AD A078217

DDC FILE COPY

LEVEL II

12

POLYTECHNIC INSTITUTE OF NEW YORK

FINAL REPORT

September 1, 1976 - September 30, 1979

TO

OFFICE OF NAVAL RESEARCH

DDC
RECEIVED
DEC 14 1979
E

CONTRACT NO. N00014-76-C-0144
(ARPA Order 2684)

Sponsored by

ADVANCED RESEARCH PROJECTS AGENCY

by

A. Papoulis

This document has been approved for public release and sale; its distribution is unlimited.

79 12 13 012

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Contributions to Signal Processing for the HALLO Program.	5. TYPE OF REPORT & PERIOD COVERED Final Technical Report Sept 1976-Sept 1979,
	6. PERFORMING ORG. REPORT NUMBER POLY-EE-79-058
7. AUTHOR(s) A. Papoulis	8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0144, ARPA Order-2684
9. PERFORMING ORGANIZATION NAME AND ADDRESS Polytechnic Institute of N. Y. Route 110, Farmingdale, N. Y. 11735	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order No. 2684
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd., Arlington, Va. 22217	12. REPORT DATE November 1979
	13. NUMBER OF PAGES 107
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Naval Research Dept. of the Navy Arlington, Va. 22217	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

clutter, target identification, deconvolution, Filtering, Image restoration, Hidden periodicities, underwater detection, spectral estimation, sampling.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A number of topics were investigated related to image restoration, clutter reduction, detection of moving targets, spatial and temporal filtering, and spectral estimation. These topics originated with various aspects of the HALLO program but have more general applications including underwater detection, target identification from undersampled radar returns, deconvolution, ultrasonics, detection of points and edges, detection of hidden periodicities and speech processing.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

411 284



Section 1

1. Introduction

We have investigated a number of topics related to image restoration, clutter reduction, detection of moving targets, spatial and temporal filtering, and spectral estimation. These topics originated with various aspects of the HALLO program but have more general applications including underwater detection, target identification from undersampled radar returns, deconvolution, ultrasonics, detection of points and edges, detection of hidden periodicities and speech processing.

In the course of this investigation, we developed and applied the following two general methods of signal processing:

Extrapolation of bandlimited signals We introduced a numerical algorithm for estimating a signal f(t) beyond a given interval (-T, T) under various assumptions involving mainly its spectrum.

Adaptive frequency domain filtering and estimation We introduced a method of filtering based on the notion of the running Fourier spectrum.

The investigation led to a number of reports, papers presented on various conferences including a prize-winning method of spectral estimation, and papers published in professional journals. We list below the principal results of the last two years:

Conference Papers

Image Restoration (section 2)

Sixth Strategic Space Symposium (S^3-W), S.R.I. International, Menlo Park, California, March, 1978.

Adaptive Extrapolation and Hidden Periodicities (section 3)

RADC Spectral Estimation Workshop, Rome Air Development Center, Rome, N. Y., May 1978. (The method was applied to the solution of problems proposed and was a winner in a spectral estimation competition.

Improvement of Range Resolution by Spectral Extrapolation

Third International Symposium on Ultrasonics, Imaging, and Tissue Characterization, National Bureau of Standards, Gaithersburg, Md., June 1978.

Adaptive Frequency Domain Estimators

IEEE International Symposium on Information Theory, Grignano, Italy, June 1979.

Spectral Estimation from Random Samples

IEEE International Conference on Information Sciences and Systems, Patras, Greece, July 1979.

Papers in Professional Journals

The Two-to-One Rule in Data Smoothing

IEEE Trans. on Information Theory, September 1977, vol. IT-23, no. 5, pp. 631-633.

Generalized Sampling Expansion

IEEE Trans. on Circuit & Systems, November 1977, vol. CAS-24, no. 11, pp. 652-654.

The Problem of Transmission Zeros in Deconvolution

IEEE Trans. on Information Theory, January 1978, vol. IT-24, no. 1, pp. 126-128.

The Factorization Problem for Time-Limited Functions and Trigonometric Polynomials

IEEE Trans. on Circuit & Systems, Jan. 1978, vol. CAS-25, no. 1, pp. 41-45.

Papers in Professional Journals (cont'd)

Identification of Systems Driven by Non-stationary Noise

IEEE Trans. on Information Theory, March 1978, vol. IT-24, no. 2, pp. 240-244.

Improvement of Range Resolution by Spectral Extrapolation

Ultrasonic Imaging, vol. 1, no. 2, pp. 121-135, April 1979. (section 4)

Detection of Hidden Periodicities by Adaptive Extrapolation

Acoustics, Speech and Signal Processing, vol. ASSP-27, no. 5, pp. 492-500, October 1979. (section 5)

Theory and Applications of Running Transforms

Submitted to the IEEE Transactions on Circuits and Systems.

Adaptive Frequency Domain Filtering and Estimation

Submitted to the IEEE Transactions on Information Theory.

In the following sections, we attach several of the above papers.

27. IMAGE RESTORATION*

By: A. Papoulis and C. Chamzas+

Abstract. An extrapolation method is presented for recovering an object from its diffraction limited image. The high frequency components of the object, lost due to the finite size of the aperture, are restored by a numerical iteration involving only FFT. The method is particularly effective for objects consisting of isolated points. Various illustrations show that the resulting resolution far exceeds the Rayleigh limit.

A modification of the method yields an algorithm for recovering a distant object obscured by low-frequency clutter. The noise is removed with suitable filtering and the lost frequency components of the object are recovered by extrapolation.

1. Spectral Extrapolation

We shall develop a method for restoring an object that is distorted by a diffraction limited imaging system. For simplicity, we assume that the object is one-dimensional and completely coherent. The results can be readily extended to two-dimensions and to incoherent objects.

In Fig. 1, we show a simplified schematic of the imaging system. We denote by $f(x)$ and $g_o(x)$ the amplitude of the object and of its image respectively, and by $F(u)$ and $G_o(u)$ their Fourier transforms. Assuming that the system is ideal, we have

$$G_o(u) = \begin{cases} F(u) & |u| < r \\ 0 & |u| > r \end{cases} \tag{1}$$

where r is the radius of the aperture.

*Work sponsored by DARPA, Contract No. N00014-67-A-0438-0017

+ A. Papoulis is professor of electrical engineering at the Polytechnic Institute of New York, Farmingdale, N.Y., 11735. C. Chamzas is a Ph.D. student at the same school.

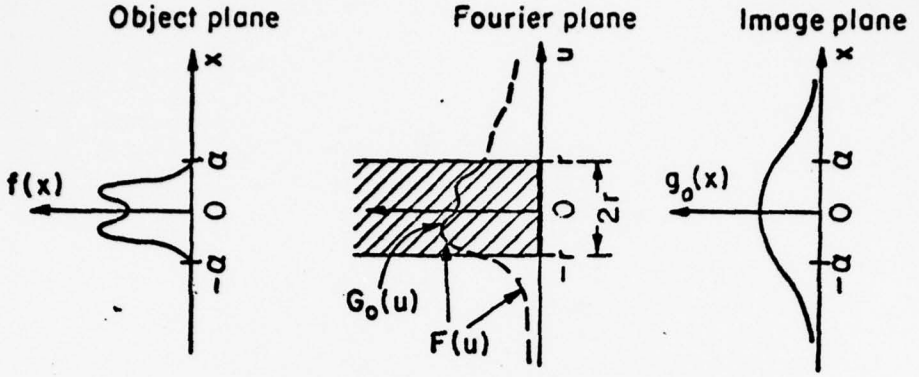


FIGURE 1 DIFFRACTION-LIMITED IMAGING SYSTEM: OBJECT f(x), IMAGE $g_o(x)$ AND FIELD F(u) ON FOURIER PLANE

We are given $g_o(x)$ and our problem is to find f(x).

a. Iteration

We shall show that, if the object f(x) vanishes for $|x| > a$, then it can be completely recovered from $g_o(x)$ by numerical iteration (Refs. 1-3).

For this purpose, we compute, first the Fourier transform $G_o(u)$ of $g_o(x)$. Clearly, $G_o(u) = F(u)$ for $|u| < r$ and $G_o(u) = 0$ for $|u| > r$.

We truncate $g_o(x)$ forming the function

$$f_1(x) = \begin{cases} g_o(x) & |x| < a \\ 0 & |x| > a \end{cases} \tag{2}$$

We find the Fourier transform $F_1(u)$ of $f_1(x)$ and form the function

$$G_1(u) = \begin{cases} F_1(u) & |u| > r \\ F(u) & |u| < r \end{cases} \tag{3}$$

We find the inverse transform $g_1(x)$ of $G_1(u)$. This completes the first step of the iteration (Fig. 2).

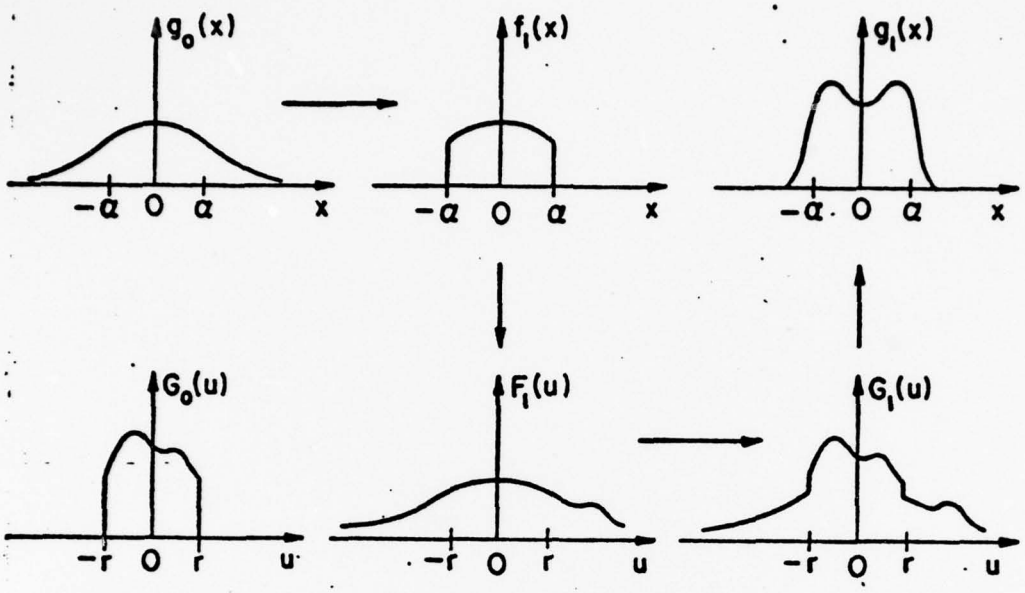


FIGURE 2 FIRST ITERATION STARTING WITH KNOWN IMAGE $g_o(x)$ AND ITS FOURIER TRANSFORM $G_o(u)$

The next step is formed by replacing in Eq. (2) the function $g_o(x)$ by the function $g_1(x)$ and proceeding as in the first step.

The nth step is similar: starting from $g_{n-1}(x)$, we form the function $f_n(x)$ as in Eq. (2), the function $G_n(u)$ as in Eq. (3), and its inverse $g_n(x)$.

We prove in Ref. 1 that, in the absence of noise, the iteration converges to the unknown object f(x). This is not the case if noise is present. However, even then the method can be used to enhance the distorted image. This is done by terminating the iteration at an appropriate step so as to minimize the overall error (Ref. 1).

b. Point objects. If additional information about the object is available, then the iteration speed increases and the noise problem can be relaxed. A case of particular interest in many applications involves objects consisting of distinct points (neighboring stars, for example). In this case, f(x) consists of sharp peaks at certain points and the problem is to find the location and the amplitude of these peaks. The number of points need not be known in advance although this knowledge improves the convergence of the iteration.

Adaptive extrapolation. To determine f(x), we start the iteration as in Fig. 2. Since f(x) consists of impulses, the function $g_n(x)$ exhibits maxima for sufficiently large n. When this is observed, a threshold level is set and only the region of $g_n(x)$ above that level is retained. The remaining part, below the threshold level, is replaced by zero yielding the function $f_{n+1}(x)$ shown in Fig. 3. The iteration step is completed as before. The threshold level is set low at first and is raised as the iteration progresses. As we show in the following illustrations, the object f(x) can be recovered even in the presence of considerable noise.

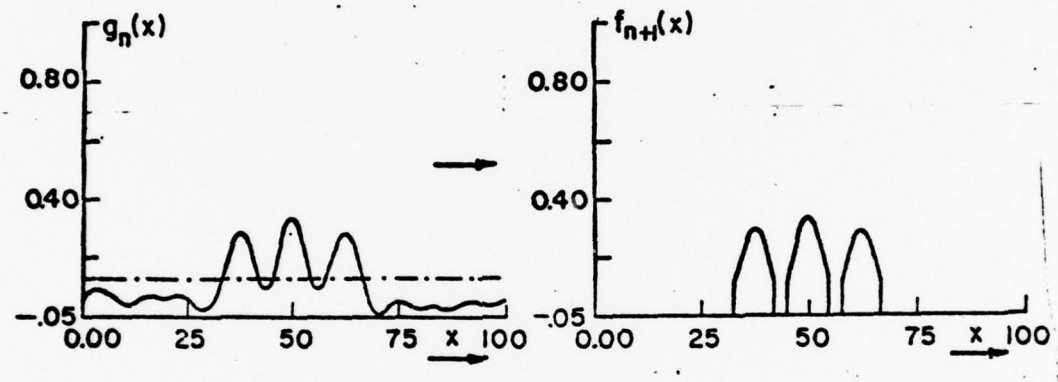


FIGURE 3 TRUNCATION OF $g_n(x)$ BELOW A THRESHOLD LEVEL YIELDING $f_{n+1}(x)$

Example 1. In Fig. 4, we show an object f(x) consisting of two points in a noisy background, and its image $g_0(x)$. It is evident from the figure that the two points are completely merged. The results of the iteration are shown for n=4 and n=30. As we see, the amplitudes and locations of the two points are recovered for n=30.

Example 2. In Fig. 5, we show an object consisting of 6 points in a noisy background and its image $g_0(x)$. At the 12th iteration, we observe 6 maxima but the amplitudes are not correct. The 6 points are essentially recovered at the 100th iteration.

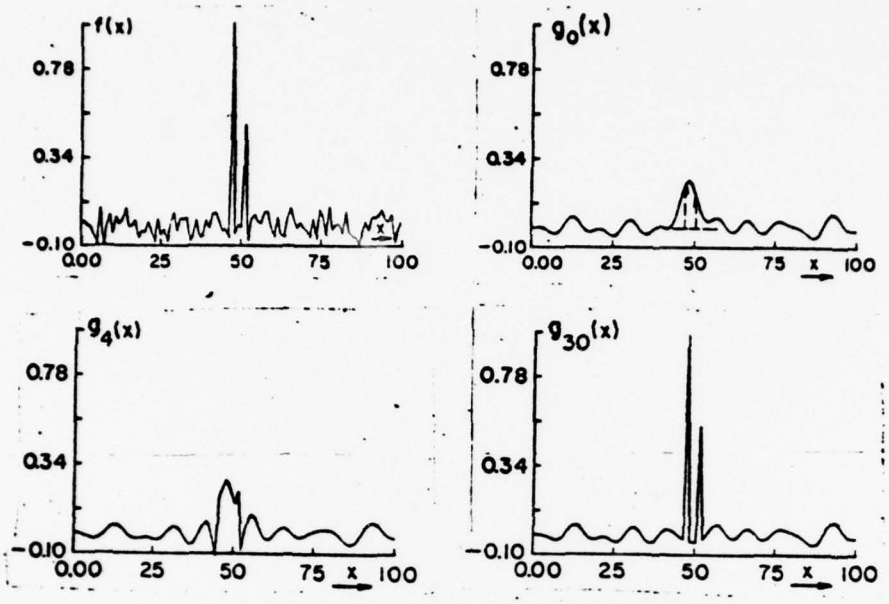


FIGURE 4 f(x): UNKNOWN OBJECT CONSISTING OF TWO POINTS CONTAMINATED BY NOISE. $g_0(x)$: KNOWN IMAGE, $g_4(x)$: RESULT OF 4TH ITERATION, $g_{30}(x)$: RESULT OF 30TH ITERATION

Example 3. In Fig. 6, we show the application of the method for the recovery of a two-dimensional object consisting of two points. The image is shown in Fig. 6b and the results of the iteration for n=4 and n=24 in Fig. 6c and d. At the 24th iteration the object is recovered completely.

2. Clutter elimination and object restoration

An object f(x) seen through a noisy medium yields an image

$$y(x) = f(x) + n(x)$$

We assume that the noise component n(x) has only low-frequency components and we seek to determine f(x). We can eliminate the noise n(x) by high-pass filtering. This can be done optically by blocking the center portion of the aperture of the viewing system. The resulting image $w_0(x)$ is free of noise but is distorted because it consists only of the high frequencies of f(x). To recover f(x) from

$w_0(x)$, we apply the preceding method suitably modified: Reversing the role of high and low frequencies, we can show that the iteration, starting from $w_0(x)$ yields f(x). The details are omitted.

