For S = 5, the value of A = 2.05 is of special significance. As A is increased from 0.50, the first two minima shift towards each other and merge together when A equals 2.05. The resultant cut-off characteristic or minimum is very deep, about 100 dB. As A is increased above 2.05, the cut-off frequency remains the same while the overall transfer function broadens out and becomes smoother.

In figures 5.11 and 5.12 examples of composite filtering are shown. Figure 5.13 illustrates the S, A filter cut-off frequency dependence on the values of S and A, and can be used as a design tool for selecting S and A to achieve a given cut-off frequency. The cut-off frequency, fco, can be scaled with Δf; in the present case Δf = 50 MHz/div.

In figure 5.13 for a given fco there are multiple choices for the pair of S, A values. For example, if fco is desired to be 120 MHz, figure 5.13 indicates that two possible choices for S, A are 12, 4.4 and 13, 5.61. Some contours of limited extent are shown for A equal to 1.0 through 6.0 along with single points for A over the range of 2.51 through 6.49.

For $1.0 \leq A \lesssim 2.5$, the cut-off frequencies of the S, A filters in figures 5.9 and 5.10 can be compared to those in figure 5.13. Also by using a time window of 200 ns (Δf = 50 MHz) in figures 5.11 and 5.12, the cut-off frequencies in figures 5.11 and 5.12 can be compared to those in figure 5.13. For example, from figure 5.11 fco = 250 MHz occurs for S, A equal to 6, 2.40. Inspection of figure 5.13 shows that for S = 6, A would be between 2 and 3.

5.4.3 Application of Time Domain Filtering

5.4.3.1 Practical Procedure

Time domain deconvolution is a progressive process (one point after another). Knowledge or selection of the first point in the input sequence, x(1), is critical as was shown in section 5.2.5. The selection of x(1) is accomplished by trial and error using the deconvolution algorithm; however, this is not as random an exercise as it may appear to be. The search for x(1) is performed using only a small part of the available data points, say 50 points out of 1024. Typically, a choice is made for x(1) and the deconvolution algorithm is run for 50 points or so. If the result is stable, then the algorithm is applied to increasingly larger numbers of data points until the entire data span is included, with x(1) being adjusted until a value is found which provides an overall stable deconvolution result.

There are two factors that strongly interact with the selection of x(1): (1) data offset error and (2) noise in the x and y data. The offset error correction can be applied before commencing the deconvolution operation (refer to section 4.6.3). Noise in the data can only be eliminated by filtering; consequently, some combination of offset correction, x(1) selection, and filtering is necessary to achieve an acceptalbe deconvolution result. All of this is accomplished in an iteractive mode of operation. First, the operator selects the offset correction, x(1), and the filter applied to x(k) and y(k), and then judges the quality of the displayed deconvolution result. The step by step procedure is as follows:

1. Preparation of the waveforms, x(k) and y(k).
   a. Offset correction
   b. Choose x(1)

2. Filtering
   a. Choose the three filter parameters, L, S and A
   b. Apply the filter to both waveforms, x(k) and y(k)

3. Deconvolution
   a. Is the result stable, i.e., does it converge?
      If "no" go back to step 1 or step 2.
   b. Is the result smooth?
      If "no" go back to step 2
   c. Continue

4. Error Evaluation
   a. Is the error acceptable?
      If "no" go back to step 2
   b. Continue

5. Print out results: d(k) with related quality parameters (mean error, standard deviation, etc.). d(k) is the estimate of the impulse response.

The development of the appropriate filter (in 2a above) should nominally adhere to the following rules:

1. Start with the filter L = 2, S = 3, A = 100

2. Keeping L = 2, and A = 10, increase S until the result is stable (converges).

3. Keeping S and A constant increase L until the result is smooth.

4. After L and S are determined, compute the error criteria (section 3.4).
   - if the mean value of the error, $\bar{e}$, is greater than zero, decrease the value of A.
   - if the mean value of $\bar{e}$ is less than zero, increase S until $\bar{e}$ becomes greater than zero.
   The smaller the value of $\bar{e}$, the better is the result.

5. When $\bar{e}$ is close to zero, compute the standard deviation, σ.
   - if σ is too large, increase L or S and adjust A to reduce σ.

In summary, as it is evident from the above described procedures, time domain deconvolution is a progressive process. The programs used to implement time domain deconvolution were written to work in an iteractive mode. Knowing that the result must be convergent, smooth, and close to the true solution (small error function), the operator selects the offset correction, x(1), and the filter, judges the quality of the displayed result, and then modifies those selections until an acceptable result is obtained.

5.4.3.2 Estimation of the Insertion Impulse Response of 300 Meters of RG 58 C/U Coaxial Cable

The following examples of time domain deconvolution employs the same x(k) and y(k) data as were used in the frequency domain deconvolution experiments in sections 4.4.2 (two parameter method) and 4.6.3.2 (one parameter method). Consequently, the time domain results can be compared to the two frequency domain methods, and vice-versa.

The input and output waveforms x(k) and y(k), respectively, are shown in figure 5.14; these are the signals without additional noise added to them. Also, the time scales in figure 5.14 are expanded (50 ns./div.) as compared to those in figures 4.9 and 4.10 (200 ns/div.).

Figure 5.15 shows the estimated impulse response d(k) using the filter parameters L = 2, S = 5, A = 3.15 applied to x(k) and y(k). Compare this result with those of figures 4.7b and 4.12 (2-parameter F. D. method) and figure 4.34 (1-parameter F.D. method).

Figure 5.16 shows the estimated impulse response d(k) using the same filter L = 2, S = 5, A = 3.15, but with noise added to x(k) and y(k) to yield SNR-x and SNR-y values of 40 dB each. Compare this result with figures 4.7f and 4.23 (2-parameter F.D. method) and figure 4.35 (1-parameter F.D. method).

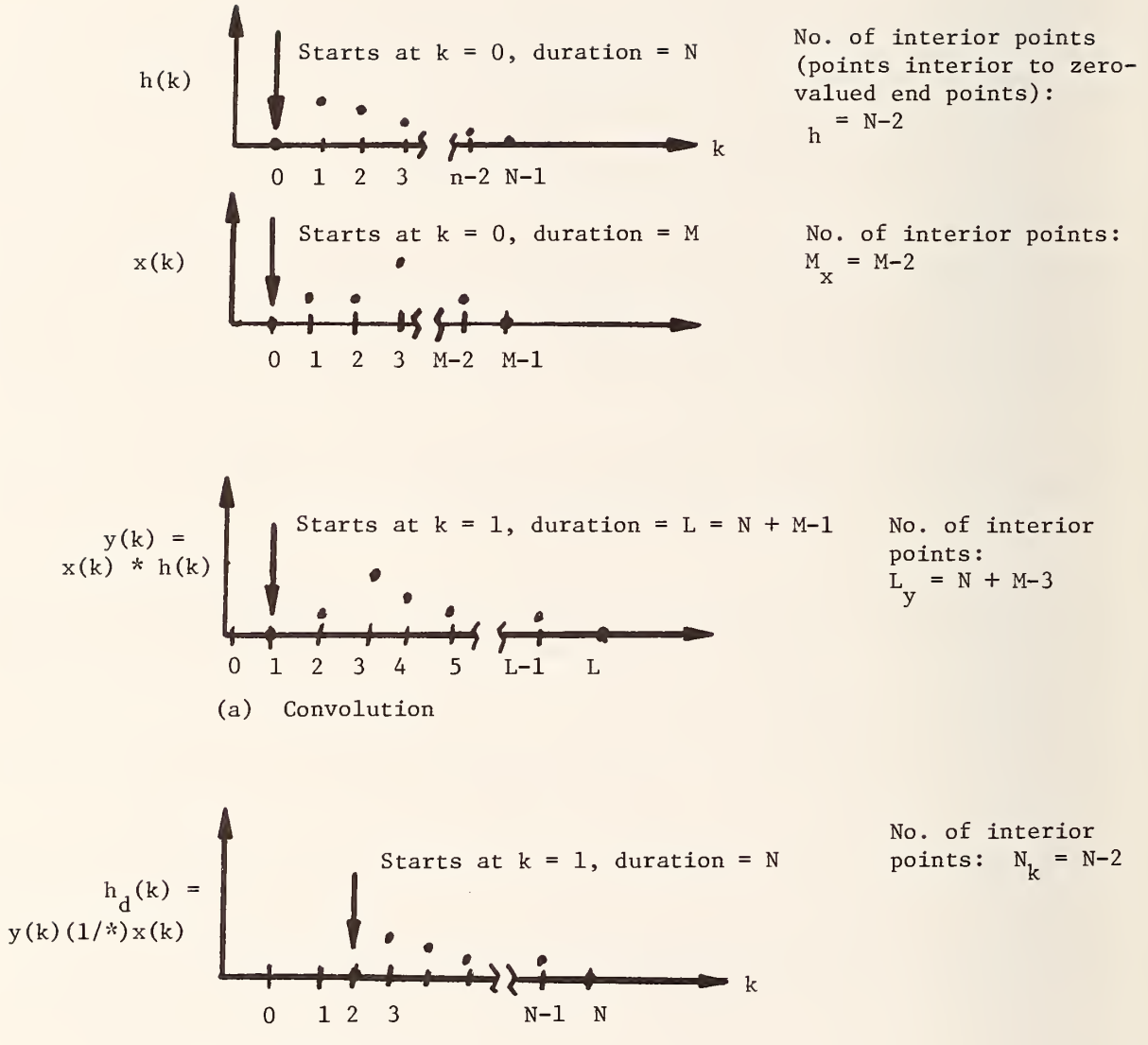


Figure 5.1 The time positions and durations of x(k), h(k), y(k) and $h_1(k)$. x(k) and h(k) are two given sequences, each commencing at k = 0; their durations are N and M, respectively. y(k) is calculated by convolving x(k) and h(k), and the result commences at k = 1. $h_1(k)$ is h(k) calculated from y(k) and x(k) using deconvolution. The calculations employ eqs. (5.2.1-8) and (5.2.2-6) or (5.2.2-21) for convolution and deconvolution, respectively.