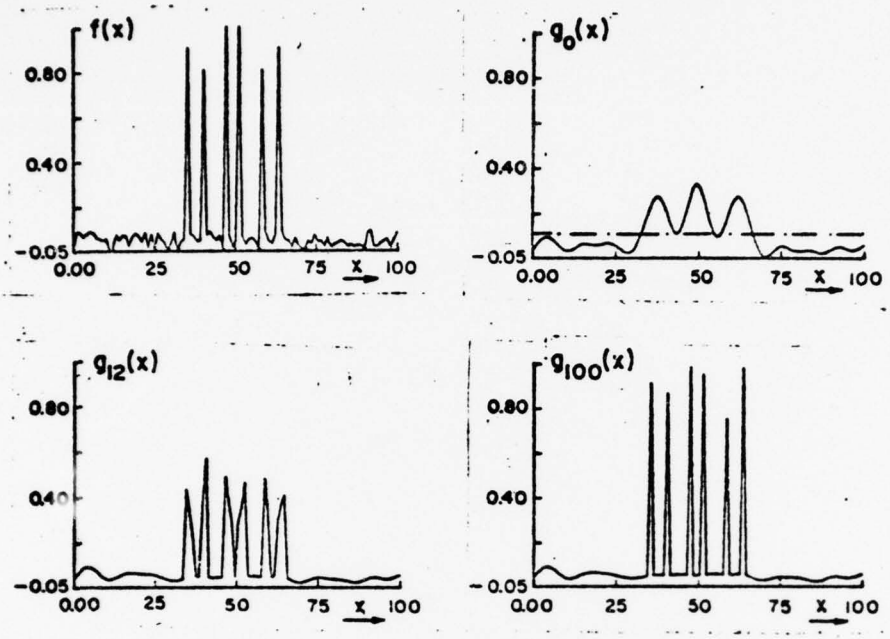


FIGURE 5 f(x): UNKNOWN OBJECT CONSISTING OF 6 POINTS CONTAMINATED BY NOISE, $g_0(x)$: KNOWN IMAGE, $g_{12}(x)$: RESULT OF 12TH ITERATION, $g_{100}(x)$: RESULT OF 100TH ITERATION

Example 4. In Fig. 7, we show the unknown object f(x) and the date y(x) containing the noise n(x). Filtering of y(x) yields the signal $w_0(x)$. As we see from the figure, the noise is eliminated but the signal f(x) is strongly distorted. To find f(x), we apply the iteration method. At the 10th step, we obtain the signal $w_{10}(x)$ shown in Fig.7. This is essentially identical to the unknown object f(x).

UNCLASSIFIED

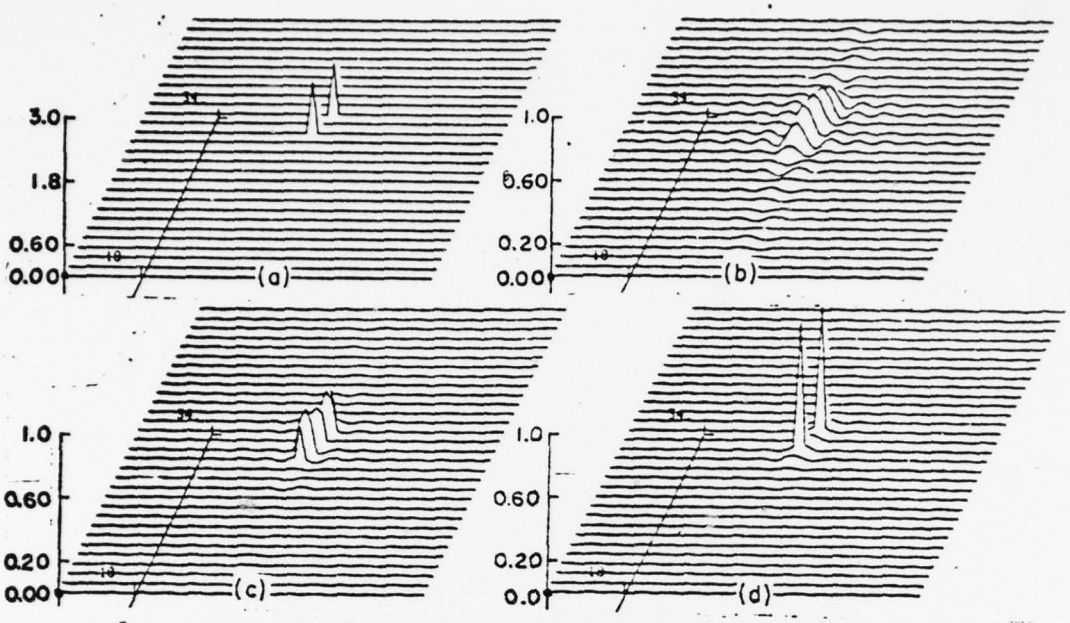


FIGURE 6 (a) UNKNOWN OBJECT, (b) KNOWN IMAGE
(c) 4TH ITERATION, (d) 24TH ITERATION

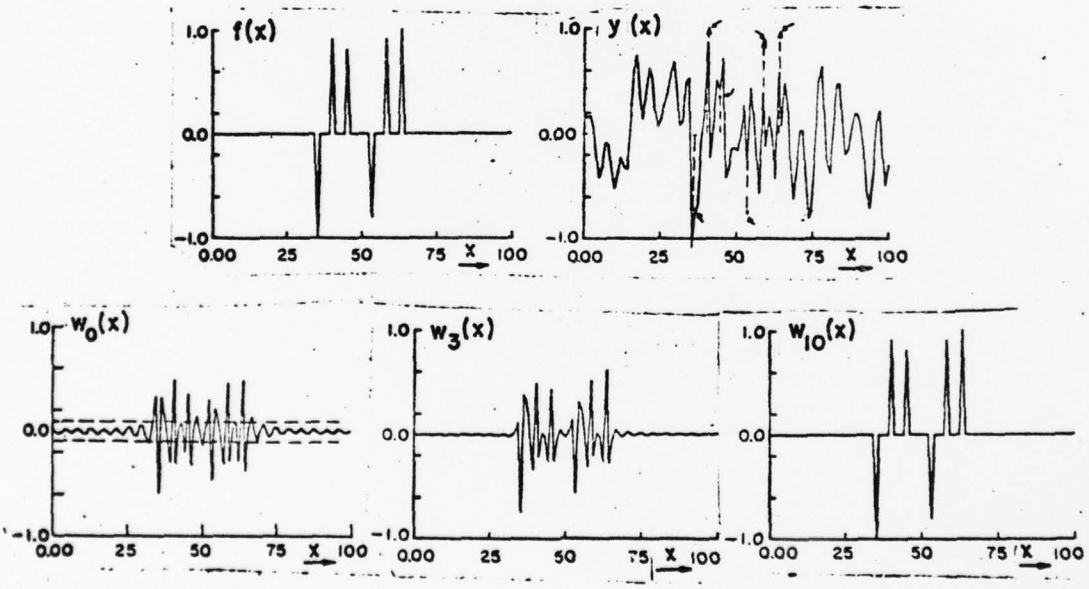


FIGURE 7 f(x): UNKNOWN OBJECT, y (x): IMAGE OF f(x) VIEWED THROUGH CLUTTER, $w_0(x)$: FILTER OUTPUT, $w_3(x)$: 3RD ITERATION, $w_{10}(x)$: 10TH ITERATION

References

1. A. Papoulis, "A New Algorithm in Spectral Analysis and Bandlimited Extrapolation" vol. CAS-22, no. 9, pp. 735-742, September 1975.

2. R. W. Gerchberg, "Super-Resolution through Error Energy Reduction" Opt. Acta, vol. 21, no. 9, pp. 709-720, 1974.

3. A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.

Section 3

ADAPTIVE EXTRAPOLATION AND HIDDEN PERIODICITIES

A. PAPOULIS and C. CHAMZAS

Department of Electrical Engineering
Polytechnic Institute of New York
Route 110, Farmingdale, N. Y. 11735

1. Introduction

An important problem in many applications is the determination of the frequency components of a signal

$$f(t) = \sum_{i=1}^{m} c_i e^{j\omega_i t} \tag{1}$$

in terms of the segment

$$w_1(t) = \begin{cases} f(t)+n(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{2}$$

of f(t) containing the noise component n(t). The signal f(t) is not known for every t for a variety of reasons:

The signal f(t) can be written as a sum of exponentials for a limited time only (Voice; non-stationary processes).

The available time of observation is limited (sun spots; weather trends).

Measurements are limited by instrument constraints (Michelson interferometer; band-limited channels).

The unknown frequencies ω_i and coefficients c_i can be determined simply with ordinary Fourier transforms if the time of observation 2T is large compared to all the periods $T_i=2\pi/\omega_i$. This is not, however, the case if T is of the order T_i particularly if the noise component of $w_1(t)$ is not negligible. In this paper, we present a method which, as we hope to

show, is reliable even in such extreme cases.

The method involves only FFT and it is based on earlier results dealing with the problem of extrapolating band-limited functions [1, 2]. We review for easy reference the relevant parts of these results.

2. Extrapolation of band-limited functions

Consider a function f(t) with Fourier transform F(ω) such that

$$F(\omega) = 0 \qquad |\omega| > \sigma \tag{3}$$

We form the function

$$w_1(t) = \begin{cases} f(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{4}$$

obtained by truncating f(t) as in Fig. 1. We shall determine f(t) in terms of $w_1(t)$ by numerical iteration.

First step. We compute the Fourier transform $W_1(\omega)$ of $w_1(t)$, form the function

$$F_1(\omega) = \begin{cases} W_1(\omega) & |\omega| < \sigma \\ 0 & |\omega| > \sigma \end{cases} \tag{5}$$

Compute the inverse transform $f_1(t)$ of $F_1(\omega)$, and form the function

$$w_2(t) = \begin{cases} w_1(t) = f(t) & |t| < T \\ f_1(t) & |t| > T \end{cases} \tag{6}$$

and its Fourier transform $W_2(\omega)$.

This completes the first step of the iteration (Fig. 1).

nth step. We form the function

$$F_n(\omega) = \begin{cases} W_n(\omega) & |\omega| < \sigma \\ 0 & |\omega| > \sigma \end{cases} \tag{7}$$

where $W_n(\omega)$ is the function obtained at the end of the preceding step. We compute the inverse transform $f_n(t)$ of $F_n(\omega)$, form the function

$$w_{n+1}(t) = \begin{cases} f(t) & |t| < T \\ f_n(t) & |t| > T \end{cases} \tag{8}$$

and compute its Fourier transform $W_{n+1}(\omega)$.

If f(t) is approximated by $f_n(t)$, the resulting mean-square error is given by

$$E_n = \int_{-\infty}^{\infty} \left[f(t)-f_n(t)\right]^2 dt = \frac{1}{2\pi} \int_{-\sigma}^{\sigma} \left|F(\omega)-F_n(\omega)\right|^2 d\omega \tag{9}$$

We maintain that this error decreases twice at each iteration step. Indeed,

$$E_n = \int_{|t|<T} \left[f(t)-f_n(t)\right]^2 dt + \int_{|t|>T} \left[f(t)-f_n(t)\right]^2 dt$$

But [see (8) and (7)]

$$\int_{|t|>T} \left[f(t)-f_n(t)\right]^2 dt = \int_{-\infty}^{\infty} \left[f(t)-w_{n+1}(t)\right]^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|F(\omega)-W_{n+1}(\omega)\right|^2 d\omega$$

$$= \frac{1}{2\pi} \int_{|\omega|>\sigma} \left|F(\omega)-W_{n+1}(\omega)\right|^2 d\omega + \frac{1}{2\pi} \int_{-\sigma}^{\sigma} \left|F(\omega)-F_{n+1}(\omega)\right|^2 d\omega$$

Hence,

$$E_n - E_{n+1}$$

$$= \int_{|t|<T} \left[f(t)-f_n(t)\right]^2 dt + \frac{1}{2\pi} \int_{|\omega|>\sigma} \left|F(\omega)-W_{n+1}(\omega)\right|^2 d\omega \qquad (10)$$

In [1] and [2] we show that $f_n(t) \to f(t)$ as $n \to \infty$. This is not true if the given segment $w_1(t)$ of $f(t)$ is noisy as in (2). In this case, a satisfactory estimate of $f(t)$ can be found by early termination of the iteration. [2]

Note. From (10) it follows that the mean-square error E_n is a monoton decreasing function and since it is positive, it tends to a limit. This does not prove the convergence of (9) because the limit need not be zero. It shows, however, that

$$E_n - E_{n+1} \to 0 \qquad n \to \infty$$

Hence,

$$\int_{|t|<T} \left[f(t)-f_n(t)\right]^2 dt \to 0 \qquad n \to \infty \qquad (11)$$

Although the functions $f(t)$ and $f_n(t)$ are band-limited, (11) does not imply that $f(t) \to f_n(t)$ because there is no lower bound on the energy concentration of band-limited functions in a finite interval [1, 3]. For example, the prolated spheroidal functions $\varphi_n(t)$ are band-limited, their energy equals one but their energy concentration in the interval $(-T, T)$ tends to zero as $n \to \infty$. This is the case because the eigenvalues λ_n of the underlying integral equation tend to zero as $n \to \infty$.

We mention without elaboration that, in the discrete version of the problem, the convergence of the iteration can be deduced from (11) under suitable conditions. The reason is that the corresponding eigenvalues are finitely many, therefore, they have a positive minimum.

3. Adaptive extrapolation

The preceding method was based on the assumption that the unknown function f(t) is band-limited. This information was used to reduce the error in the estimation of f(t) twice at each iteration step. The speed of iteration can be increased and the effects of noise can be reduced if additional a priori information about f(t) is available. Suppose, for example, that the size of the band of $F(\omega)$ is known but its precise location is unknown. We then choose a constant σ sufficiently large for $F(\omega)$ to vanish outside the integral $(-\sigma, \sigma)$ and proceed as in Sec. 2. As the iteration progresses, the form of $W_n(\omega)$ suggests appropriate reduction of the assumed band of f(t).

The adaptive extrapolation method is particularly effective if f(t) is a sum of exponentials as in (1). In this case, $F(\omega)$ consists of impulses (lines) as in Fig. 2:

$$F(\omega) = 2\pi \sum_{i=1}^{m} c_i \, \delta(\omega - \omega_i) \tag{12}$$

and our problem is to determine their locations ω_i and amplitudes c_i in terms of the known segment $w_n(t)$ of f(t).

To solve this problem, we select a constant σ larger than the largest possible value of ω_i and we proceed with the iteration until $W_n(\omega)$ takes significant values only in a subset B_n of the band $(-\sigma, \sigma)$ of f(t) (Fig. 3). This suggests that the unknown frequencies are in B_n. When this is observed, the function $F_n(\omega)$ of the nth iteration step is obtained from the following modification of (7):

$$F_n(\omega) = \begin{cases} W_n(\omega) & \omega \in B_n \\ 0 & \omega \in \bar{B}_n \end{cases} \tag{13}$$

(Fig. 3) In the above, B_n is the set of points such that $W_n(\omega)$ exceeds a threshold level ϵ_n.

$$|W_n(\omega)| \begin{cases} > \epsilon_n & \omega \in B_n \\ < \epsilon_n & \omega \in \bar{B}_n \end{cases} \tag{14}$$

and $\bar{B}_n$ its complement . The process is repeated until $W_n(\omega)$ approach the unknown spectrum. This can be checked by comparing the inverse $w_n(t)$ of $W_n(\omega)$ with the known segment of f(t).

Notes. As the following examples show, the unknown frequencies can be found even if the data are noisy and the constant T is smaller than the smallest period T_i .

The choice of the threshold level ϵ_n is dictated by two conflicting factors: For a speedy convergence and reduction of the noise component, ϵ_n must be large. It must be sufficiently small so that all frequency components of $F(\omega)$ are in the set B_n. Thus ϵ_n is small at first and it increases as the iteration progresses.

The accuracy of the method depends on the number m of the unknown components and their relative locations and amplitudes. If some components are small compared to the maximum c_i, it is possible that they could be lost. However, if the noise is sufficiently small, they can be recovered by substracting the significant components and repeating the iteration.

A priori knowledge of the number m of the unknown frequencies is useful but not essential.

If, at the nth iteration step, all frequency components of f(t) are in the set B_n, then the resulting mean-square error reduction is given by (10) mutatis mutandis.