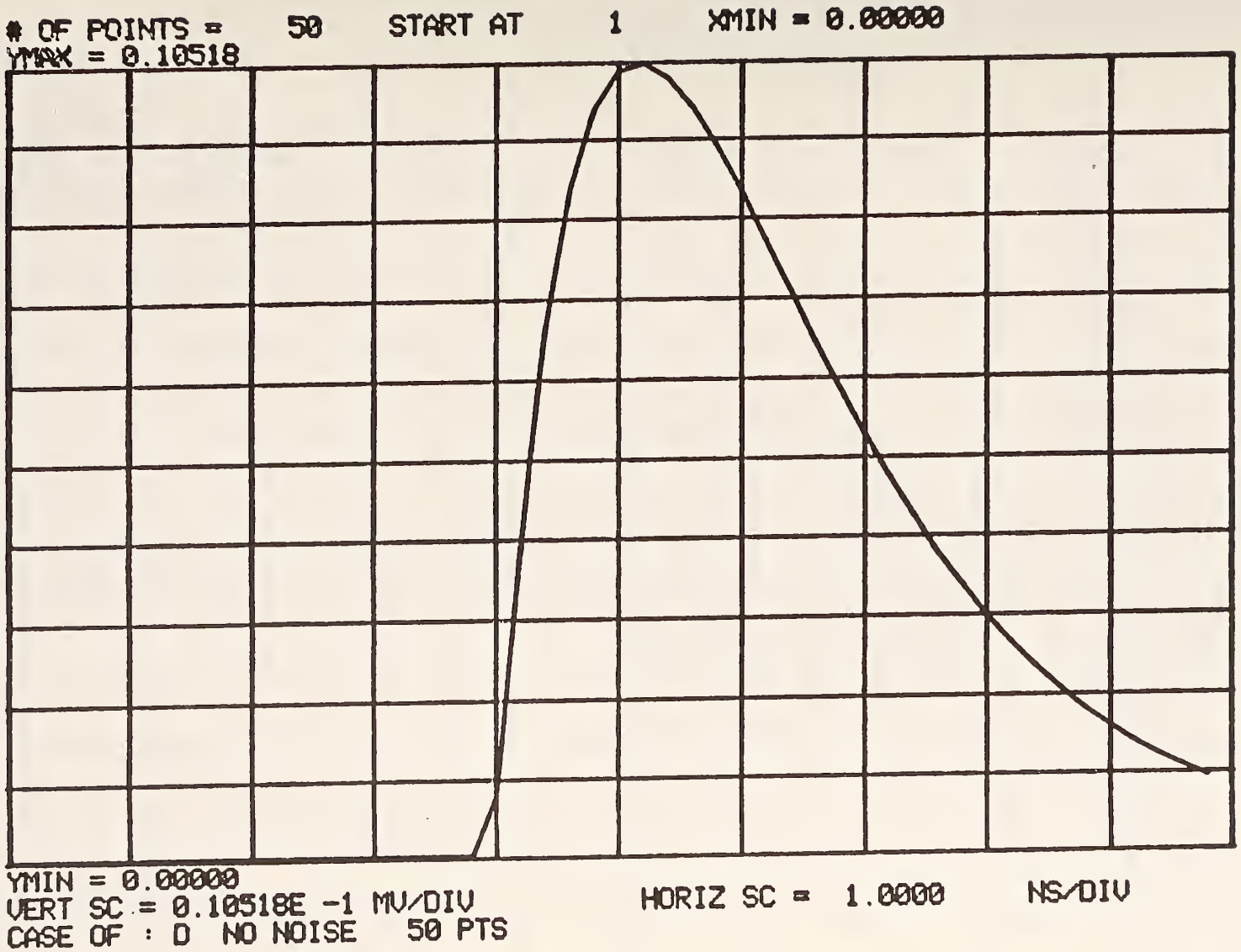


Figure 5.2 The impulse response h(k) obtained by the direct discrete classical deconvolution method from the noise free signals x(k) and y(k).

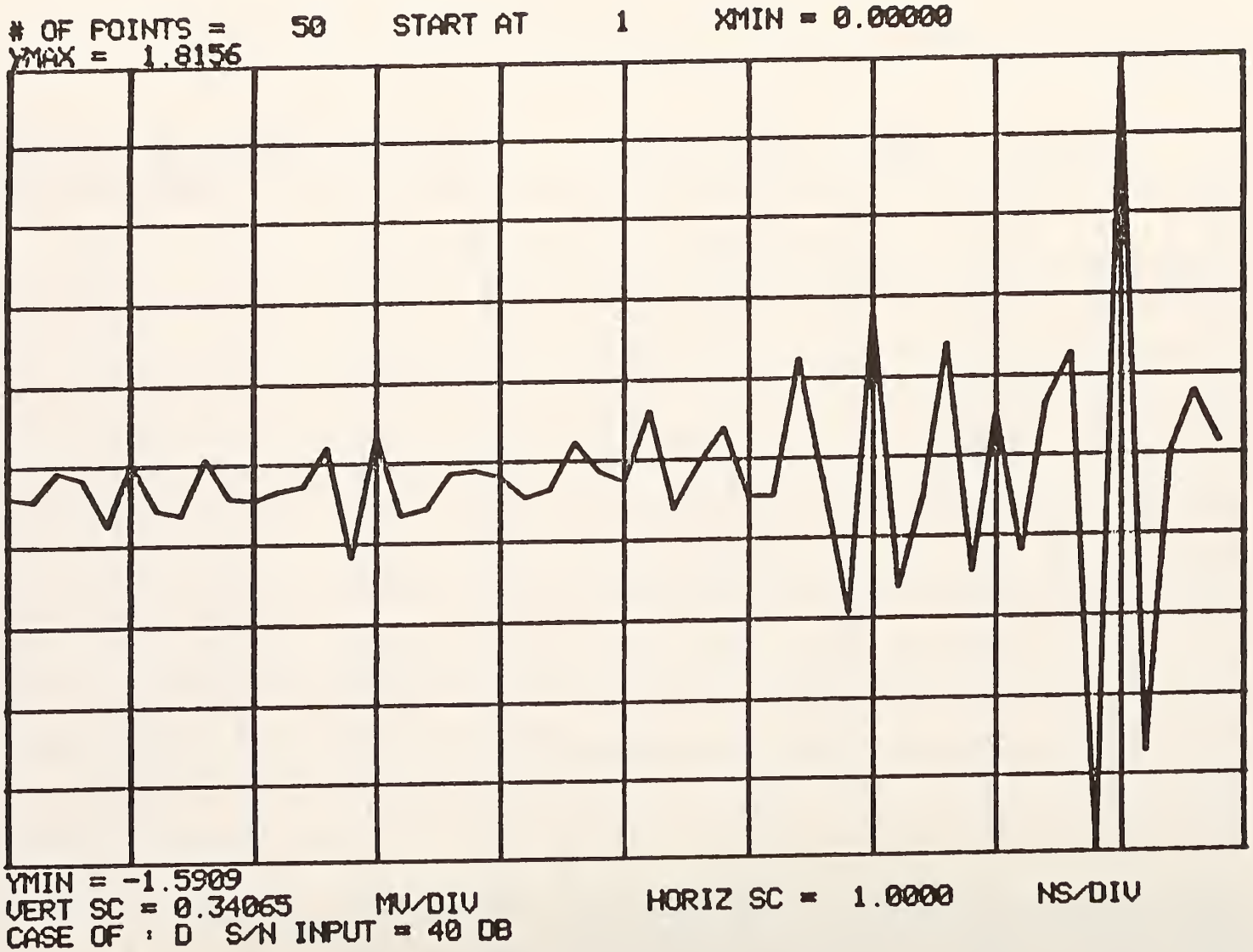


Figure 5.3 The impulse response $\tilde{h}(k)$ obtained by the direct discrete classical deconvolution method from the corrupted input and output signals $\tilde{x}(k)$ and $\tilde{y}(k)$, 40 dB signal to noise ratio for both signals. Compare to Figure 5.2.

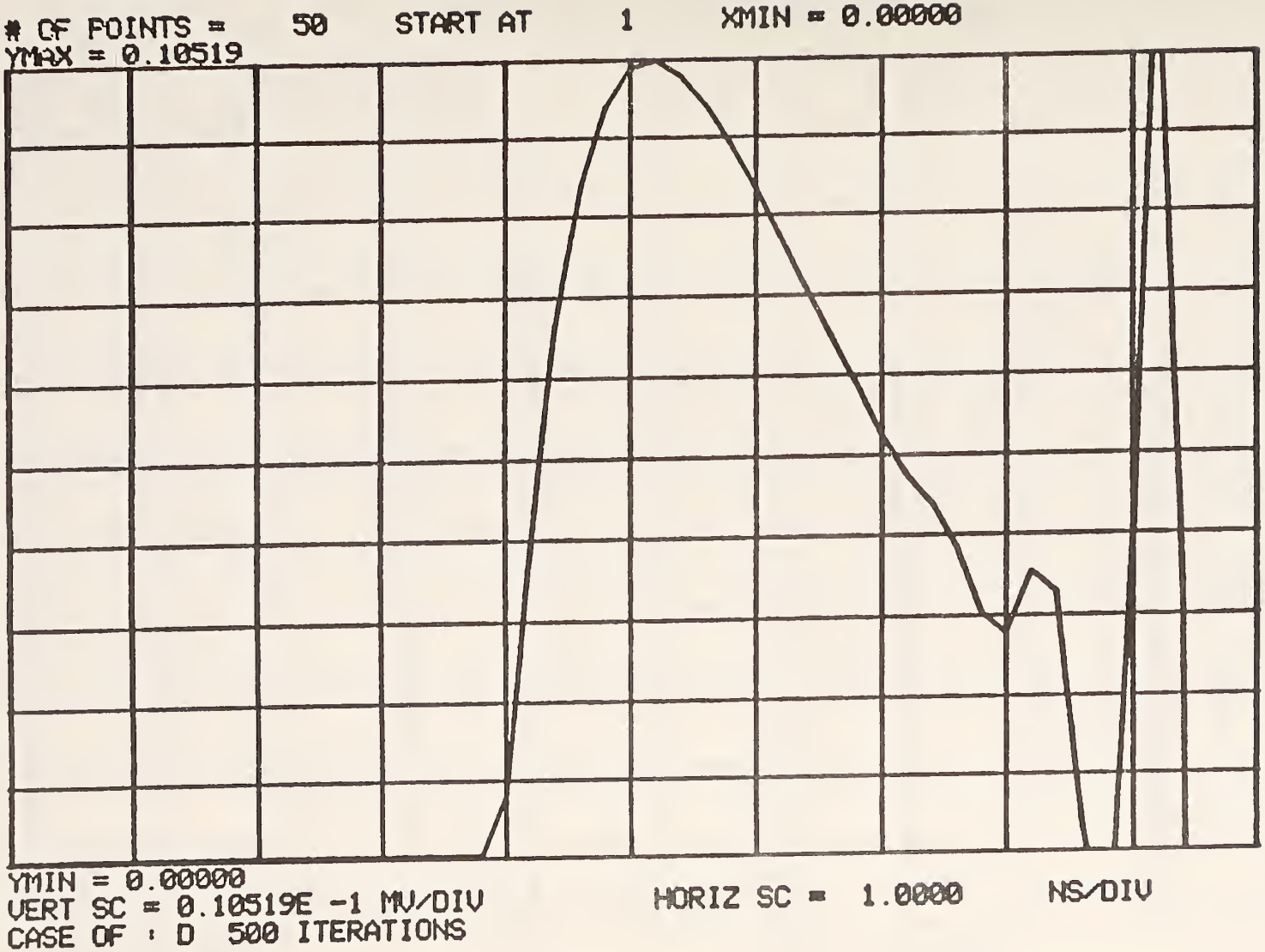


Figure 5.4 The impulse response h(k) obtained by the direct discrete iterative deconvolution method from the noise-free signals x(k) and y(k). 500 iterations.

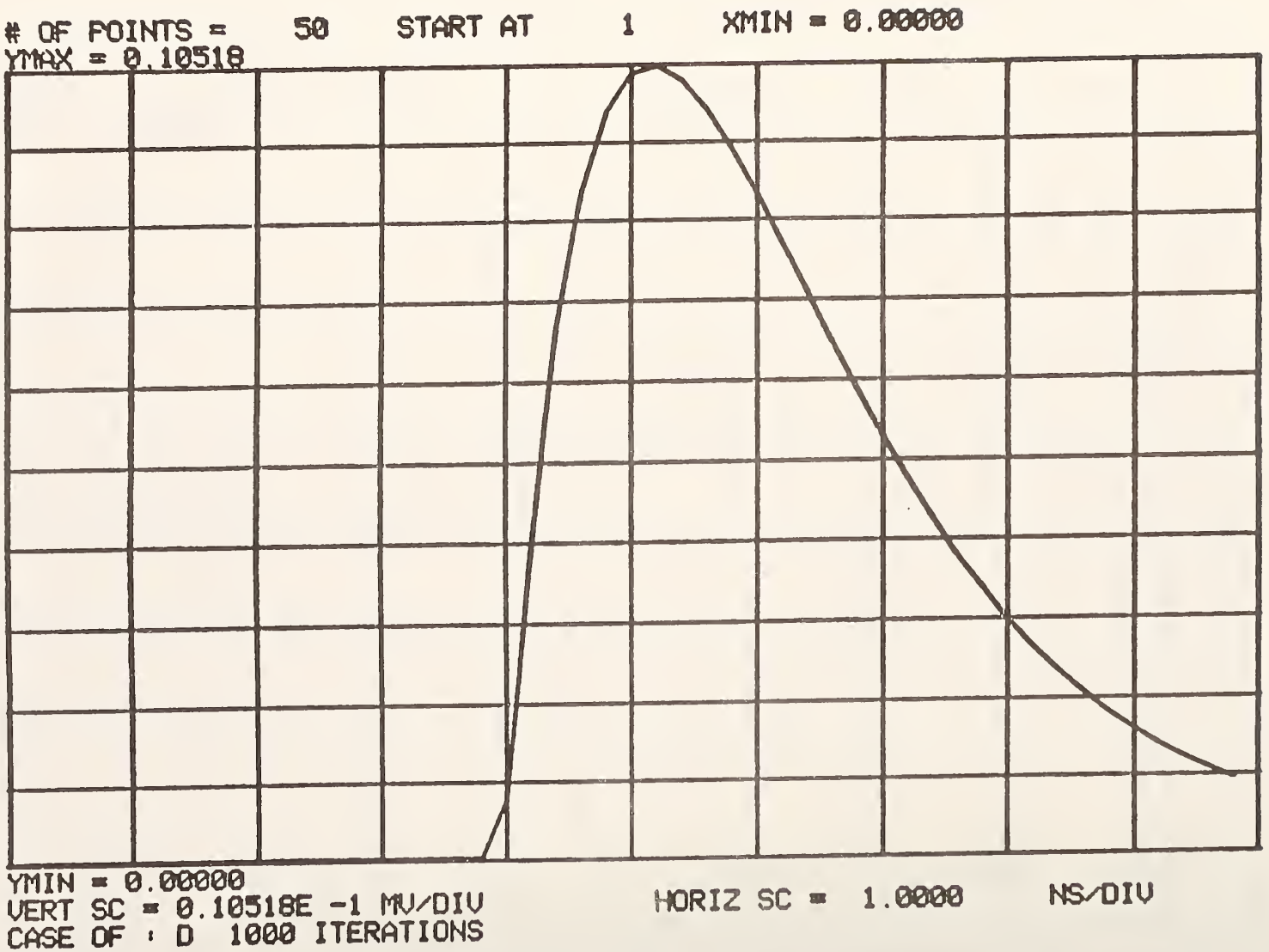


Figure 5-5 The impulse response h(k) obtained by the direct discrete iterative deconvolution method from the noise-free signals x(k) and y(k). 1000 iterations. Compare to Figure 5.2.

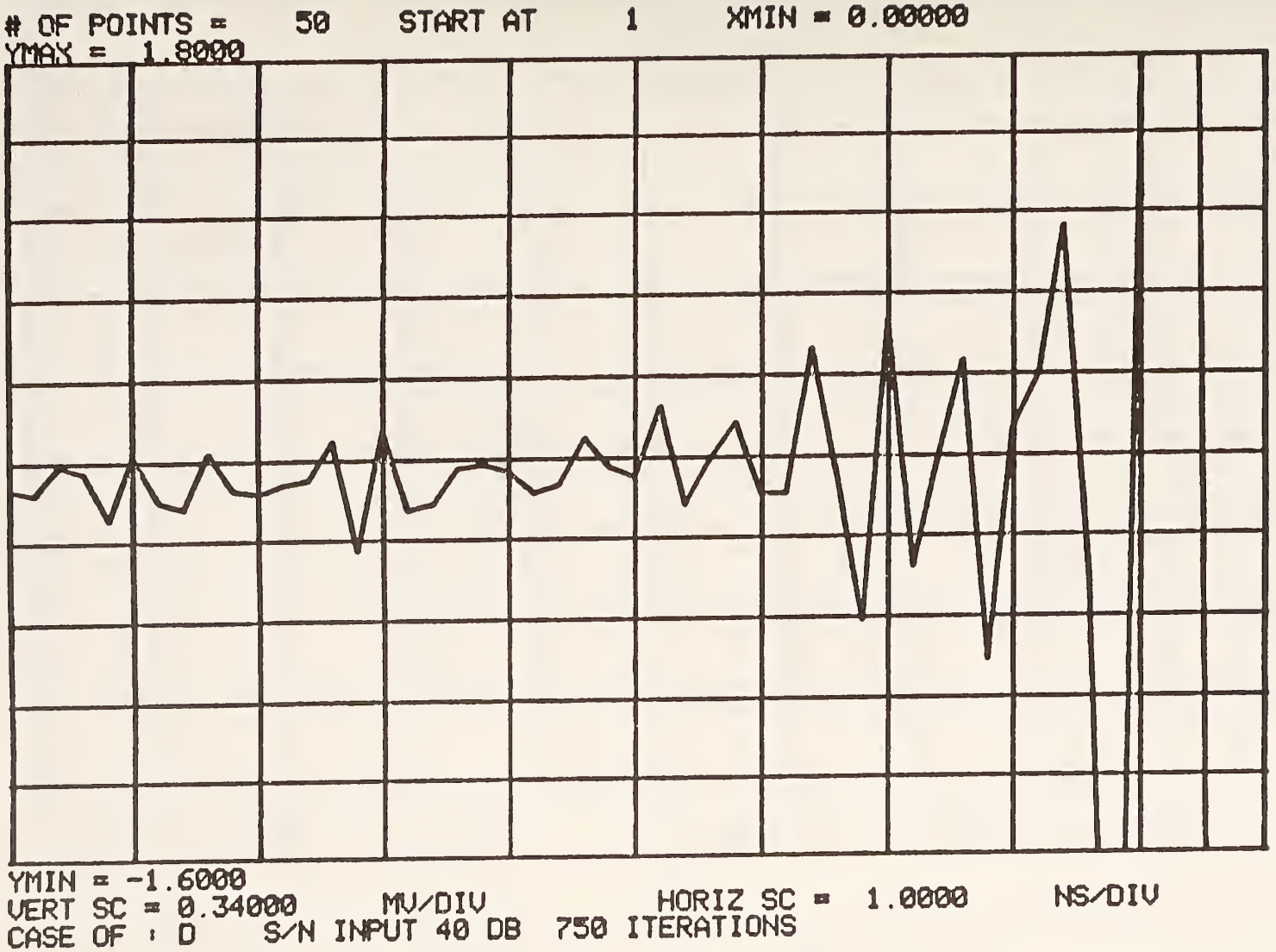


Figure 5-6 The impulse response $\tilde{h}(k)$ obtained by the direct discrete iterative deconvolution method from the corrupted input and output signals $\tilde{x}(k)$ and $\tilde{y}(k)$, 40 dB signal to noise ratio for each signal. 750 iterations. Compare to Figure 5.3.