4. Illustrations

We conclude with a digital implementation of the above method. The computations were performed with a PDP 11 minicomputer (single precision) and the FFT size was N= 256. The known segment $w_1(t)$ of f(t) contains the first 30 points. In figure 4 the unknown signal consists of 3 cosine waves. Determination of their amplitudes and frequencies is not apparent from the Fourier transform $W_1(\omega)$, figure 4b, of the given segment. The resolution is improved considerably within a few steps of the iteration. In the 12th step the three cosine terms have been revealed, figure 4c, and in the 63rd step complete estimation of frequencies and amplitudes have been achieved. (figure 4d).

In the next example, figure 5, the same signal is used but we added 10% white noise uniformly distributed. The signal is again recovered completely in 100 steps. In figure 6 the unknown signal consisting of two exponentials has been corrupted by 20% white noise. To recover the signal the iteration needs 30 steps.

Thus, as we see from these illustrations the unknown frequencies can be found even if the data are noisy and the constant T is smaller than the smallest period in f(t).

References

[1] A. Papoulis, Signal Analysis, McGraw-Hill, New York

[2] _____, "A New Algorithm in Spectral Analysis and Band-limited Extrapolation", IEEE Trans. Circuits Syst., vol. CAS-22, no. 9, pp. 735-742, September 1975.

[3] D. Slepian, H.J. Landau, and H.O. Pollack, "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty Principle I and II", Bell Syst. Tech. J., vol. 40, no. 1, 1961.

[4] A. Papoulis and M. Bertran, Digital Filtering and Prolate Functions", IEEE Trans. Circuit Theory, vol. CT-19, pp. 674-681, November 1972.

Figures

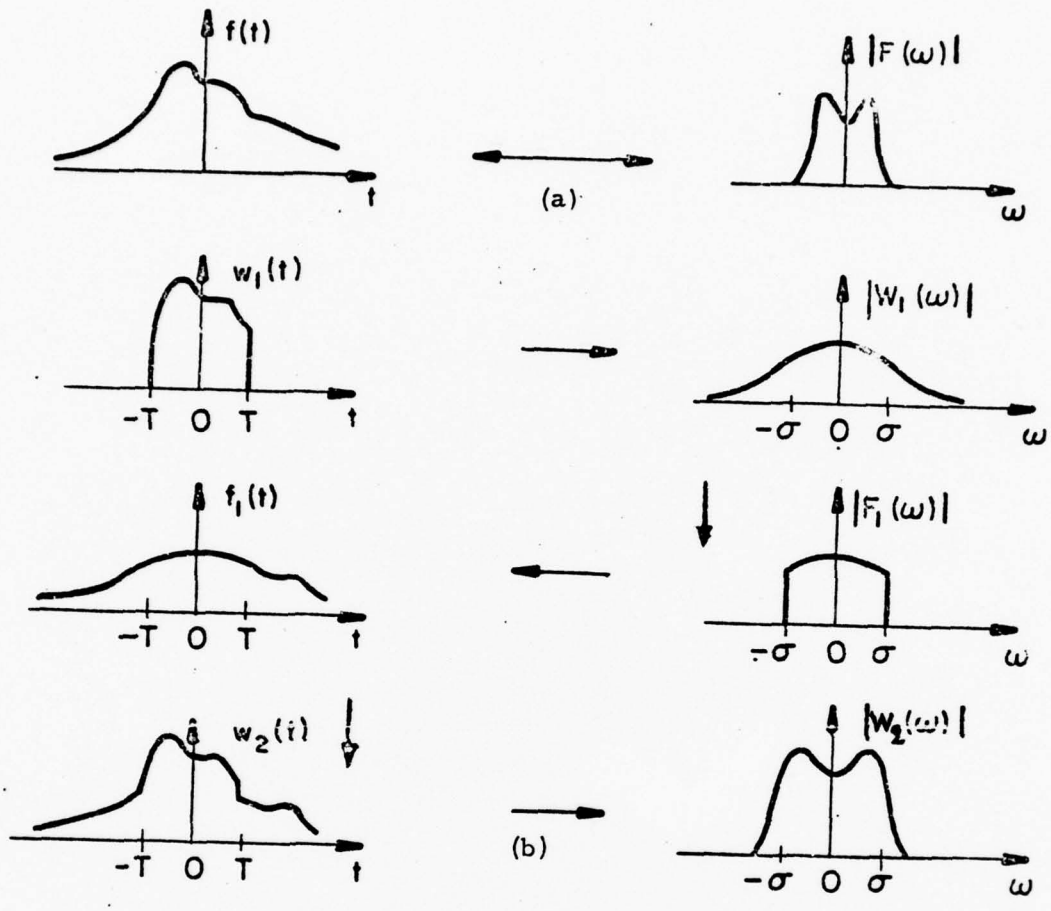


FIGURE 1. (a) The unknown signal f(t) and its Fourier transform F(ω)
(b) First iteration step starting with known segment $w_1(t)$.

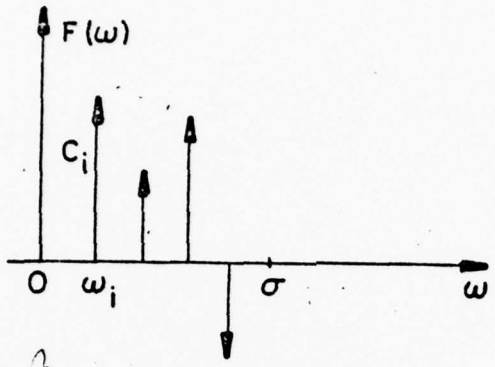


FIGURE 2. Fourier transform of the unknown signal.

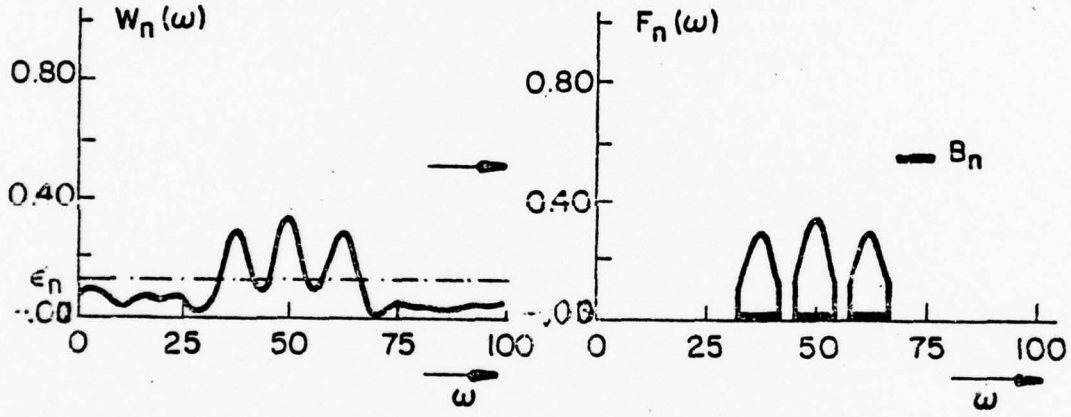


FIGURE 3. Truncation of $W_n(\omega)$ below a threshold level ϵ_n yielding $F_n(\omega)$.

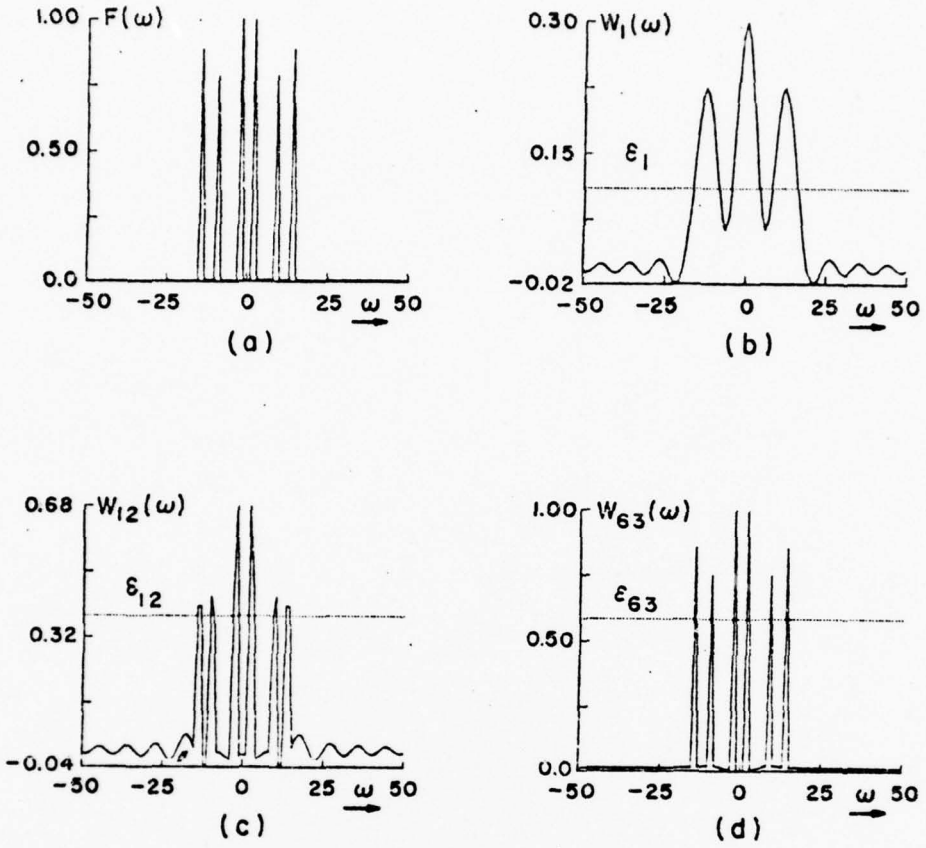


FIGURE 4. (a) $F(\omega)$: The Fourier transform of the noiseless unknown signal.
(b) $W_1(\omega)$: The Fourier transform of the given segment $W_1(t)$.
(c), (d): The result of the 12th and 63rd iteration (ϵ_n indicates the threshold level at the nth step)

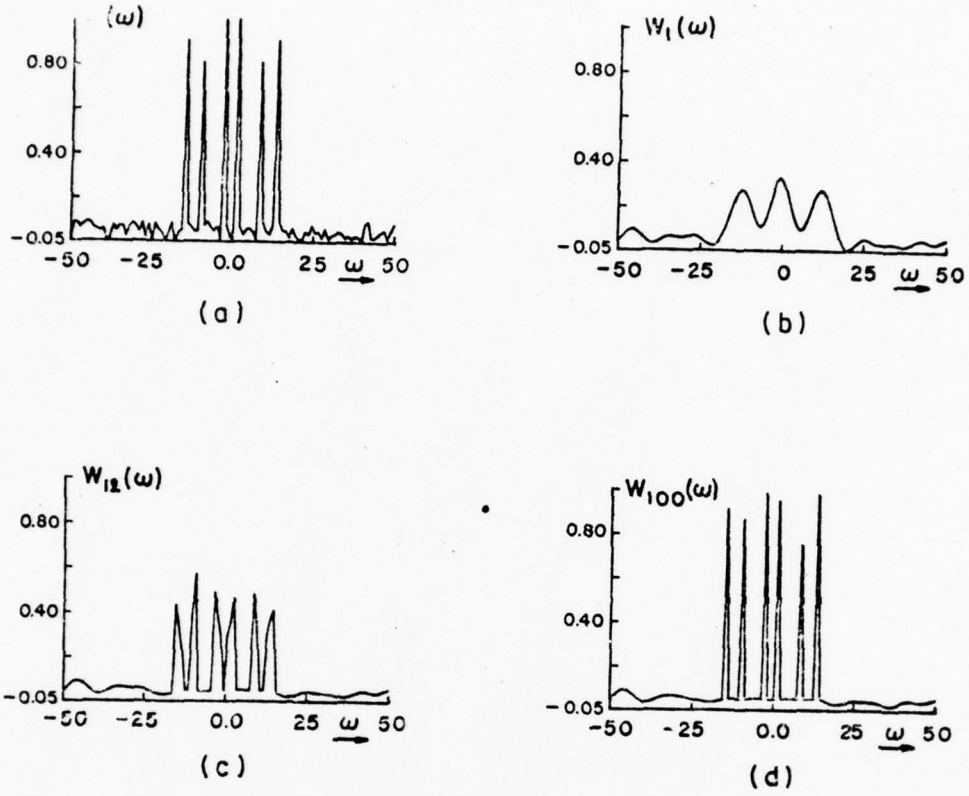


FIGURE 5. (a) $F(\omega)$: The Fourier transform of the unknown signal $f(t)$.
(b) $W_1(\omega)$: The Fourier transform of the given segment $w_1(t)$.
(c), (d): The result of the 12th and 100th iteration.

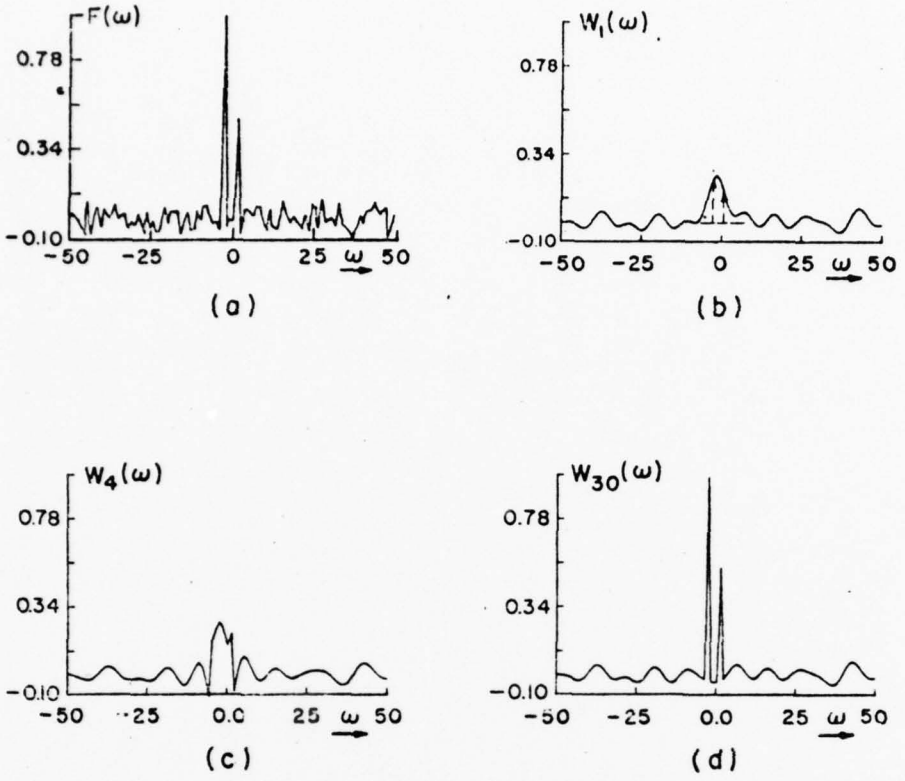


FIGURE 6. (a) $F(\omega)$: Unknown signal consisting of two impulses contaminated by noise.

(b) $W_1(\omega)$: Fourier transform of the known segment $w_1(t)$.

(c),(d): The result of the 4th and 30th iteration.

ULTRASONIC IMAGING 1, 121-135 (1979)

Section 4

IMPROVEMENT OF RANGE RESOLUTION BY SPECTRAL EXTRAPOLATION

A. Papoulis[1] and C. Chamzas

Polytechnic Institute of New York
Department of Electrical Engineering
Farmingdale, New York 11735

Under various simplifying assumptions, the reflected signal y(t) in the interrogation of a substance by an ultrasonic wave is a convolution of the transmitted signal x(t) with a function h(t) that is related to the reflection coefficient of the medium in the direction of propagation. The function h(t) can, in principle, be determined by deconvolution. However, since the band B of the spectrum $X(\omega)$ of x(t) is finite, the frequency components of h(t) outside B cannot be found reliably. In this paper, a method is presented for extrapolating $X(\omega)$ beyond B. The resulting increase in resolution is limited only by the level of noise. The method is particularly effective if h(t) is a sum of impulses.

Key words: Deconvolution; diagnostics; echoes; extrapolation; resolution; ultrasonic.

1. Introduction

Ultrasonic waves are used in metallurgy, in medicine, and in other areas to determine the structure of a medium. The medium is interrogated by a narrow beam (fig. 1) and the reflected signal y(t) is used to determine various properties of the medium, in particular, the location of its surface of discontinuity. Under various simplifying assumptions, the signal y(t) can be expressed as a convolution integral:

$$y(t) = \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau \qquad (1)$$

where x(t) is the transmitted signal (fig. 2a) and h(t) is a function related to the reflection coefficients of the medium in the direction of propagation. The variable t is proportional to the distance along the beam. The assumptions leading to eq. (1) and the relationship between h(t) and the parameters of the medium will not be considered here.

If the medium consists of homogeneous layers, then h(t) is a sum of impulses as in figure 2b:

---

[1]Address all correspondence to A. Papoulis.