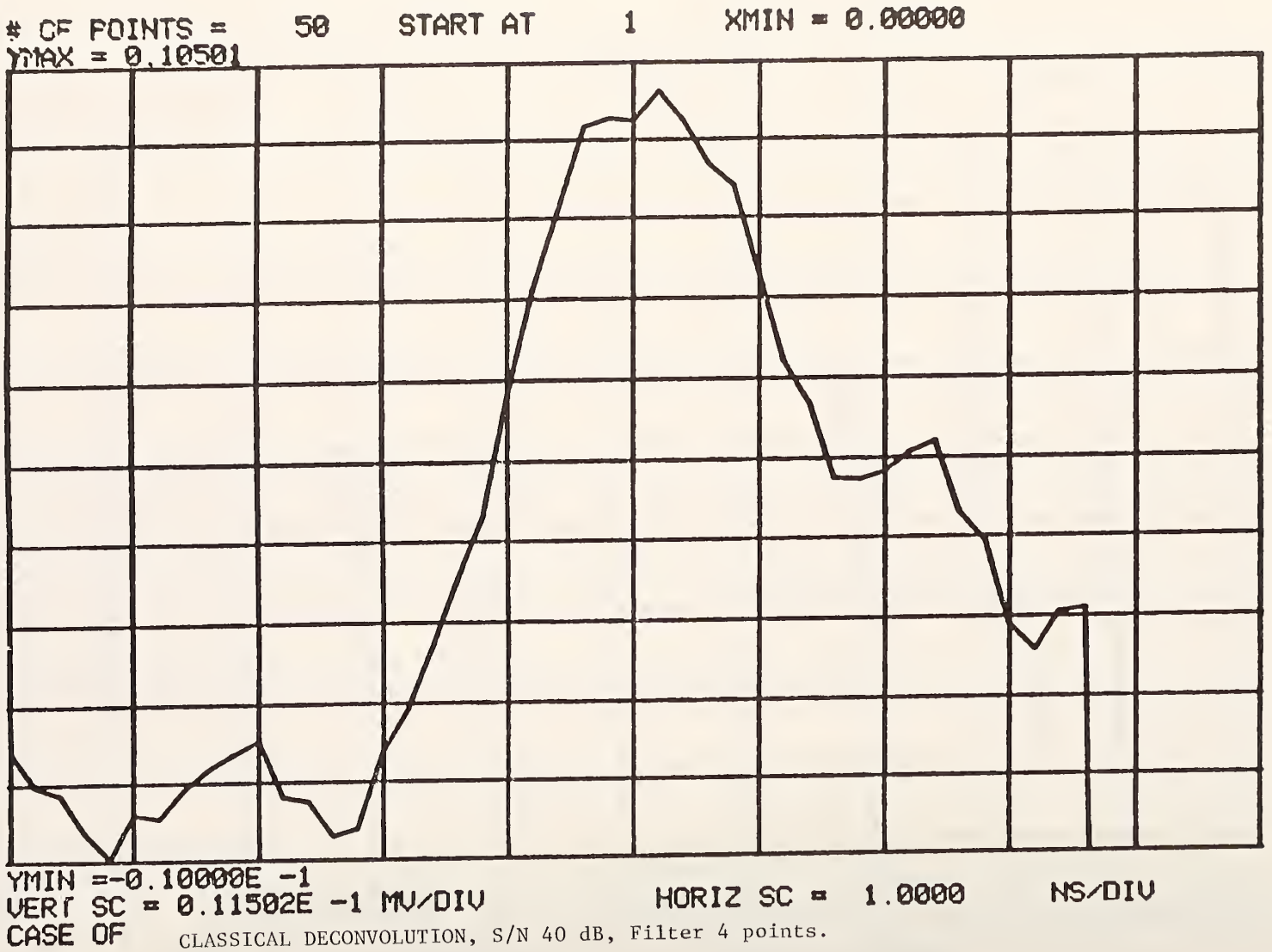


Figure 5.7 The estimate, d(k) of the impulse response h(k) obtained by the regularized discrete classical deconvolution method.

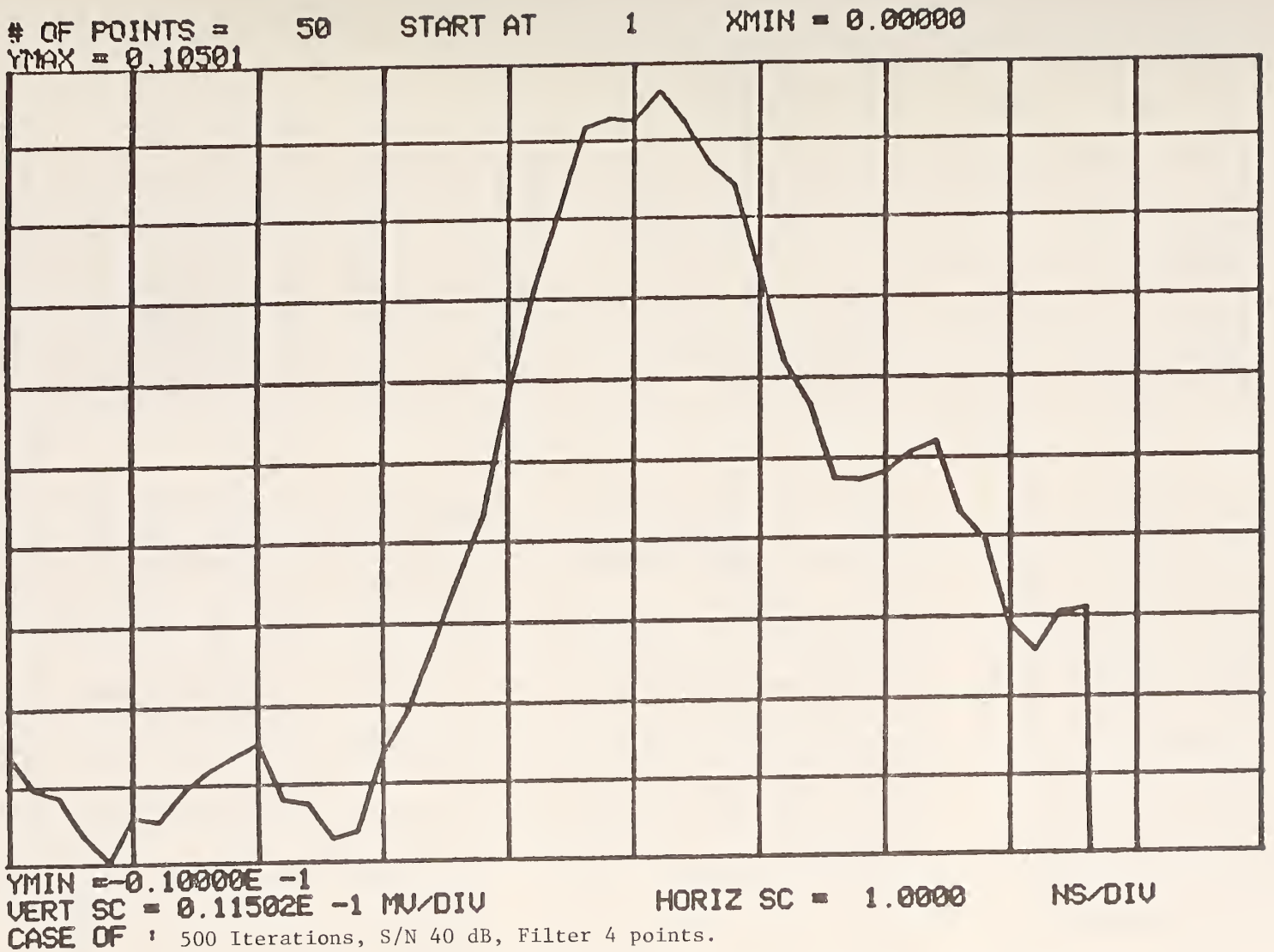
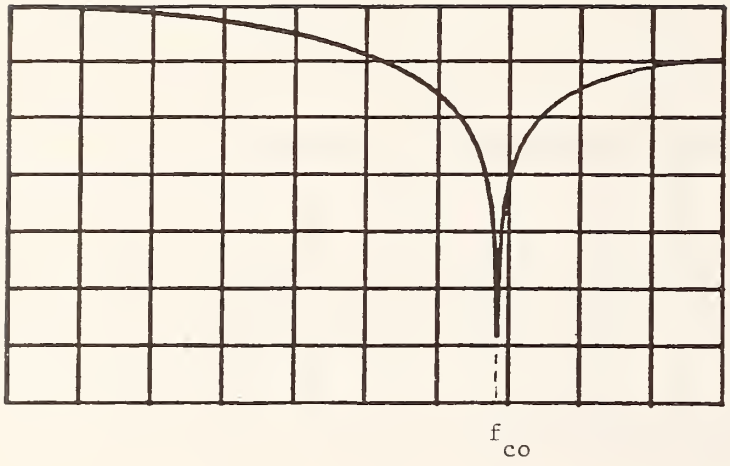
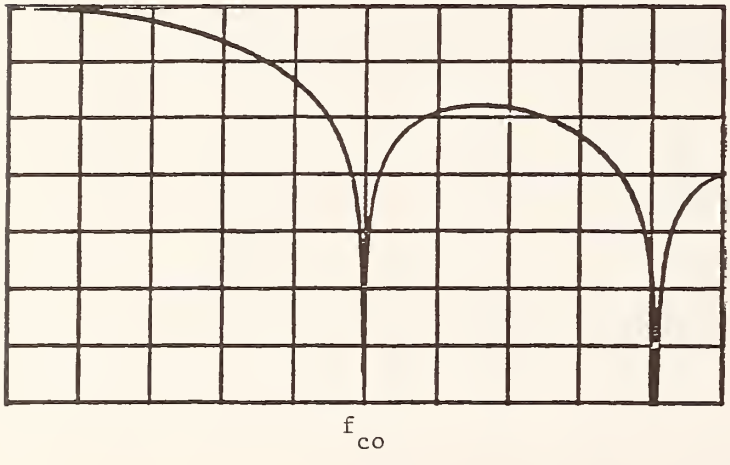