0161-7346/79/020121-15$02.00/0
Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.

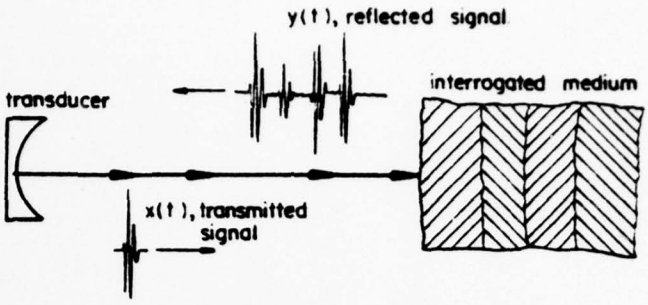


Fig. 1 Schematic of an ultrasonic system

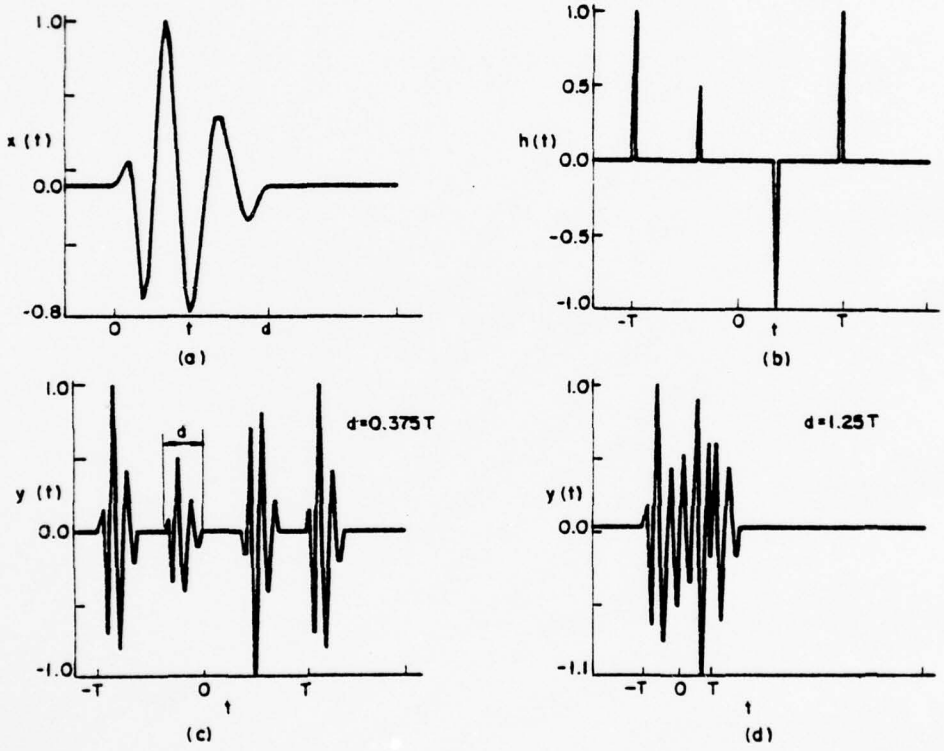


Fig. 2 (a) Transmitted signal x(t)
(b) Impulse response h(t) of idealized medium
(c) Reflected signal y(t) for d = 0.375T
(d) Reflected signal y(t) for d = 1.25T

$$h(t) = \sum_i c_i \, \delta(t-t_i) \qquad (2)$$

where the points t_i correspond to the locations of the separation surfaces, and the coefficients c_i are related to the reflection coefficients at these surfaces. In this case, eq. (1) yields

$$y(t) = \sum_i c_i \, x(t-t_i) \qquad (3)$$

and the problem is to determine the unknowns c_i and t_i in terms of the observed signal y(t).

If x(t) is a pulse whose duration d is smaller than the minimum distance $t_i - t_j$ between neighboring impulses, then the various terms in eq. (3) do not overlap (fig. 2c), hence, the constants c_i and t_i can be readily found. This is not the case, however, if d is larger than $t_i - t_j$ (fig. 2d). The resolution of the system, i.e., the smallest distance between reflecting surfaces that can be detected without special processing is thus, proportional to the duration d of x(t). To improve the resolution, we must decrease d, or, equivalently, we must design a transducer with large bandwidth. In this paper, we present a method for increasing the resolution of a given system by numerical processing of the signal y(t).

In principle, h(t) can be determined simply in terms of x(t) and y(t) by deconvolution. Indeed, taking Fourier transforms of both sides of eq. (1), we conclude from the convolution theorem that

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \qquad (4)$$

This yields the Fourier transform $H(\omega)$ of the unknown h(t) in terms of the Fourier transforms $X(\omega)$ and $Y(\omega)$ of the signals x(t) and y(t) respectively. To find h(t), it suffices, therefore, to compute the inverse transform of the ratio $Y(\omega)/X(\omega)$.

However, the resulting values of $H(\omega)$ in the region of the frequency axis where $X(\omega)$ is of the order of the background noise are not reliable. For a typical transducer, $X(\omega)$ is a curve as in figure 3 taking significant values in the band (ω_1, ω_2) only; hence, $H(\omega)$ can be determined reliably only in this band. Our problem, therefore, is to extrapolate $H(\omega)$ for $\omega < \omega_1$ and $\omega > \omega_2$. We shall do so under the assumption that h(t) is a sum of impulses as in eq. (2). We note that, since $X(\omega)$ is not strictly bandlimited, the frequencies ω_1 and ω_2 are rather arbitrary. If they are so chosen that $X(\omega)$ is significantly larger than the noise level, then the error in the estimation of $H(\omega)$ by the ratio $Y(\omega)/X(\omega)$ is small. However, the extrapolation problem is then more difficult because the length $\omega_2 - \omega_1$ of the resulting interval in which $H(\omega)$ is assumed known is small.

We note, finally, that in the presence of noise, the best estimate of $H(\omega)$ in the interval (ω_1, ω_2) is not the ratio $Y(\omega)/X(\omega)$. It

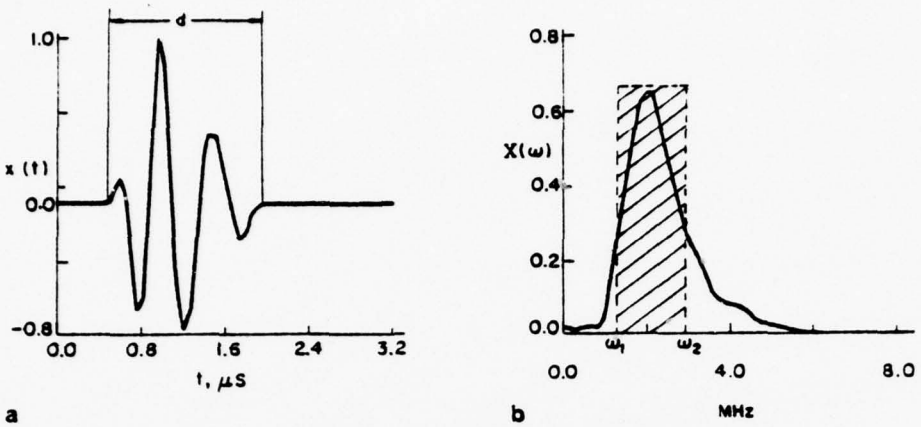


Fig. 3 (a) Typical transmitted signal x(t)
(b) Spectrum of x(t).

can be shown [4] that the estimate of $H(\omega)$ in the vicinity of the low values of $X(\omega)$ can be improved if the properties of the noise are known.

The proposed method of determining h(t) in terms of only a segment of its Fourier transform $H(\omega)$, is a nonlinear adaptive modification of an extrapolation method developed in reference [2]. In the next section, we review briefly the relevant parts of this paper.

2. Spectral extrapolation of time-limited signals

Consider a function h(t) with Fourier transform $H(\omega)$. We assume that only the segment

$$W_1(\omega) = \begin{cases} H(\omega) & \omega \in B \\ 0 & \omega \in \bar{B} \end{cases} \tag{5}$$

of $H(\omega)$ is known, where B is a certain band on the frequency axis and $\bar{B}$ is its complement. In general, h(t) cannot be determined in terms of $W_1(\omega)$. However, if h(t) is time-limited, i.e., if

$$h(t) = 0 \quad \text{for} \quad |t| > T \tag{6}$$

then [2] it can be uniquely determined in terms of $W_1(\omega)$.

In the ultrasonics problem, the duration 2T of h(t) is proportional to the thickness of the medium, and the band B is the interval (ω_1, ω_2) in which $X(\omega)$ takes significant values. The function $W_1(\omega)$ equals the ratio $Y(\omega)/X(\omega)$ for every ω in B.

Using a numerical iteration involving only FFT's, we shall determine $H(\omega)$ for every ω.

First step. We compute the inverse transform

$$w_1(t) = \frac{1}{2\pi} \int_B W_1(\omega) e^{j\omega t} \, d\omega \tag{7}$$

of $W_1(\omega)$. We form the function

$$h_1(t) = \begin{cases} w_1(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{8}$$

obtained by truncating $w_1(t)$ for $|t| > T$ (fig. 4). We compute the Fourier transform

$$H_1(\omega) = \int_{-T}^{T} h_1(t) e^{-j\omega t} \, dt \tag{9}$$

of $h_1(t)$, and form the function

$$W_2(\omega) = \begin{cases} W_1(\omega) = H(\omega) & \omega \in B \\ H_1(\omega) & \omega \in \bar{B} \end{cases} \tag{10}$$

nth iteration. We compute the inverse transform

$$w_n(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_n(\omega) e^{jn\omega t} \, d\omega \tag{11}$$

of the function $W_n(\omega)$ determined at the end of the n-1 iteration. We form the function

$$h_n(t) = \begin{cases} w_n(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{12}$$

obtained by truncating $w_n(t)$ for $|t| > T$ (fig. 5). We compute the Fourier transform

$$H_n(\omega) = \int_{-T}^{T} h_n(t) e^{-j\omega t} \, dt \tag{13}$$

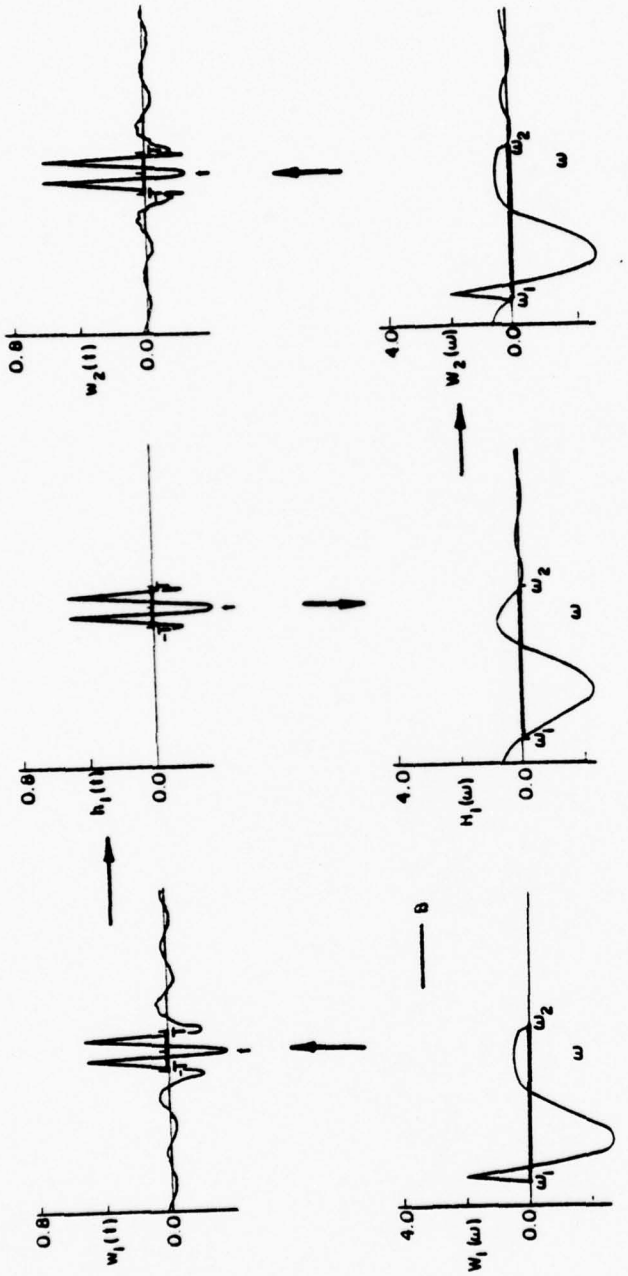


Fig. 4 Description of first iteration Upper Level: time signals; lower level; their spectra. Arrows indicate the order of the various computations starting with $W_1(\omega)$.

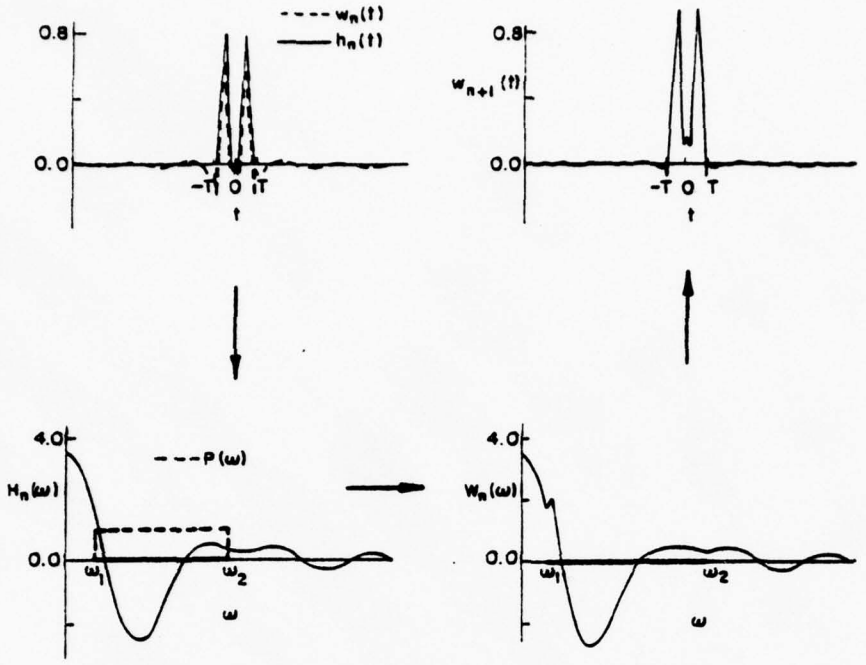


Fig. 5 Description of nth iteration

of $h_n(t)$, and form the function

$$W_{n+1}(\omega) = P(\omega)H(\omega) + [1 - P(\omega)]H_n(\omega) \qquad (14)$$

where $P(\omega)$ is a rectangular window:

$$P(\omega) = \begin{cases} 1 & \omega \in B \\ 0 & \omega \in \bar{B} \end{cases} \qquad (15)$$

as in figure 5.

We have shown [2, 3] that, in the absence of noise, the sequence $h_n(t)$ so formed tends to the unknown signal $h(t)$ as n tends to infinity. The effects of noise are considered in reference [2].

In eq. (14), we used for the term $P(\omega)$ a rectangular window. This, however, is not necessary [5, 6]. It can be shown that the iteration converges for any $P(\omega)$ provided that

$$0 \leq P(\omega) \leq 1 \quad . \qquad (16)$$

We mention also that the speed of convergence increases if eq. (14) is replaced by the equation

$$W_{n+1}(\omega) = \begin{cases} H(\omega) & \omega \in B \\ A_n H_n(\omega) & \omega \in \bar{B} \end{cases} \tag{17}$$

where A_n is a suitable factor.

3. Adaptive extrapolation

We now assume that h(t) is different from zero not in the entire interval (-T, T) but only in a subset b of this interval. Thus, h(t) = 0 for every t in the complement $\bar{b}$ of b. In the problem under consideration, b is a set consisting of a finite number of points t_i [see eq. (3)]. If the set b is known, then we determine the function $h_n(t)$ of the nth iteration such that

$$h_n(t) = \begin{cases} w_n(t) & t \in b \\ 0 & t \in \bar{b} \end{cases} \tag{18}$$

This is a modified form of eq. (12) and it results in a reduction of the effect of noise and an increase in the speed of convergence.

In our problem, the set b is not known. In fact, our objective is to find it. However, the information that it consists of a finite number of points can be used to reduce the size of the truncation interval. This is done as follows:

We start the iteration as in Sec. 2. As the iteration progresses, the function $w_n(t)$ takes significant values only in a subset of the interval (-T, T) because $w_n(t)$ tends to a sum of impulses as $n \to \infty$. This set is denoted by b_n and is so defined that

$$w_n(t) \geq \epsilon_n \qquad t \in b_n$$

$$w_n(t) < \epsilon_n \qquad t \in \bar{b}_n \tag{19}$$

where ϵ_n is a suitable threshold level. It is desirable that ϵ_n be as large as possible subject to the condition that all points t_i be included in the corresponding set b_n. We have found from a number of numerical calculations that a reasonable choice is the minimum

$$\epsilon_n = \min\{|w_{n-1}(t)|\} \qquad t \in b_{n-1} \tag{20}$$

of the signal $w_{n-1}(t)$ in the set b_{n-1} of the preceding iteration step.