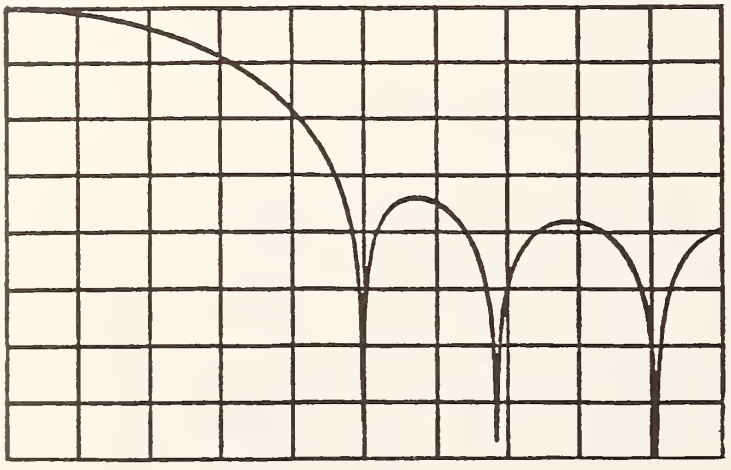
Figure 5.8 The estimate d(k) of the impulse response h(k) obtained by the regularized discrete iterative deconvolution method. Compare to Figure 5.7



(a) L- Filter. L = 3, f_{co} = 340 MHz. Vertical scale 10 dB/Div. Horizontal scale 50 MHz/Div.



(b) S, A-Filter. S = 5, A = 1 f_{co} = 250 MHz. Vertical scale 10 dB/Div. Horizontal scale 50 MHz.



(c) Combination of L and S Filters. The L filter f_{co} = 340 coincides with the S, A Filter first maximum. Vertical and horizontal scales are 10 dB/Div. and 50 MHz, Div.

Figure 5.9 Transfer properties L and S,A filters and their cascade properties.

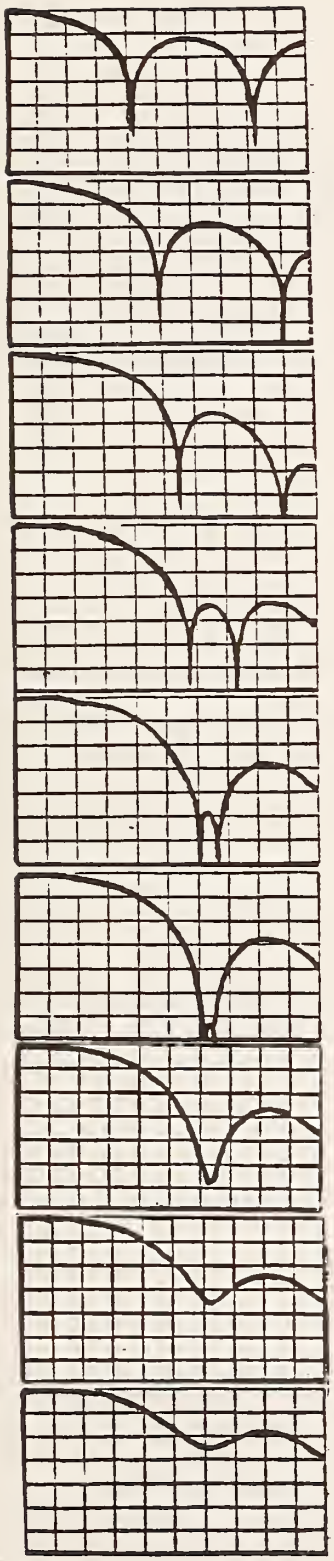
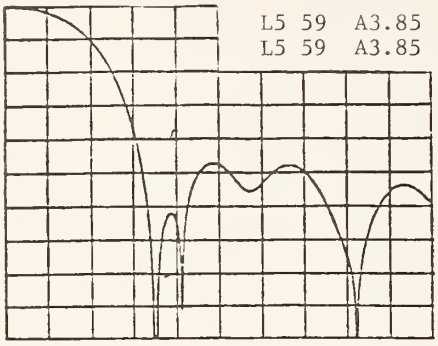
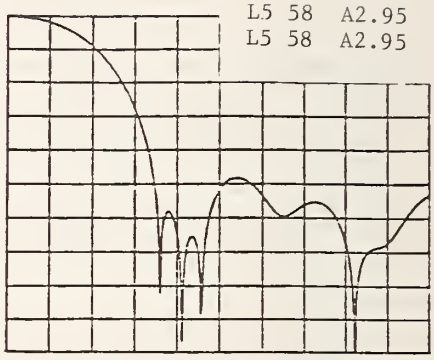
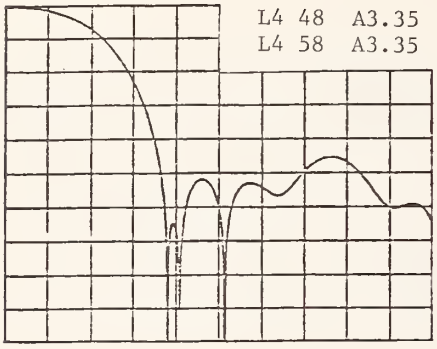
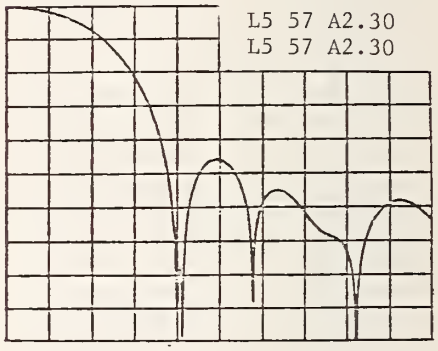


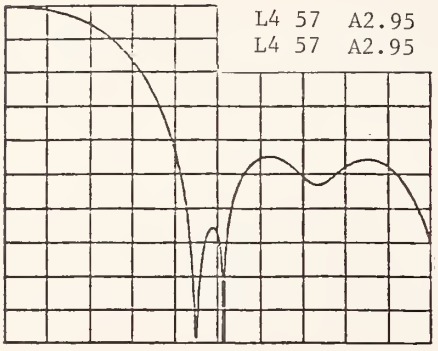
Figure 5.10 Change in the transfer properties of the S, A Filter as A is increased. S = 5.

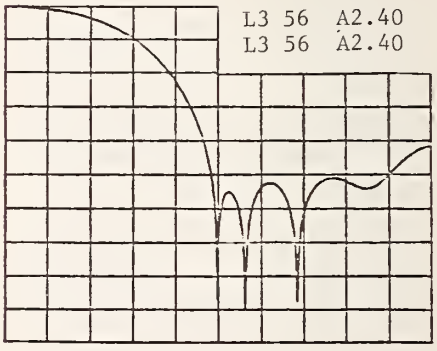


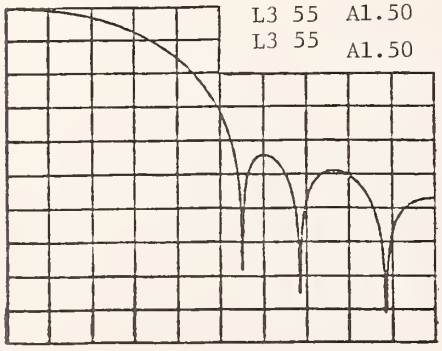












Vertical scale 10 dB/Div. Horizontal scale Δf/Div. is a function of the time window T, $\Delta f = 1/T$.

Figure 5.11 Examples of composite filters.