We next form the function $h_n(t)$ by truncating $w_n(t)$ in the complement $\bar{b}_n$ (fig. 6):

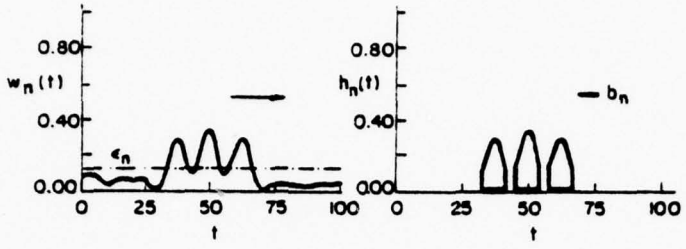


Fig. 6 Adaptive reduction of truncation set; ε_n; threshold level.

$$h_n(t) = \begin{cases} w_n(t) & t \in b_n \\ 0 & t \in \bar{b}_n \end{cases}, \tag{21}$$

we compute its Fourier transform $H_n(\omega)$, and continue as in eq. (14). A nested sequence of sets b_n

$$b_{n-1} \supseteq b_n \supseteq b_{n+1} \supseteq \cdots$$

results which for reasonably small noise levels converges to the set of points t_i. If, however, some of the coefficients c_i of h(t) [see eq. (2)] are small, the corresponding points t_i might not be in all the sets b_n. This depends on the level of noise but precise conditions cannot be easily established in general. We are in the process of determining sufficient conditions for the complete recovery of h(t) for various special cases involving a small number of points t_i.

The following modification of the method results in a reduction of the effects of noise and round-off errors: the set b_n does not necessarily change at each iteration step, i. e., it is possible that $b_{n-1} = b_n$. We continue the iteration until b_n is a proper subset of b_{n-1}. When this is observed, we replace the signal $w_n(t)$ by the signal $w_1(t)$ of the first iteration step, and start again where now the interval (-T, T) is replaced by the set b_n. This reduces the total energy of the noise, and it eliminates the accumulation of errors from the preceding steps.

4. Numerical illustration

We shall use the method to determine the location and the reflection coefficients of a synthetic medium consisting of six homogeneous layers. The corresponding impulse response is an impulse train consisting of six unequal impulses as in figure 7. The observed signal is the sum

$$z(t) = y(t) + n(t) \tag{22}$$

shown in figure 8a. The component y(t) is obtained by convolving h(t) with the signal x(t) of fig. 8b.

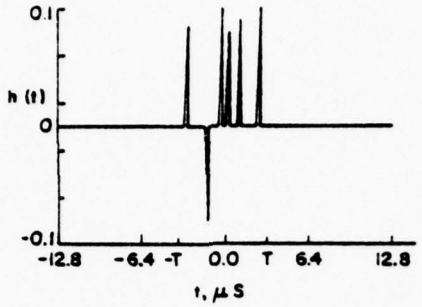


Fig. 7 Unknown impulse response

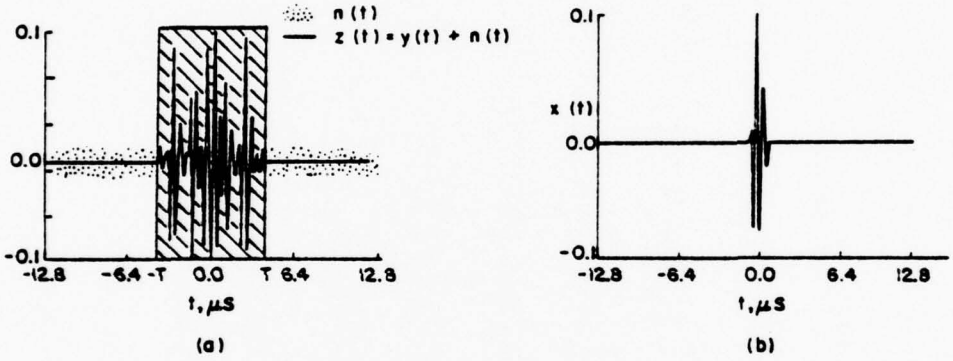


Fig. 8 (a) Observed signal z(t) = y(t) + n(t) consisting of reflected signal y(t) and noise component n(t); (b) Transmitted signal x(t)

$$y(t) = x(t) * h(t) = \sum_{i=1}^{6} c_i \, x(t-t_i) \tag{23}$$

The component n(t) is white noise with uniform distribution, zero mean and standard derivation $\sigma = 0.05$. The signal x(t) is the output of a physical transducer (see also fig. 3), and it determines the scale of the time axis. All other signals are computer generated.

Our objective is to determine numerically h(t) in terms of z(t) and x(t). The computations are carried out digitally on a single-precision minicomputer using as sampling interval $\delta = 0.1 \mu s$. It is assumed that the unknown impulses are in the

interval (-T, T) where T = 4μs. This interval is generally known from the physical description of the problem (outer surfaces of interrogated medium), or it can be determined from the observed data. In our case, the interval 2T contains 80 sample points. All Fourier transforms are computed with an FFT size N = 256. The total processing interval is, thus, Nδ = 25.6μs, that is, about three times the truncation interval 2T. The interval Nδ must be sufficiently large to avoid aliasing errors. The sampling interval δ determines the limit of resolution. The parameters N and δ are not critical.

Since it is known that the segment of z(t) outside the interval (-T, T) is noise, it is truncated prior to any processing.

We shall first attempt to determine h(t) by direct deconvolution. It will become evident from the computations that the results do not give an adequate estimate of h(t). For this purpose, we compute the Fourier transforms Z(ω) and X(ω) of z(t) and x(t) respectively and form the ratio

$$H_a(\omega) = \frac{Z(\omega)}{X(\omega)} \tag{24}$$

(fig. 9a). The inverse transform $h_a(t)$ of $H_a(\omega)$ is the result of direct deconvolution (fig. 9b). The unknown impulses are not recognizable.

The effect of the noise is most pronounced outside the interval (ω_1, ω_2). To reduce it, we truncate $H_a(\omega)$ outside this interval. The resulting function $W_1(\omega)$ and its inverse $w_1(t)$ are shown in figure 10. Again, $w_1(t)$ is not a satisfactory estimator of h(t) because the frequencies outside the interval (ω_1, ω_2) are eliminated.

To determine the missing frequencies of H(ω), we apply the iteration described in Sec. 3 starting with the function $W_1(\omega)$ of figure 10. The results of the iteration are shown in figure 11a for various values of n. In figure 11b, we show in detail the function $w_n(t)$ for n = 40. In figure 12 , we plot the normalized mean-square error

$$e_n = \frac{1}{E} \sum_{k=0}^{255} [h(k\delta) - w_n(k\delta)]^2 \qquad E = \sum_{k=0}^{255} h^2(k\delta) \tag{25}$$

as a function of n. As we see from these results, the unknown h(t) is essentially fully recovered for n = 40. In fact, the locations t_i of the six impulses are located exactly (see figure 7).

We next repeat the iteration using as the window not the pulse P(ω) of eq. (15), but the function

$$P_1(\omega) = \frac{1}{k} |X(\omega)| \tag{26}$$

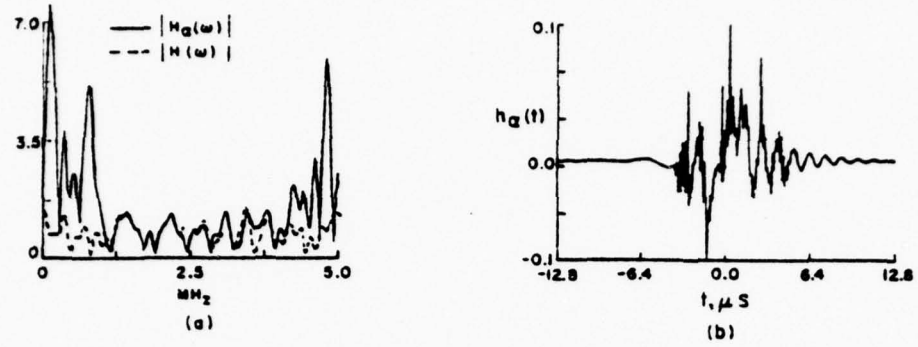


Fig. 9 Estimate of h(t) obtained through direct deconvolution
(a) Ratio $H_a(\omega) = Z(\omega)/X(\omega)$ of spectra of z(t) and x(t)
(b) Inverse transform $h_a(t)$ of $H_a(\omega)$.

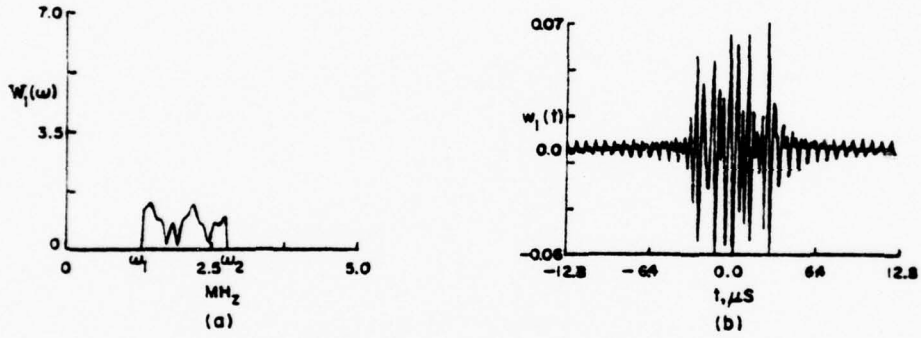


Fig. 10 Estimate of h(t) obtained through deconvolution and truncation (a) Reliable segment $W_1(\omega)$ of $H_a(\omega)$
(b) Inverse transform $w_1(t)$ of $W_1(\omega)$

where k equals the maximum of $|X(\omega)|$. This function is between zero and one and is chosen because it favors the frequencies for which $|X(\omega)|$ is large, i. e., the portion of the frequency axis where $H_a(\omega)$ is reliably determined [see eq. (24)]. The window $P_1(\omega)$, however, has the drawback that it distorts the component of $Z(\omega)$ due to the useful signal y(t).

In figure 13a, we show the results of the iteration. The function $w_{40}(t)$ is shown in figure 13b and the normalized mean-square error e_n in figure 12 . As we see from those figures, the error is reduced rapidly at the first 15 steps, but it re-

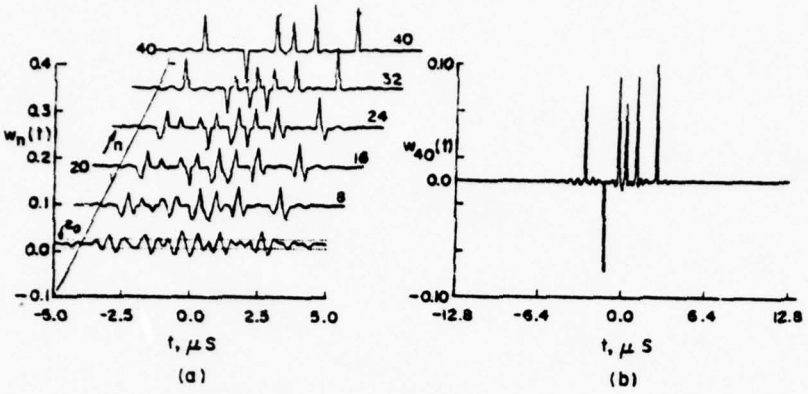


Fig. 11 (a) Results of adaptive iteration with rectangular window for n = 8, 16, 24, 32, and 40. (b) Detailed plot of $w_n(t)$ for n = 40.

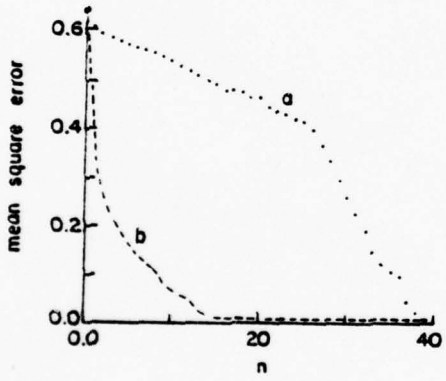


Fig. 12 Mean-square error

$$e_n = \frac{1}{E} \sum_{k=0}^{N-1} [h(k\delta) - w_n(k\delta)]^2$$

a: With rectangular window. b: With window $P_1(\omega)$ proportional to $|X(\omega)|$

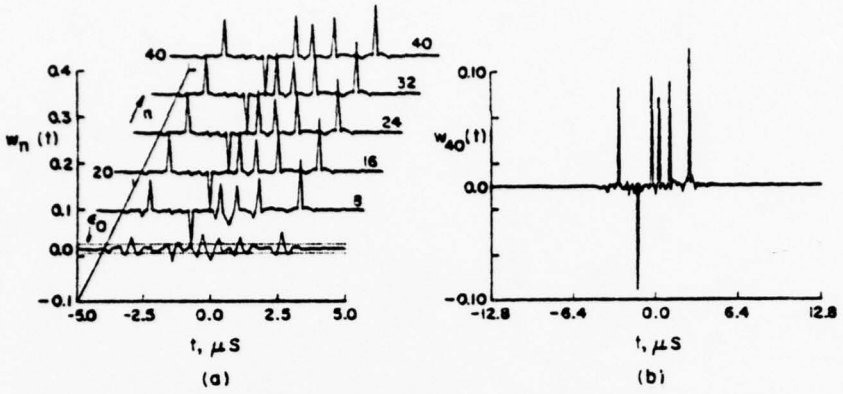


Fig. 13 (a) Results of adaptive iteration with window $P_1(\omega)$ for n = 8, 16, 24, 32, 40. (b) Detailed plot for n = 40.

mains essentially constant for $n > 20$. The reason is that in the early stages of the iteration, the distortion due to the noise is rapidly reduced, and what remains is the distortion of the useful component of $Z(\omega)$ due to the fact that $P_1(\omega)$ is not constant in truncation interval. It seems, therefore, that a combination of the two windows leads to better results: We start with the window $P_1(\omega)$ until the set b_n of the iteration is only a small part of the interval $(-T, T)$. We then continue with the rectangular window of eq. (15).

We conclude with an observation concerning the meaning of the term "resolution". This term is used extensively in several areas, and it has been given various definitions. These definitions, however, lead to a measure of resolution that is of the order of the duration d of the transmitted signal x(t), or, equivalently, inversely proportional to the bandwidth of x(t). (In optics, the resulting number is inversely proportional to the diameter of the aperture.) The measure so described is adequate if we rely on the direct observation of the data for the determination of the minimum spacing between impulses (point objects, in optics). It is not, however, useful if the data are subjected to elaborate processing. As we have seen, if the data contain no noise, then we can separate two or more arbitrarily close points no matter how large d is. The resolution is limited only by the size of the sampling interval δ. This is not so if noise is present; however, in this case also the limitation is not simply the size of d, but it depends on the level of the noise and on the form of the unknown signal h(t).

Acknowledgement

This work was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research, under Contract N00014-76C-0144.

References

[1] Papoulis, A., The Fourier Integral and Its Applications (McGraw-Hill, New York 1962).

[2] Papoulis, A., A New Algorithm in Spectral Analysis and Band-limited Extrapolation, in IEEE Trans. Circuits Syst., Vol. CAS-22, (9), pp. 735-742, September 1975.

[3] Papoulis, A., Signal Analysis (McGraw-Hill, New York, 1977).

[4] Papoulis, A., The Problem of Transmission Zeros in Deconvolution, in IEEE Trans. on Information Theory, Vol. IT-24, (1) pp. 126-128, January 1978.

[5] Papoulis, A., Ultrasonics in Medicine-Recent Advances, in Proceeding of the First International Conference on Information Sciences and Systems, (University of Patras, Greece, 1976).

[6] Papoulis, A. and Beretsky, I., Improvement of range resolution of a pulse-echo system, Ultrasound in Medicine 3B, 1613-1625 (1977).

Section 5

Detection of Hidden Periodicities by Adaptive Extrapolation

ATHANASIOS PAPOULIS, FELLOW, IEEE, AND CHRISTODOULOS CHAMZAS

Abstract—A method is presented for determining the harmonic components of a noisy signal by nonlinear extrapolation beyond the data interval. The method is based on an algorithm that adaptively reduces the spectral components due to noise.