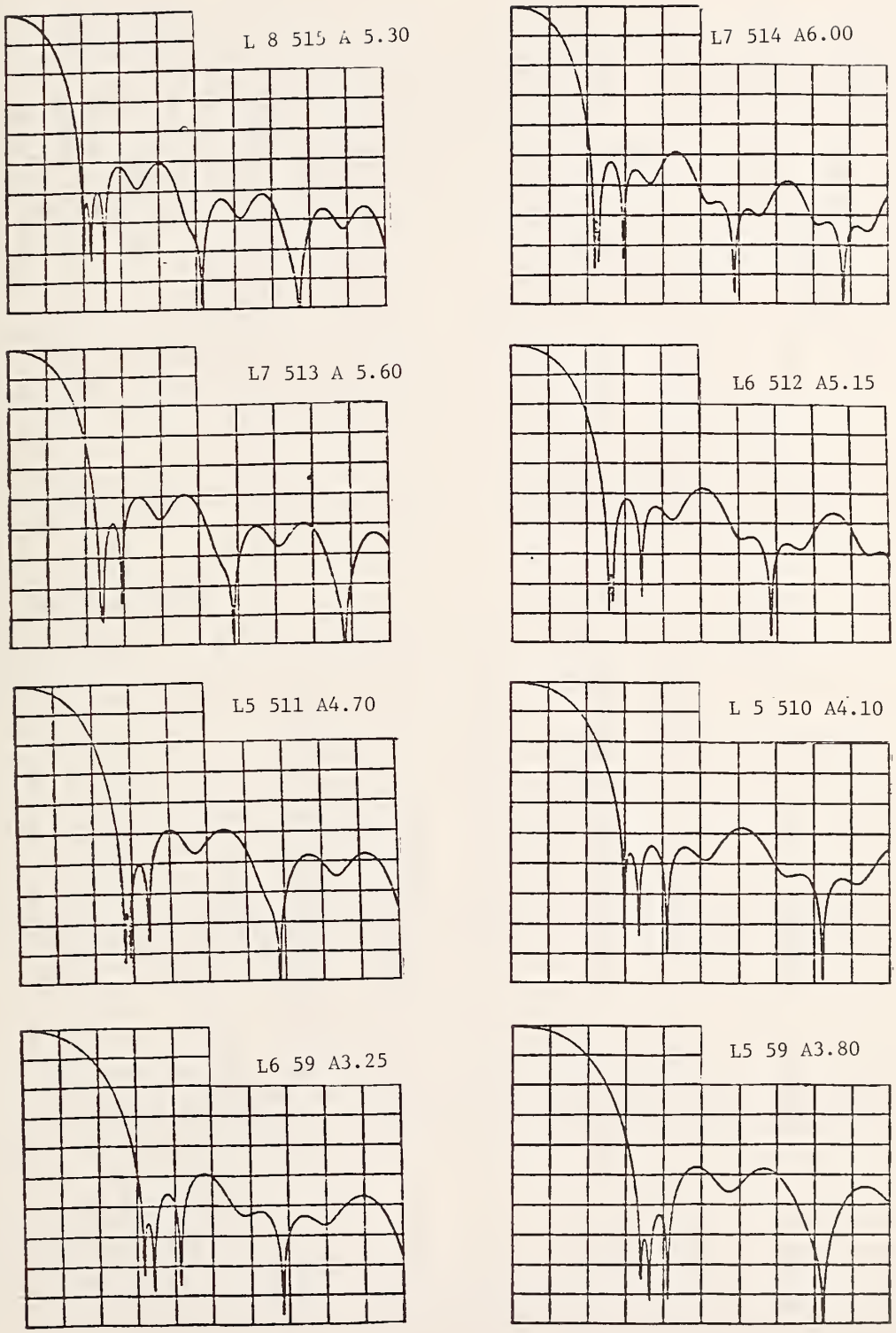


Figure 5.12 Additional examples of composite filtering scales are the same as in figure 5.12.

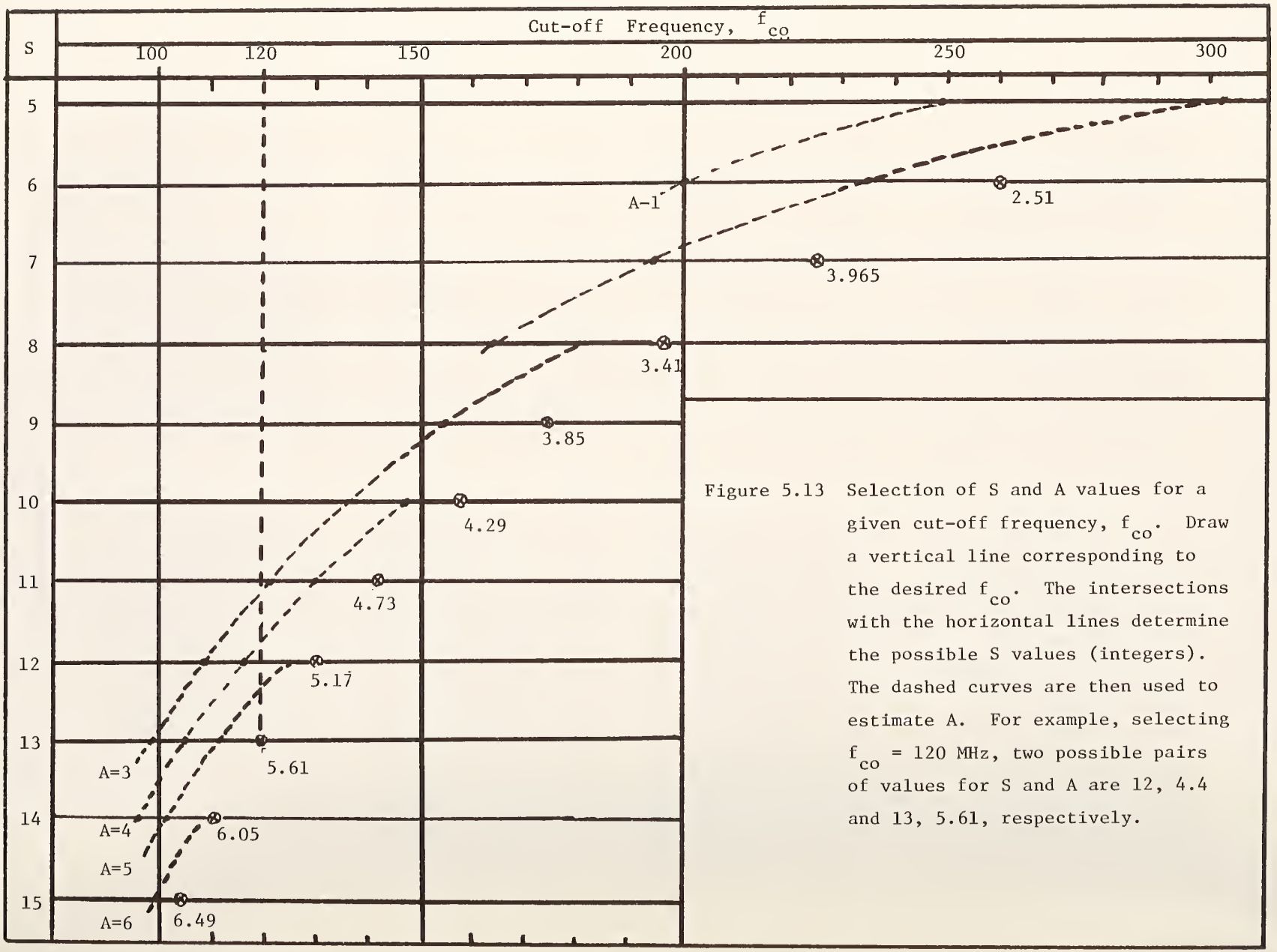


Figure 5.13 Selection of S and A values for a given cut-off frequency, f_{co}. Draw a vertical line corresponding to the desired f_{co}. The intersections with the horizontal lines determine the possible S values (integers). The dashed curves are then used to estimate A. For example, selecting f_{co} = 120 MHz, two possible pairs of values for S and A are 12, 4.4 and 13, 5.61, respectively.