I. Introduction

An important problem in many applications is the determination of the frequency components of a signal

$$f(t) = \sum_{i=1}^{m} c_i e^{j\omega_i t} \tag{1}$$

in terms of the segment (data)

$$w_1(t) = \begin{cases} f(t) + n(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{2}$$

of $f(t)$ containing the noise component $n(t)$. The data are known for $|t| < T$ only for a variety of reasons:

1) The signal $f(t)$ can be written as a sum of exponentials for a limited time only (voice; nonstationary processes).

2) The available time of observation is limited (sun spots; weather trends).

3) Measurements are limited by instrument constraints (Michelson interferometer; diffraction-limited imaging).

Manuscript received November 22, 1978; revised January 25, 1979 and March 20, 1979. This work was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract N00014-76C 0144. This paper is in part from a Ph.D. dissertation submitted by C. Chamzas to the Faculty of the Polytechnic Institute of New York, Farmingdale, NY.

The authors are with the Department of Electrical Engineering, Polytechnic Institute of New York, Farmingdale, NY 11735.

0096-3518/79/1000-0492$00.75 © 1979 IEEE

-40

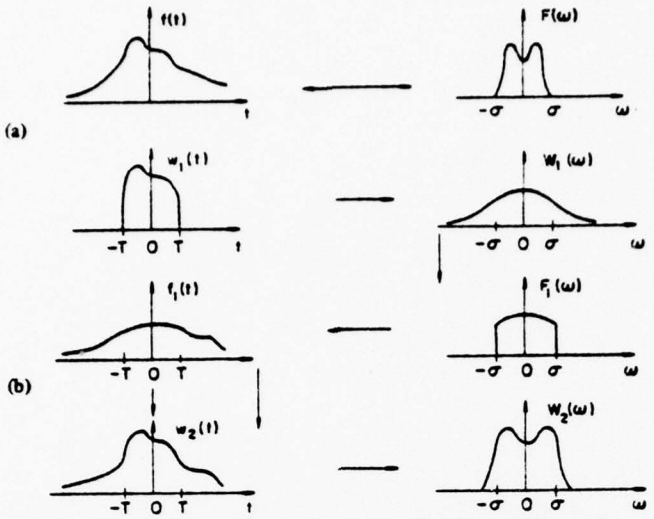


Fig. 1. (a) The unknown signal $f(t)$ and its Fourier transform $F(\omega)$. (b) First iteration starting with known segment $w_1(t)$.

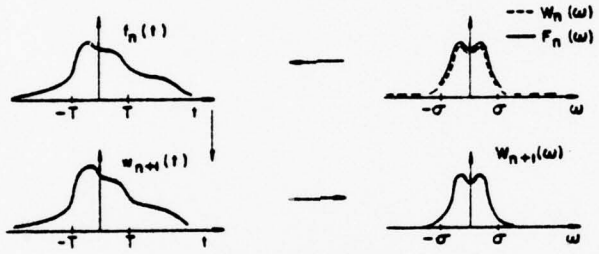


Fig. 2. nth iteration.

The unknown frequencies ω_i and coefficients c_i can be determined simply with ordinary Fourier transforms if the time of observation $2T$ is large compared to all the periods $T_i = 2\pi/\omega_i$ and their differences. This is not, however, the case if T is of the order $T_i - T_j$, particularly if the noise component $n(t)$ is not negligible. In this paper we present a method which, as we hope to show, is reliable even if T is small and the data are noisy.

The method involves only FFT and it is based on earlier results dealing with the problem of extrapolating band-limited functions [1], [2]. We review (for easy reference) the relevant parts of these results.

II. Extrapolation of Band-Limited Functions

Consider a function $f(t)$ with the Fourier transform $F(\omega)$ such that

$$F(\omega) = 0 \qquad |\omega| > \sigma. \tag{3}$$

We form the function

$$w_1(t) = \begin{cases} f(t) & |t| < T \\ 0 & |t| > T \end{cases} \tag{4}$$

obtained by truncating $f(t)$ as in Fig. 1. We shall determine $f(t)$ in terms of $w_1(t)$ by numerical iteration.

First Step: We compute the Fourier transform $W_1(\omega)$ of $w_1(t)$ and form the function

$$F_1(\omega) = \begin{cases} W_1(\omega) & |\omega| < \sigma \\ 0 & |\omega| > \sigma. \end{cases} \tag{5}$$

We compute the inverse transform $f_1(t)$ of $F_1(\omega)$, and form the function

$$w_2(t) = \begin{cases} w_1(t) = f(t) & |t| < T \\ f_1(t) & |t| > T \end{cases} \tag{6}$$

and its Fourier transform $W_2(\omega)$.

This completes the First Step of the iteration (Fig. 1).

nth Step: We form the function (Fig. 2)

$$F_n(\omega) = \begin{cases} W_n(\omega) & |\omega| < \sigma \\ 0 & |\omega| > \sigma \end{cases} \tag{7}$$

where $W_n(\omega)$ is the function obtained at the end of the preceding step and compute the inverse transform $f_n(t)$ of $F_n(\omega)$. We form the function

$$w_{n+1}(t) = \begin{cases} f(t) & |t| < T \\ f_n(t) & |t| > T \end{cases} \tag{8}$$

and compute its Fourier transform $W_{n+1}(\omega)$.

If $f(t)$ is approximated by $f_n(t)$, the resulting mean-square error is given by

$$E_n = \int_{-\infty}^{\infty} [f(t) - f_n(t)]^2\, dt = \frac{1}{2\pi}\int_{-\sigma}^{\sigma} |F(\omega) - F_n(\omega)|^2\, d\omega. \tag{9}$$

We maintain that this error decreases twice at each iteration step. Indeed,

$$E_n = \int_{|t|<T} [f(t) - f_n(t)]^2\, dt + \int_{|t|>T} [f(t) - f_n(t)]^2\, dt.$$

But [see (7) and (8)]

$$\int_{|t|>T} [f(t) - f_n(t)]^2\, dt = \int_{-\infty}^{\infty} [f(t) - w_{n+1}(t)]^2\, dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega) - W_{n+1}(\omega)|^2\, d\omega$$

$$= \frac{1}{2\pi}\int_{|\omega|>\sigma} |F(\omega) - W_{n+1}(\omega)|^2\, d\omega$$

$$+ \frac{1}{2\pi}\int_{-\sigma}^{\sigma} |F(\omega) - F_{n+1}(\omega)|^2\, d\omega.$$

And since the last integral E_{n+1} [see (9)], we obtain

$$E_n - E_{n+1} = \int_{|t|<T} [f(t) - f_n(t)]^2\, dt + \frac{1}{2\pi}\int_{|\omega|>\sigma} |W_{n+1}(\omega)|^2\, d\omega \tag{10}$$

because $F(\omega) = 0$ for $|\omega| > \sigma$.

In [1] and [2] we show that $f_n(t) \to f(t)$ as $n \to \infty$. This is not true if the given segment $w_1(t)$ of $f(t)$ is noisy as in (2). In this case, a satisfactory estimate of $f(t)$ can be found by early termination of the iteration [2].

Note: From (10) it follows that the mean-square error E_n is a monoton decreasing function and since it is positive it tends to a limit. This does not prove the convergence of (9) because the limit need not be zero. It shows, however, that

$$E_n - E_{n+1} \to 0 \qquad n \to \infty.$$

Hence,

$$\int_{|t|<T} [f(t) - f_n(t)]^2\, dt \to 0 \qquad n \to \infty. \tag{11}$$

Although the functions $f(t)$ and $f_n(t)$ are band limited, (11) *does not imply* that $f(t) \to f_n(t)$ because there is no lower bound on the energy concentration of band-limited functions in a finite interval [1], [3]. For example, the prolate spheroidal functions $\varphi_n(t)$ are band limited; their energy equals one but their energy concentration in the interval $(-T, T)$ tends to zero as $n \to \infty$. This is the case because the eigenvalues λ_n of the underlying integral equation tend to zero as $n \to \infty$.

We mention without elaboration that, in the discrete version of the problem, the convergence of the iteration can be deduced from (11) under suitable conditions. The reason is that the corresponding eigenvalues are finitely many, therefore, they have a positive minimum [4].

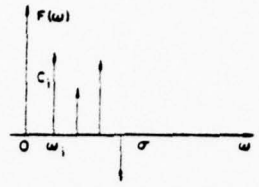


Fig. 3. Fourier transform of the unknown signal.

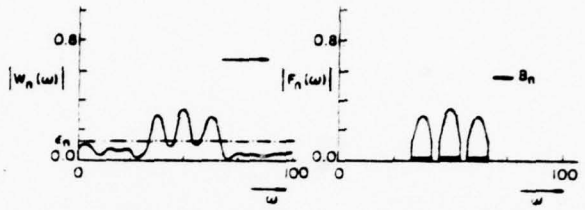


Fig. 4. Truncation of $W_n(\omega)$ below a threshold level ϵ_n yielding $F_n(\omega)$.

III. Adaptive Extrapolation

The preceding method was based on the assumption that the unknown function $f(t)$ is band limited. This information was used to reduce the error in the estimation of $f(t)$ twice at each iteration step. The speed of iteration can be increased and the effects of noise can be reduced if additional *a priori* information about $f(t)$ is available. Suppose, for example, that the size of the band of $F(\omega)$ is known but its precise location is unknown. We then choose a constant σ such that $F(\omega)$ vanishes outside the integral $(-\sigma, \sigma)$ and proceed as in Section II. As the iteration progresses, the form of $W_n(\omega)$ suggests appropriate reduction of the assumed band of $f(t)$.

The adaptive extrapolation method is particularly effective if $f(t)$ is a sum of exponentials as in (1). In this case, $F(\omega)$ consists of impulses (lines) as in Fig. 3:

$$F(\omega) = 2\pi \sum_{i=1}^{m} c_i \delta(\omega - \omega_i), \tag{12}$$

and our problem is to determine their locations ω_i and amplitudes c_i in terms of the known segment $w_1(t)$ of $f(t)$.

To solve this problem, we select a constant σ larger than the largest possible value of ω_i and we proceed with the iteration until $W_n(\omega)$ takes significant values only in a subset B_n of the band $(-\sigma, \sigma)$ of $f(t)$ (Fig. 4). This suggests that the unknown frequencies are in B_n. When this is observed, the function $F_n(\omega)$ of the nth iteration step is obtained from the following modification of (7)

$$F_n(\omega) = \begin{cases} W_n(\omega) & \omega \in B_n \\ 0 & \omega \in \overline{B}_n \end{cases} \tag{13}$$

(Fig. 4) where $\overline{B}_n$ is the complement of B_n. The process is repeated with occasional reduction of the size of B_n as further evidence suggests, and it terminates when $w_n(t)$ is essentially a sum of exponentials. Another application of the method is discussed in [5] in the context of deconvolution.

Discussion

The adaptive extrapolation method is essentially empirical. Although, as we see in the following examples, it works well in a number of cases, there is no *a priori* certainty that in a given problem it will converge to the unknown signal. In fact, if some of the components c_i of $f(t)$ are relatively small, they might be lost.

The accuracy and reliability of the method depends on a number of parameters: total number of unknown frequencies, possibly prior knowledge of this number, relative sizes of amplitudes c_i and frequencies ω_i, noise level, length $2T$ of the data interval, and available FFT size N. A precise statement, even empirical, of the importance of all these factors cannot be made; it would depend on many parameters. We are in the process of determining, empirically, the limits of the method for a number of special cases. We comment below, briefly, on certain empirical criteria for selecting the set B_n and on the limitations due to sampling.

For the subset B_n introduced in (13) we select the set of points such that the magnitude of $W_n(\omega)$ exceeds a threshold level ϵ_n:

$$\begin{aligned} |W_n(\omega)| &\geq \epsilon_n \qquad \omega \in B_n \\ |W_n(\omega)| &< \epsilon_n \qquad \omega \in \bar{B}_n. \end{aligned} \tag{14}$$

The choice of ϵ_n is dictated by two conflicting requirements: for a speedy convergence and noise reduction, ϵ_n must be large; it must be sufficiently small so that all frequency components of $f(t)$ remain in B_n. In the examples given below we used the following method for determining ϵ_n.

We first find the minimum M_{n-1} of $|W_{n-1}(\omega)|$ in the set B_{n-1}:

$$M_{n-1} = \min |W_{n-1}(\omega)| \qquad \omega \in B_{n-1}. \tag{15}$$

If ϵ_{n-1} is greater than μM_{n-1}, where μ is a constant less than one, then we do not change the threshold level. If ϵ_{n-1} is less than μM_{n-1} then we choose $\epsilon_n = \mu M_{n-1}$. Thus,

$$\epsilon_n = \max \{\epsilon_{n-1}, \mu M_{n-1}\}. \tag{16}$$

In the examples, μ is chosen between 0.9 and 0.99.

Numerical Considerations

The numerical implementation of the method involves the discrete signals

$$f_n = f(nt_o) \qquad F_n = F(n\omega_o)$$

obtained by sampling $f(t)$ and $F(\omega)$.

Suppose, first, that the problem is inherently discrete, i.e., that we wish to find the spectrum of a sequence f_n from incomplete data. Clearly, the discrete version of the iteration and of the band-limited assumption are self-evident. However, the assumption that $f(t)$ is a sum of sine waves has no obvious discrete version. It corresponds, loosely, to the assumption that the smallest distance of the nonzero frequencies is large compared to one (no "neighboring frequencies" are present). If this is the case, then the unknown frequencies can be determined exactly, provided that the data interval is not too small and the noise level is reasonable.

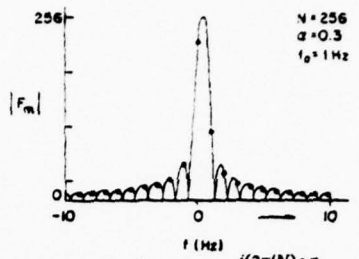


Fig. 5. Discrete spectrum $|F_m|$ of $f_{n_n} = e^{j(2\pi/N)\alpha n}$ ($\alpha = 0.3, N = 256$).

We turn now to our main problem: the numerical determination of the frequencies of an analog signal. We assume that the FFT size N is specified. It suffices, therefore, to select the size t_o of the sampling interval. As we know [1], the frequency interval is then determined because $\omega_o = 2\pi/Nt_o$. Since the data interval is $2T$, the number M of available samples equals $2T/t_o$. The choice of M is guided by the following considerations: if $M \ll N$, then the iteration might converge to the wrong frequencies. If M is large, then the aliasing errors are large.

It appears from our experience that $M = N/4$ is a reasonable compromise and it leads to $t_o \simeq 8T/N$. However, as we shall see, to increase the resolution we might use a larger value for t_o.

The accuracy of the method and the attainable resolution depend on the relationship between the unknown frequencies ω_i and the sampling frequency ω_o. If all unknown frequencies are multiples of ω_o

$$\omega_i = r_i \omega_o$$

then the problem is essentially discrete. If the unknown frequencies and their differences are large compared to ω_o, then the error is small because it is of the order of ω_o.

The problem of determining ω_i is difficult if ω_o is of the order of ω_i, and ω_i is not an integer multiple of ω_o

$$\omega_i = (r_i + \alpha)\, \omega_o \qquad |\alpha| < \tfrac{1}{2}.$$

In this case, the resolution error $\omega_o/2$ is of the order of ω_i. Furthermore, aliasing generates spurious frequencies in the vicinity of ω_i. Indeed, if

$$f(t) = e^{j\omega_i t}$$

then

$$f_n = e^{j\omega_i n t_o} = w^{(r_i+\alpha)n} \qquad w = e^{j2\pi/N}$$

yielding the discrete spectrum (Fig. 5)

$$F_m = \sum_{n=0}^{N-1} f_n w^{-mn} = \frac{1 - w^{(m - r_i - \alpha)N}}{1 - w^{(m - r_i - \alpha)}}.$$

To improve the accuracy, we can repeat the process with a larger value of t_o, using as starting B_o the set containing only the estimated frequencies ω_i and their neighbors.

IV. Illustrations

We illustrate the method with several examples involving signals whose unknown frequencies cannot be determined

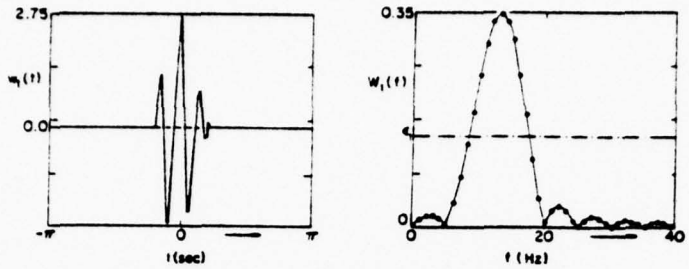


Fig. 6. Given segment $w_1(t)$ and its Fourier transform $W_1(\omega)$.

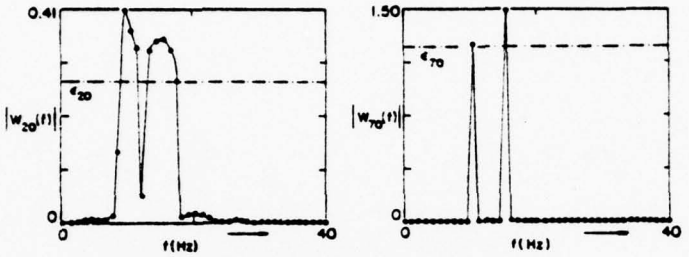


Fig. 7. Result of the iteration for $n = 20$ and $n = 70$.

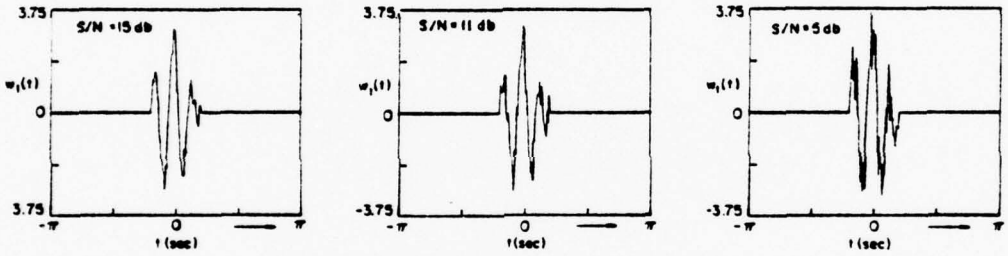


Fig. 8. Given data segment for $S/N = 15$, 11, and 5 dB.

from direct Fourier analysis. In these illustrations we consider several noise levels. With

$$w_1(t) = f(t) + n(t)$$

the given data, we define the signal-to-noise ratio S/N as the ratio of the energies of $f(t)$ and $n(t)$ in the data interval. In all examples, the noise is white and is uniformly distributed in the interval ($-c$ to c). The ratio S/N is changed by changing the size of c.

The computations are carried out with

$$N = 256 \quad f_o = 1 \text{ Hz} \quad t_o = 1/256 \text{ s}.$$

To avoid large scaling factors, we divided all frequency components by $N/2$. In the examples we show also the value of the parameter μ [see (16)] and of the initial threshold level ϵ_1.

Example 1: The unknown signal is a sum of two sine waves

$$f(t) = 1.5 \cos(30\pi t + 60^\circ) + 1.25 \cos(20\pi t + 30^\circ)$$

and the unknown frequencies $f_1 = 10$ Hz and $f_2 = 15$ Hz are integral multiples of the sampling frequency ω_o.

a) We first assume that the data interval contains $M = 51$ sampling points and $n(t) = 0$.

In Fig. 6 we show the given segment of the unknown signal and its spectrum. As we see from the figure, the frequencies f_1 and f_2 are not visible. The initial threshold is $\epsilon_1 = 0.15$ and its value at the nth iteration is obtained from (16) with $\mu = 0.99$. In Fig. 7 we show the results of the iteration for $n = 20$ and $n = 70$. At the 70th iteration the frequencies, amplitudes, and phases of $f(t)$ are recovered exactly.

We note that, in this case, the values of ϵ_1 and μ are not critical. Any value of μ between 0.9 and 0.99 and of ϵ_1 between 0.05 and 0.15 is adequate. The iteration was performed also with a data interval containing $M = 41$ sampling points. In this case, the results are similar but the speed of convergence is slower.

b) We consider, next, noisy data with various S/N ratios as in Fig. 8. In all cases,

$$M = 51 \quad \mu = 0.99 \quad \epsilon_1 = 0.15.$$

The iteration was performed several times with the same signal but with different samples of noise. As the following indicates, the results are not the same for all samples: $S/N = 15$ dB ($c = 0.375$). Six samples were tried. In five of these, the frequencies f_1 and f_2 were found exactly. $S/N = 11$ dB ($c = 0.625$). Fourteen samples were tried. In nine, we ob-

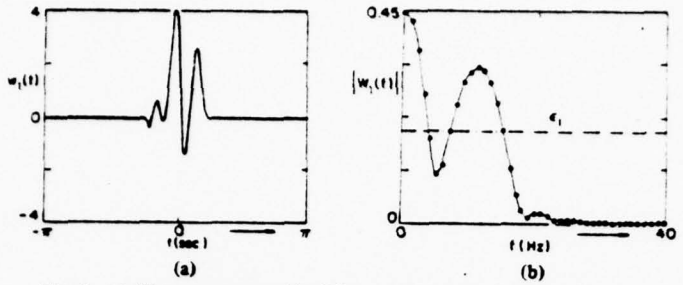


Fig. 9. (a) Given segment $w_1(t)$ of $f(t)$. (b) Fourier transform of $w_1(t)$.

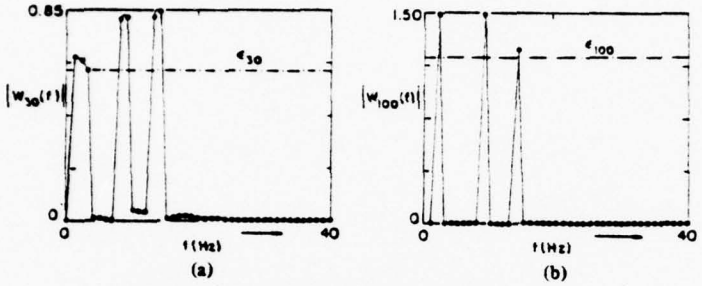


Fig. 10. Result of the iteration for $n = 30$ and $n = 100$.

tained f_1 and f_2 exactly. In four cases, an error of 1 Hz developed. In one case, the iteration yielded not two but three frequencies: $f_1 = 9$ Hz, $f_2 = 14$ Hz, and $f_3 = 15$ Hz. $S/N = 5$ dB ($c = 1.25$). This is a very noisy case. Of the eleven samples tried, three gave the correct answer, two yielded 1 Hz error, five resulted in 2 Hz error, and in one case the frequency $f_2 = 15$ Hz was lost.

Example 2: In this example $f(t)$ consists of three sine waves and the data are noiseless. We consider two cases. In the first case, the unknown frequencies are multiples of ω_o. In the second case, they are not.

a)

$$f(t) = 1.5 \cos 4\pi t + 1.5 \cos (18\pi t + 60^\circ) + 1.25 \cos (28\pi t + 30^\circ).$$

We start with the following values of the relevant parameters:

$$M = 59 \quad \mu = 0.95 \quad \epsilon_1 = 0.20.$$

In Fig. 9 we plot the given segment $f(t)$ and its spectrum. Fig. 10 shows the results of the iteration for $n = 30$ and $n = 100$. At the 100th iteration the frequencies, amplitudes, and phases of $f(t)$ are recovered exactly.

Again the values of μ and ϵ_1 are not critical. Essentially the same results are obtained if the data interval is reduced to $M = 51$ provided that μ is not less than 0.95.

The method has been tried also for a smaller data interval. However, the convergence is slow and the result inaccurate. With $M = 41$, $\mu = 0.99$, $\epsilon_1 = 0.20$ the component with the lowest frequency is lost.

b)

$$f(t) = 1.5 \cos 4.8\pi t + 1.5 \cos (18\pi t + 60^\circ) + 1.25 \cos (29.2t + 30^\circ).$$

In this case,

$$f_1 = (2 + 0.4) f_o \quad f_2 = 9 f_o \quad f_3 = (14 + 0.6) f_o.$$

We used $M = 59$, $\mu = 0.95$, and $\epsilon_1 = 0.20$.

With an FFT size $N = 256$, we obtained after 350 iteration steps the frequencies 2 Hz, 9 Hz, and 15 Hz (Fig. 11c).

Increasing the FFT size to $N = 512$, we found in 200 steps the frequencies 2.5 Hz, 8.75 Hz, 9.25 Hz, and 14.5 Hz. (Fig. 11d).

We note that the accuracy in the evaluation of coefficients of different levels can be improved if the threshold level ϵ_n is not constant through the band but it takes different values in the vicinity of each frequency. This is demonstrated in the next example.

Example 3: The unknown signal is a sum of five sine waves.

$$f(t) = 1.5 \cos 4\pi t + 1.25 \cos (12\pi t + 30^\circ) + 0.375 \cos (40\pi t + 60^\circ) + 0.625 \cos 50\pi t + 1.25 \cos (60\pi t + 45^\circ)$$

with frequencies 2, 6, 20, 25, and 30 Hz; the noise is zero. In Fig. 12 we show the given data, obtained with $M = 71$, and their spectrum. In the iteration we assume that $\mu = 0.99$ and $\epsilon_1 = 0.04$. The level of the threshold level at the nth iteration is defined as in (16). However, it is not constant throughout the band. Its value is determined from the behavior of $W_n(\omega)$ in the vicinity of each maximum (Fig. 13).

In Fig. 13 we show the iteration for $n = 10$ and $n = 20$. At the 50th step, (Fig. 14) we recover the frequencies 2, 6, 25, and 30. As it is clear from the figure, $W_n(\omega)$ contains a peak in the vicinity of $f = 20$. To determine its exact location we introduce the following variation to the method: we subtract from the given data the recovered portion of $f(t)$ and repeat the iteration starting with the new data $d(t)$ so obtained.

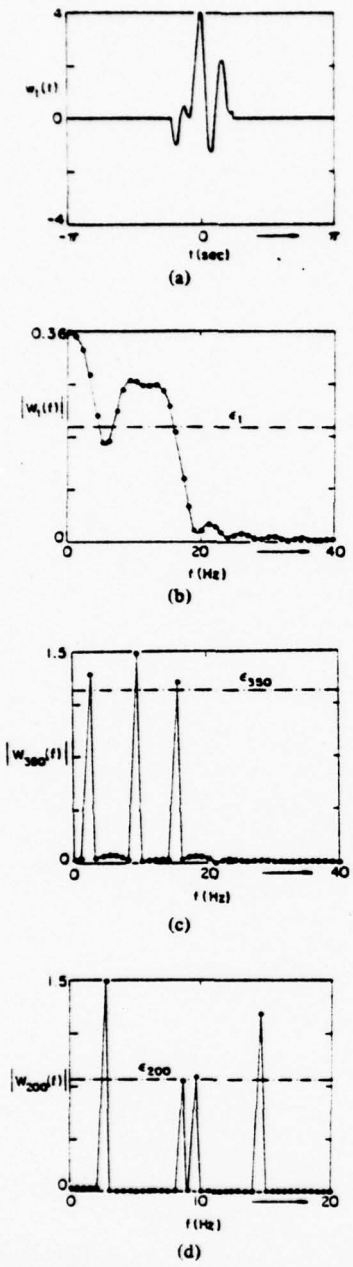


Fig. 11. (a) Given segment $w_1(t)$. (b) Fourier transform of $w_1(t)$. (c) Result of the iteration for $n = 350$ and FFT size $N = 256$. (d) Result of the iteration for $n = 200$ and FFT size $N = 512$.

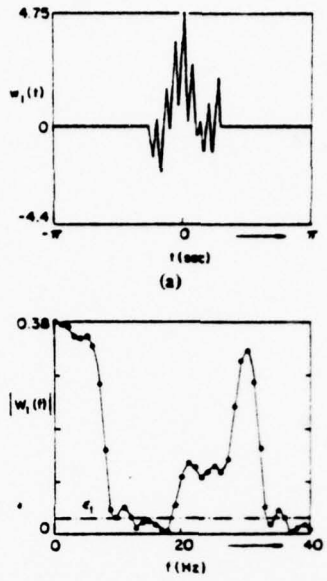


Fig. 12. (a) Given segment $w_1(t)$. (b) Fourier transform of $w_1(t)$.

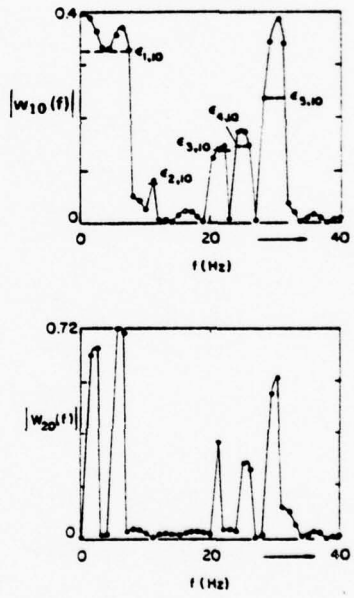


Fig. 13. Result of the iteration for $n = 10$ and $n = 20$.

In Fig. 15 we show $d(t)$ and its spectrum $D(\omega)$. The unknown frequency $f = 20$ is recovered at the 20th step (Fig. 16).

The iteration was performed also with a smaller data segment ($M = 61$). The results, however, were similar but the convergence slower.

Example 4: To test the limits of the method, we consider as a last case an example where the data interval is less than one-half the unknown period, and the unknown frequency is not a multiple of ω_o so that the aliasing is significant. We assume that

$$f(t) = 1.25 \cos (5.4\pi t + 30^\circ) \qquad T = 0.08 \text{ s}.$$

This yields $M = 41$ sampling points in the data interval.

The iteration was performed with $\mu = 0.99$ and $\epsilon_1 = 0.05$. We considered four different signal levels (Fig. 17).

a) $n(t) = 0$. At the 40th iteration we recover the frequency

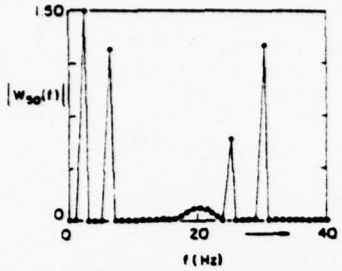
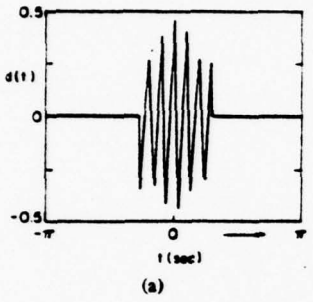


Fig. 14. Result of the iteration for $n = 50$.



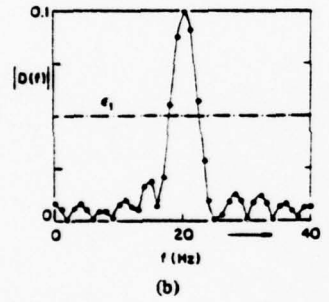


Fig. 15. (a) New data segment $d(t)$. (b) Fourier transform of $d(t)$.

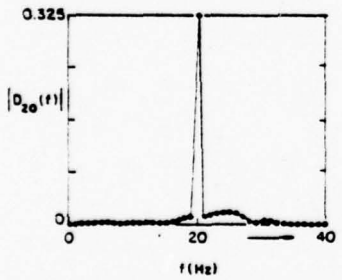
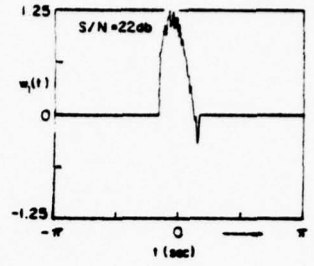
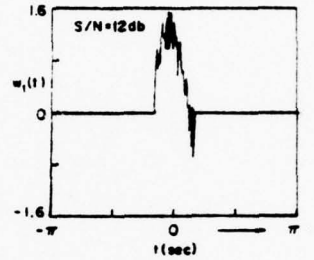
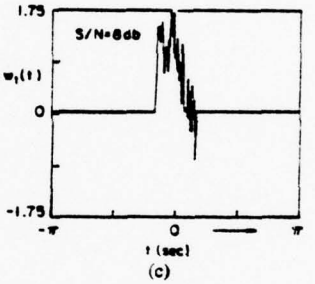


Fig. 16. Result of the iteration for $n = 20$.







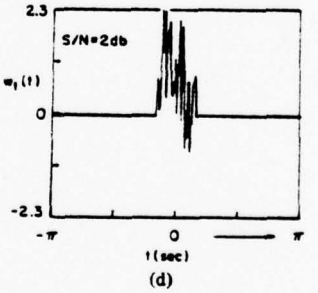


Fig. 17. Given data segment $w_1(t)$ for various noise levels. (a) S/N = 22 dB. (b) S/N = 12 dB. (c) S/N = 8 dB. (d) S/N = 2 dB.

$f = 3$ Hz. This is the nearest sampling frequency to the unknown $f_1 = 2.7$. However, since the resolution frequency $f_o = 1$ is of the order of f_1, the error is large. To reduce it, we increase the sampling interval from $t_o = 1/256$ to $t_o = 1/128$. This yields $f_o = 1/2$ Hz but the number of sampling points is reduced to $M = 21$. The iteration starts from the band B_o consisting of the location $f = 3$ of the recovered frequency and its two neighbors $f = 2.5$ and $f = 3.5$. After $n = 150$ steps, we recovered the frequency $f = 2.5$ (nearest to the unknown $f_1 = 2.7$).