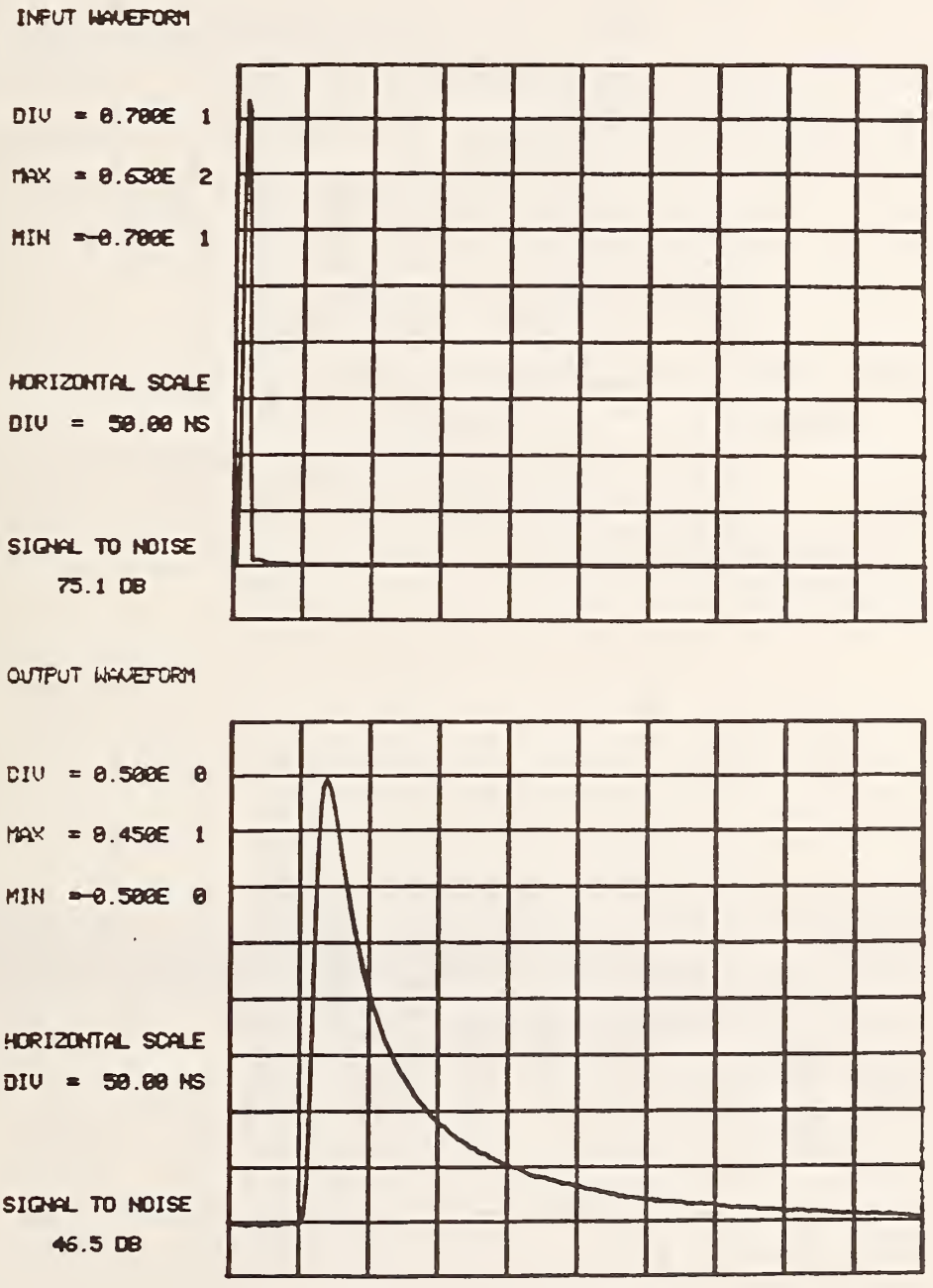


Figure 5.14 Input and output waveforms x(k) and y(k) used in the frequency and time domain deconvolution experiments before noise is added to the signals. These waveforms are identical to those in figures 4.9 and 4.10 except that the abscissae are 50 ns/Div. (as compared to 200 ns/Div.).

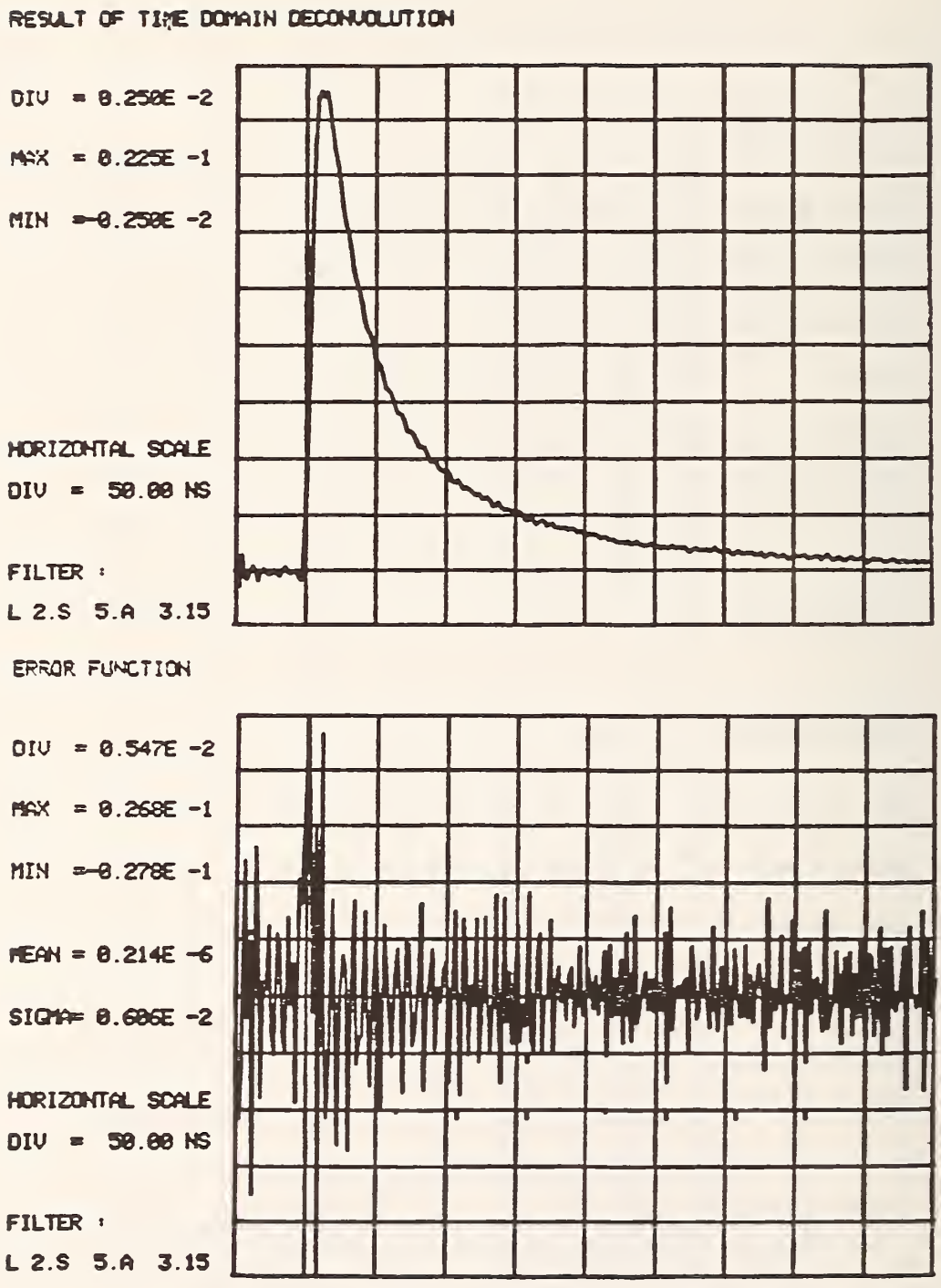


Figure 5.15 Result of the time domain deconvolution d(k) the L2 and S=5, A=3.15 filters applied to the same waveforms used in the frequency domain deconvolution experiments (figure 5.14). Compare to the frequency domain deconvolution result, figure 4.12.

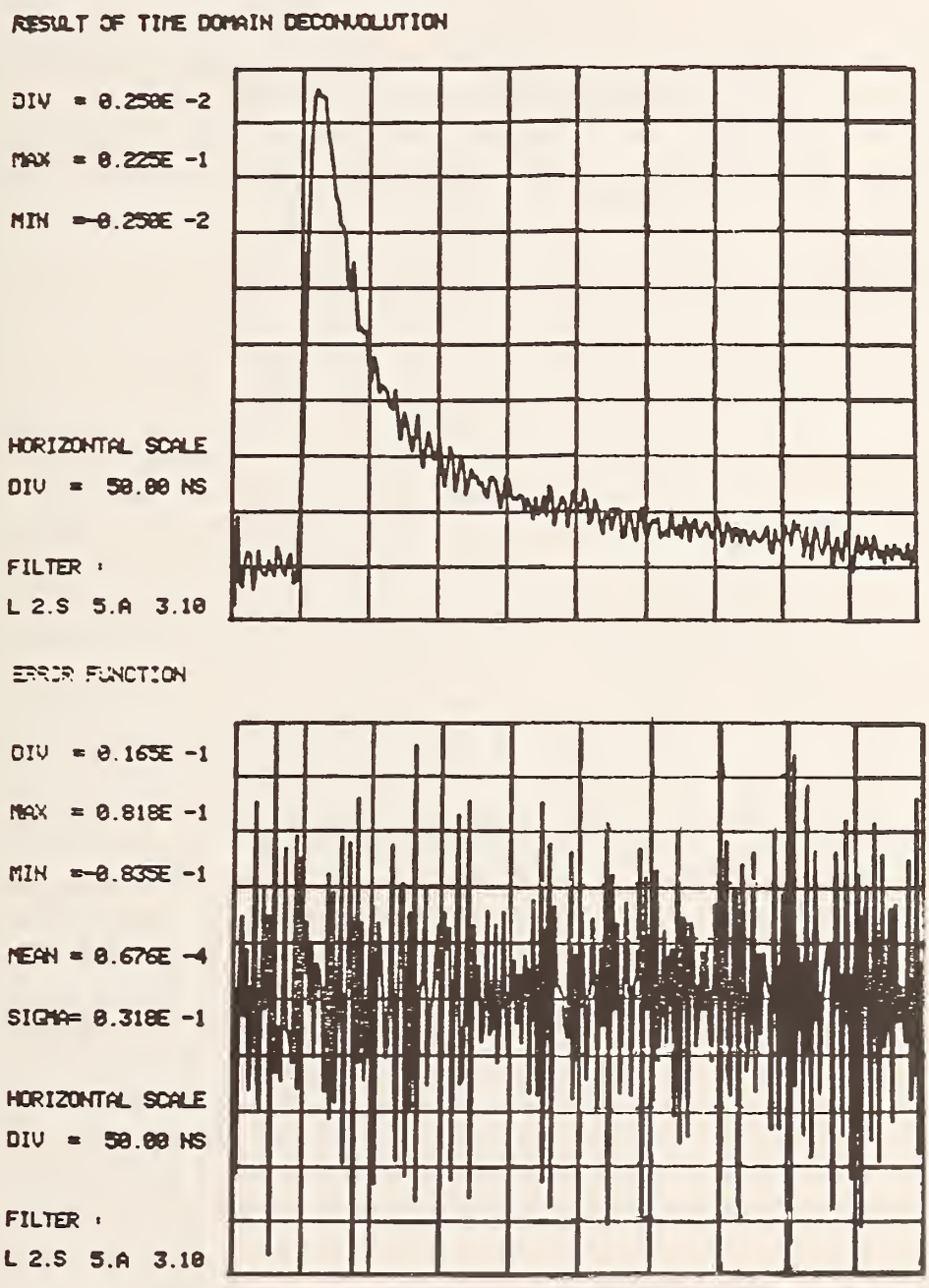


Figure 5.16 Result of time domain deconvolution d(k) when the input and output waveforms are corrupted by the addition of 35 and 6 dB of noise, respectively, (figures 4.20 and 4.21) to yield SNR-x 40 dB and SNR-y 40 dB. The filters applied to each waveform are the same as those in figure 5.15. Compare this result to the frequency domain result, figure 4.23.

6. SUMMARY AND CONCLUSIONS

6.1 Summary

The objective of the studies reported here was to investigate methods of deconvolution to be applied to time domain sequences such as those that represent the signals in sampled or discrete data systems.

In the preceeding chapters several deconvolution methods have been explored analytically and experimentally. After some introductory comments and some general presentations of the deconvolution problem and error criteria, two distinct domains for deconvolution of time domain sequences were investigated: (1) the frequency transform sequence domain, or frequency domain, and (2) the time sequence domain, or time domain.

Two methods of discrete frequency domain deconvolution were studied; the first used a two parameter filter whose parameters were subjectively selected by the operator to filter the transfer function H(n) prior to its inversion to the time domain to yield the estimated impulse response, d(k). The second method was one-parameter optimal technique in which the frequency domain filter was optimally adjusted to minimize an objective function consisting of the weighted sum of an error-energy and roughness-energy (departure from smoothness). The two methods were applied to the measurement of the impulse response of a 300 meter long coaxial cable for which experimental input and output time domain sequences were known; various degrees of computer generated noise (pseudo-random sequences) were added to the relatively noise free experimental data sequences in order to simulate the deconvolution of noisy data. Also, in addition to the coaxial transmission line experiments, the two-parameter method was applied to the measurement of the impulse response of a broadband antenna, while the one-parameter method was employed to obtain the impulse response of a waveshaping filter.

In the time domain studies, two methods were also investigated, discrete classical deconvolution and discrete iterative deconvolution; however, the latter method was not applied to any experimental data. The classical method directly inverted the discrete convolution equation to obtain the unknown impulse response. In the iterative method, the discrete convolution equation is used to form successive approximations to the impulse response; the resultant approximate impulse responses are convolved with the input signal to yield an output sequence which is then compared to the known output sequence. When the two sequences agree within some specified limits, the iteration process is stopped and the N-th iterated impulse response is declared to be equal to the sought-after or unknown impulse response. Both methods were applied in a simulation study using specified x(k), h(k), and y(k) sequences uniquely related to each other through the convolution equation, in the absence and presence of simulated noise. Also, a simple regularization filter was used to demonstrate the improvement provided by filtering to offset the effects of noise. The classical method and a more sophisticated (composite) filter were applied to the same coaxial cable experimental data which were used in the frequency domain deconvolution studies; thus the results of the frequency and time domain methods could be compared to each other. The synthesis and properties of the composite discrete time domain were discussed and developed.

6.2 Conclusions

For the deconvolution methods considered here the frequency domain one-parameter method provided the smoothest results without any obvious aberations such as Gibbs-effect. The two-parameter method may be useful in some interactive filtering situations, but for the most part, it is only of academic interest. The one-parameter method is an optimal filtering method, and as such the weighting between error and smoothness (energies) can be controlled.

The time domain deconvolution methods are still in their infancy and are purely experimental. Consequently, much work remains to be done to bring them to fruition so that they can be easily applied.

Nothing encountered in this investigation rules out the promise held for time domain methods, i.e., the elimination of transform filtering.

7. REFERENCES

[1] McGillem, C. D.; Cooper, G. R. Continuous and discrete signal and system analysis. New York: Holt, Reinhart and Winston; 178; 188; 1974.

[2] Brigham, E. O. The fast fourier transform. New Jersey: Prentice-Hall; 1974.

[3] Tikhonov, A. V.; Arsenin, V. Y. Solutions of ill-posed problems. Washington, D.C.: 1977.

[4] Biraud, Y. G. Les méthodes de déconvolution et leurs limitations fondamentales. Revue de physique appliquée. Tome 11, No. 2: 203-214; 1976 March.

[5] Jones, A. F.; Misell, D. L. The problem of error in deconvolution. J. Phys. A., Gen. Phys. 3: 462-472; 1970.

[6] Hildebrand, F. B. Introduction to numerical analysis. New York: McGraw-Hill; 1956.

[7] Von Blaricum, M. L.; Mittra, R. Problems and solutions associated with Prony's method for processing transient data. IEEE Trans. EM Compat. EMC-20(1): 174-182; 1978 Feb.

[8] Riad, S. M.; Nahman, N. S. Application of the homomorphic deconvolution for the separation of TDR signals occurring in overlapping time windows. IEEE Trans. Instrum. & Meas. IM-25(4): 388-391; 1976 Dec.

[9] Gans, W. L. Present capabilities of the NBS automatic pulse measurement system. IEEE Trans. Instrum. Meas. IM-25(4): 384-388; 1976 Dec.

[10] Gans, W. L.; Andrews, J. R. Time domain automatic network analyzer for the measurement of RF and microwave components. Nat. Bur. Stand. (U.S.) Tech Note 672; 1975 September.

[11] Op. Cit. [1]. Section 5-18.

[12] Op. Cit. [2]. Chapter 8.

[13] Digital Signal Processing Committee, ed. Programs for digital signal processing. New York: IEEE Press; 1979.

[14] Oppenheim, A. V.; Schafer, R. W. Digital signal processing. New Jersey: Prentice-Hall; 1975.

[15] Op. Cit. [1]. Section 5-15.

[16] Op. Cit. [13]. 405, 410.

[17] Kuo, Franklin F. Network analysis and synthesis. New York: John Wiley and Sons; 1962; Chap. 13.

[18] Twomey, S. The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements. J. Franklin Inst. 279: 95-108, 1965 Feb.

[19] Hunt, B. R. Deconvolution of linear systems by constrained regression and its relationship to the Weiner Theory. IEEE Trans. on Automatic Control AC-17(5): 703-705: 1972 October.

[20] Op. Cit. [14]. 204.

[21] Personal communication from Dr. Yves Balcou, Dept. of Crystalline Physics and Chemical Structure, University of Rennes, Rennes, 35000, France.

[22] Op. Cit. [14], p. 125.

[23] Op. Cit. [1], p. 37

[24] Gans, W. L.; Nahman, N. S. A note on the conversion of an experimentally obtained step-like waveform into a duration limited waveform for implementing Fourier transform analysis. Submitted for publication IEEE I&M Transactions.

[25] Andrews, J. R.; Pulse reference waveform standards development at NBS. Proceedings of the ATE Seminar/Exhibit, Pasadena, CA., Jan. 1981, pp. IV-13-IV-19. Benwill Publishing Corp., Boston, MA. 02215.

[26] Op. Cit. [1], p. 89.

[27] Janson, Peter A. Method for determining the response function of a high-resolution infrared spectrometer. J. of the Opt. Soc. Amer., Volume 60, No. 2., 194-191, Feb. 1970.

[28] van Cittert, P. H. Z. Physik, Vol. 69, 298, (1931).

[29] Op. Cit. [1], Chapter 7.

[30] Pipes, L. A.; Harwill, L. R. Applied mathematics for engineers and physicists", 3rd Ed., pp. 562-567, McGraw-Hill Book Co. 1970.

NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBS TN-1047	2. Performing Organ. Report No.	3. Publication Date October 1981

4. TITLE AND SUBTITLE

DECONVOLUTION OF TIME DOMAIN WAVEFORMS IN THE PRESENCE OF NOISE

5. AUTHOR(S) Norris S. Nahman and M. E. Guillaume

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

Deconvolution or inverse filtering was used to determine the impulse response of a system using noisy input and output time domain sequences (discrete data). Frequency and time domain methods were studied along with the synthesis of the filters required to obtain stable and smooth results. For the methods studied it was concluded that the superior technique was provided by an optimal frequency domain method implemented via the FFT. Also, it is pointed out that the time domain methods are only in their infancy and still retain the promise of avoiding transform domain filtering. Examples are presented in which the impulse responses are determined in the presence of varying degrees of noise for a coaxial transmission line, a wave-shaping filter, and a broad-band antenna.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

deconvolution; impulse response; inverse filtering; noise contamination; time domain; waveforms.

13. AVAILABILITY

☒ Unlimited
☐ For Official Distribution. Do Not Release to NTIS
☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES 122

15. Price $11.00

USCOMM-DC 6043-P80

☆U.S. GOVERNMENT PRINTING OFFICE:1981—576-001 / 1206

NBS TECHNICAL PUBLICATIONS

PERIODICALS

JOURNAL OF RESEARCH—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. As a special service to subscribers each issue contains complete citations to all recent Bureau publications in both NBS and non-NBS media. Issued six times a year. Annual subscription: domestic $13; foreign $16.25. Single copy, $3 domestic; $3.75 foreign.

NOTE: The Journal was formerly published in two sections: Section A "Physics and Chemistry" and Section B "Mathematical Sciences."

DIMENSIONS/NBS—This monthly magazine is published to inform scientists, engineers, business and industry leaders, teachers, students, and consumers of the latest advances in science and technology, with primary emphasis on work at NBS. The magazine highlights and reviews such issues as energy research, fire protection, building technology, metric conversion, pollution abatement, health and safety, and consumer product performance. In addition, it reports the results of Bureau programs in measurement standards and techniques, properties of matter and materials, engineering standards and services, instrumentation, and automatic data processing. Annual subscription: domestic $11; foreign $13.75.

NONPERIODICALS

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The principal publication outlet for the foregoing data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the **above** *NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

Order the **following** *NBS publications—FIPS and NBSIR's—from the National Technical Information Services, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Services, Springfield, VA 22161, in paper copy or microfiche form.

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Washington, D.C. 20234

OFFICIAL BUSINESS

Penalty for Private Use, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF COMMERCE
COM-215

U.S.MAIL

SPECIAL FOURTH-CLASS RATE
BOOK