The process was repeated with $t_o = 1/64$, that is, for $f_o = 1/4$ and $M = 11$ sampling points. The iteration yielded two frequencies: $f_a = 2.5$ and $f_b = 2.75$ with amplitudes $|c_a| = 0.616$, $|c_b| = 0.630$ and phases $\varphi_a = 30.06°$, $\varphi_b = 31.44°$, re-

TABLE I
RESULTS OF THE ITERATION FOR 20 NOISE SAMPLES

S/N	22 db (c=0.125)	12 db (c=0.375)	8 db (c=0.625)	2 db (c=1.25)
f_o (Hz)				
1 (M=41)	3 (20)	3 (18) 2 (2)	3 (15) 2 (4) 1 (1)	3 (12) 2 (3) 1 (1) 0 (1) - (3)
0.5 (M=21)	2.5 (8) 2.5, 3.0(8) 3.0(4)	2.5 (7) 2.5, 3.0(4) 3.0(9)	2.5 (7) 2.5, 3.0(1) 3.0(8) 3.5(2) - (2)	-
0.25 (M=11)	2.75(7) 2.5,2.75(13)	2.5 (9) 2.5, 2.75(7) 2.25,2.5(2) - (2)	-	-

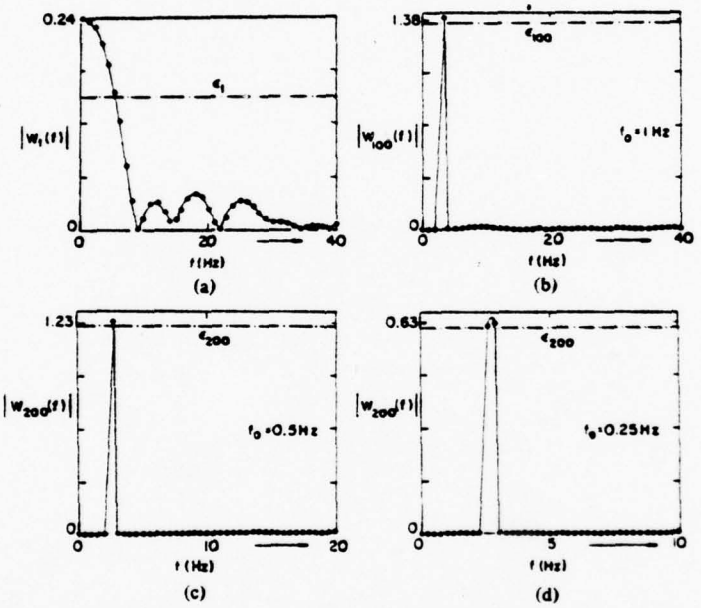


Fig. 18. $w_1(t) = 1.25 \cos(5.4\pi t + 30^\circ) + n(t)$ $|t| < 0.08$ s. (a) Fourier transform of $w_1(t)$. (b) Result of the iteration for $f_o = 1$ Hz ($M = 41$) and $n = 100$. $f(t) = 1.38 \cos(6\pi t + 25.5^\circ)$. (c) Result of the iteration for $f_o = 0.5$ Hz ($M = 21$) and $n = 200$. $f(t) = 1.23 \cos(5\pi t + 31.8^\circ)$. (d) Result of the iteration for $f_o = 0.25$ Hz ($M = 11$) and $n = 200$. $f(t) = 0.62 \cos(5\pi t + 30.1^\circ) + 0.63 \cos(5.5\pi t + 31.4^\circ) \simeq 1.246 \cos(2.63\pi t + 30.7^\circ)$.

spectively. The location $\hat{f}$ and amplitude $\hat{c}$ of the unknown frequencies was finally estimated by interpolation, yielding

$$\hat{f} = f_a + \frac{|c_b|}{|c_a| + |c_b|} f_o = 2.63 \text{ Hz}$$

$$\hat{c} = c_a + c_b = 1.246 \angle 30.76^\circ.$$

b) $n(t) \neq 0$. We considered four different signal-to-noise ratios. All cases were performed 20 times using different samples for the noise. The statistical conclusions are described in Table I. The numbers in the table are the estimates in Hz of the unknown frequency. Numbers in parentheses indicate in how many samples out of 20 the cited estimates were obtained.

In Fig. 18 we show the results for $S/N = 8$ dB and $f_o = 1$, 0.5, 0.25 Hz.

REFERENCES

[1] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill.
[2] —, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 735-742, Sept. 1975.
[3] D. Slepian, H. J. Landau, and H. O. Pollack, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty principle I and II," *Bell Syst. Tech. J.*, vol. 40, no. 1, 1961.
[4] A. Papoulis and M. Bertran, "Digital filtering and prolate functions," *IEEE Trans. Circuit Theory*, vol. CT-19, pp. 674-681, Nov. 1972.
[5] A. Papoulis and C. Chamzas, "Improvement of range resolution by spectral extrapolation," *Ultrasonic Imaging*, vol. 1, pp. 121-135, Apr. 1979.

Section 6

Theory and Applications of Running Transforms

6. 1. Introduction

In the early years of filter design, the objective was to find a realizable system whose frequency response $H(\omega)$ met certain requirements (low-pass, high-pass, or band-pass, for example). The input $f(t)$ was viewed as a deterministic signal characterized in terms of its Fourier transform $F(\omega)$ and the output $g(t)$ was evaluated from the inversion formula [1]

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)H(\omega)e^{j\omega t}\, d\omega \tag{1}$$

This is still a design method, however, in the last two or three decades, statistical considerations (optimum estimation or detection, for example) led often to design requirements based on the formula [2]

$$E\{ g^2(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)|H(\omega)|^2\, d\omega \tag{2}$$

expressing the average intensity of the output $g(t)$ of the filter in terms of the power spectrum $S(\omega)$ of the input $f(t)$.

In both cases, the design involves frequency domain considerations. However, the pertinent spectra are global: In (1) $F(\omega)$ is the Fourier transform of the entire signal $f(t)$; in (2) it is assumed that $f(t)$ is stationary, i.e., its statistical properties are the same for all t.

In a number of applications, it is desirable to design filters whose parameters depend not on global but on local properties of the signals to be processed. This leads to a time-varying system adaptively controlled according to some criterion based on local properties of the input.

A typical non-recursive discrete realization of such a system is shown in Fig. 1. The input x[n] is a sequence whose N values x[n],..., x[n-N+1] form the input vector X[n]. The resulting output g[n] is a weighted sum

$$y[n] = \sum_{k=0}^{N-1} a_k[n]\, x[n-k] = X^T[n]\, A[n] \tag{3}$$

where A[n] is a vector whose N components are the time-varying weights $a_k[n]$ of the system.

The above system has finite memory, hence, its response y[n] depends only on local properties of the input x[n]. However, since it is essentially a time-domain filter, it does not utilize directly the frequency-domain properties of x[n]. For a number of applications (noise or clutter rejection, for example) this is a disadvantage. Another disadvantage is the fact that the number N of its adaptively controlled parameters equals the length of its memory. This number can be reduced if the processing is performed recursively. However, adaptively controlled recursive filters have other disadvantages.

In this paper, we show that localized frequency domain processing is possible if the spectral properties of the input f(t) are expressed in terms of its running transform $F(t, \omega)$. The main objective of the paper is the development of running transforms. However, to demonstrate their value in filter design, we include two applications.

The results are presented in analog form. Discrete and two-dimensional signals are considered only briefly.

2. The running Fourier transform

The running Fourier transform of a signal f(t) is the integral [3]

$$F_c(t,\omega) = \int_{-c}^{c} f(t+\tau)\, e^{-j\omega\tau}\, d\tau \tag{4}$$

where c is a given constant. For a fixed t, $F_c(t,\omega)$ is the Fourier transform in the variable τ of the segment $f(t+\tau)$ of f(t) shown in Fig. 2.

The concept of the running transform is not new [4]. However, in earlier years, it was of limited use because of the then underlying computational difficulties.

The function $F_c(t,\omega)$ describes the local spectral properties of f(t) and their variation with time. As c increases from 0 to ∞, it changes from a time signal to a frequency signal. At the two extremes,

$$F_\infty(t,\omega) = F(\omega) \qquad \frac{1}{2c} F_c(t,\omega) \underset{c\to 0}{\to} f(t) \tag{5}$$

We shall use the notation

$$f(t) \longmapsto F_c(t,\omega)$$

to indicate that $F_c(t,\omega)$ is the running transform of f(t). The subscript c will often be omitted.

From the definition it follows readily that

$$f(t-t_o) \longmapsto F(t-t_o,\omega) \tag{6}$$

$$e^{j\omega_c t} f(t) \longmapsto e^{j\omega_c t} F(t,\omega-\omega_c)$$

A simple illustration follows:

$$1 \quad \longmapsto \quad \frac{2 \sin c\omega}{\omega}$$

$$e^{j\omega_c t} \quad \longmapsto \quad e^{j\omega_c t} \; \frac{2 \sin c(\omega-\omega_c)}{\omega-\omega_c}$$

More generally, if f(t) is a periodic function with period 2c:

$$f(t) = \sum_{r=-\infty}^{\infty} A_r \, e^{jr\omega_o t} \qquad\qquad \omega_o = \frac{\pi}{c}$$

then

$$F(t,\omega) = \sum_{r=-\infty}^{\infty} A_r \; \frac{2 \sin c(\omega - r\omega_o)}{\omega - r\omega_o} \, e^{jr\omega_o t} \tag{7}$$

Hence [see (6)],

$$e^{j\omega_c t} \sum_{r=-\infty}^{\infty} A_r \, e^{jr\omega_o t} \quad \longmapsto \quad e^{j\omega_c t} \sum_{r=-\infty}^{\infty} A_r \; \frac{2 \sin(\omega-\omega_c-r\omega_o)}{\omega-\omega_c-r\omega_o} e^{jr\omega_o t} \tag{8}$$

This result will be used later. We note that for the example in (8),

$$F(t, \omega_c + n\omega_o) = 2c \, A_n \, e^{j(\omega_c + n\omega_o)t} \tag{9}$$

Properties. Suppose first that f(t) is a deterministic signal with Fourier transform F(w). In this case, $F(t,\omega)$ has a Fourier transform in the variable t:

$$\overline{F}(w,\omega) = \int_{-\infty}^{\infty} F(t,\omega) \, e^{-jwt} \, dt \tag{10}$$

From (4) and (10) it follows readily that

$$\bar{F}(w,\omega) = \int_{-c}^{c} F(w)\, e^{jw\tau}\, e^{-jw\tau}\, d\tau = F(w)\, \frac{2\sin c(w-\omega)}{w-\omega} \tag{11}$$

This shows that for a specified ω, $F(t,\omega)$ can be considered as the output of a linear system with input $f(t)$ and system function

$$J(w,\omega) = \frac{2\sin c(w-\omega)}{w-\omega} \tag{12}$$

where w is the frequency variable and ω a constant parameter (Fig. 3a).

We now assume that $f(t)$ is a stationary random process with autocorrelation $R(\tau)$ and power spectrum $S(\omega)$. In this case, $F(t,\omega)$ is also a random process stationary in the variable t and non-stationary in the variable ω. We shall determine its autocorrelation

$$R_{FF}(\tau,\omega) = E\{F(t+\tau,\omega)\, F^{*}(t,\omega)\} \tag{13}$$

and power spectrum

$$S_{FF}(w,\omega) = \int_{-\infty}^{\infty} R_{FF}(\tau,\omega) e^{-jw\tau}\, d\tau \tag{14}$$

Since $F(t,\omega)$ is the output of the system of Fig. 3a with input $f(t)$, it follows from a well known theorem [2] that

$$S_{FF}(w,\omega) = S(w)\, \frac{4\sin^2 c(w-\omega)}{(w-\omega)^2} \tag{15}$$

The inverse transform of the above in the variable w yields

$$R_{FF}(\tau, \omega) = \int_{-2c}^{2c} (2c - |x|)\, R(x+\tau) e^{-jwx}\, dx \qquad (16)$$

This follows from the convolution theorem and the fact that the inverse transform of the fraction in (15) is a triangular pulse [1] multiplied by $e^{-j\omega x}$.

For a fixed t, $F(t, \omega)$ is a non-stationary process in ω. We shall show that

$$E\{F(t, u)F^*(t, v)\} = \frac{2}{\pi} \int_{-\infty}^{\infty} S(w)\, \frac{\sin c(u-w)\sin c(v-w)}{(u-w)(v-w)}\, dw \qquad (17)$$

Clearly, F(t, u) and F(t, v) are the outputs of two systems with common input f(t) and system functions

$$\frac{2 \sin c(w-u)}{w-u} \qquad \text{and} \qquad \frac{2 \sin c(w-v)}{w-v}$$

respectively (Fig. 3b). Hence, their cross-power spectrum in the variable t equals

$$S(w)\, \frac{4 \sin c(w-u)\sin c(w-v)}{(w-u)(w-v)}$$

This leads to the conclusion that

$$E\{F(t+\tau, u)F^*(t, v)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(w)\, \frac{4 \sin c(w-u)\sin c(w-v)}{(w-u)(w-v)}\, e^{jw\tau}\, dw \qquad (18)$$

and (17) follows with $\tau = 0$.

We note finally with $u=v=\omega$ that

$$E\{|F(t,\omega)|^2\} = \frac{2}{\pi}\int_{-\infty}^{\infty} S(w)\,\frac{\sin^2 c(w-\omega)}{(w-\omega)^2}\,dw \tag{19}$$

Inversion formula. We shall now show that $f(t)$ can be expressed in terms of the samples $F(t, m\omega_o)$ of its running transform:

$$f(t) = \frac{1}{2c}\sum_{m=-\infty}^{\infty} F(t, m\omega_o) \qquad \omega_o = \frac{\pi}{c} \tag{20}$$

Proof. We expand the function $f(t+\tau)$ into a Fourier series in the variable τ. The coefficients of this expansion in the interval $(-c, c)$ are given by [see (4)]

$$\frac{1}{2c}\int_{-c}^{c} f(t+\tau)e^{-jm\omega_o\tau}\,d\tau = \frac{1}{2c}\,F(t, m\omega_o) \tag{21}$$

Hence,

$$f(t+\tau) = \frac{1}{2c}\sum_{m=-\infty}^{\infty} F(t, m\omega_o)e^{jm\omega_o\tau} \tag{22}$$

for $|\tau| < c$. And with $\tau=0$, (20) follows.

Sampling expansion. The preceding result is a version of the sampling theorem and is based on the fact that $F(t,\omega)$ is the Fourier transform of the time-limited function shown in Fig. 2. Thus, $f(t)$ and, therefore, $F(t,\omega)$ is uniquely determined in terms of $F(t, m\omega_o)$. In fact, $F(t,\omega)$ can be expressed explicitly in terms of its samples $F(t, m\omega_o)$:

$$F(t,\omega) = \sum_{m=-\infty}^{\infty} F(t, m\omega_o)\,\frac{\sin c(\omega-m\omega_o)}{c(\omega-m\omega_o)} \tag{23}$$

Proof. Inserting (22) into (4) and integrating term-wise, we obtain (23).

Interpolation. From (23) it follows with $\omega = r\omega_o + \omega_o/2$ that

$$F(t, r\omega_o + \omega_o/2) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} \frac{(-1)^{r-k}}{r-k+1/2} F(t, k\omega_o) \tag{24}$$

This result can be used to increase the resolution in the spectral representation of f(t). Indeed, suppose that we know the samples $F_c(t, m\omega_o)$ of $F_c(t, \omega)$. If we double the interval (-c, c), we obtain the running transform $F_{2c}(t, \omega)$ whose samples are $F_{2c}(t, m\omega_o/2)$. We shall express these samples in terms of $F_c(t, m\omega_o)$.

From the definition (4) it follows readily that

$$F_{2c}(t, \omega) = F_c(t-c, \omega) + F_c(t+c, \omega) \tag{25}$$

If $m = 2r$ is even, then (25) yields

$$F_{2c}(t, m\omega_o/2) = F_c(t-c, r\omega_o) + F_c(t+c, r\omega_o) \tag{26}$$

If $m = 2r+1$ is odd, then [see (24)]

$$F_{2c}(t, m\omega_o/2) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} \frac{(-1)^{r-k}}{r-k+1/2} \left[F_c(t-c, k\omega_o) + F_c(t+c, k\omega_o) \right] \tag{27}$$

Reasoning similarly, we can express the samples $F_{nc}(t, m\omega_o/n)$ of $F_{nc}(t, \omega)$ in terms of the samples $F_c(t, m\omega_o)$ of $F_c(t, \omega)$.

Existance and uniqueness. An arbitrary function $F(t, \omega)$ of two variables is not, in general, a running transform. It must satisfy the sampling expansion (23). We shall examine the properties of its samples

$$a_m(t) = F(t, m\omega_o) \tag{28}$$

Suppose, first, that f(t) is a deterministic signal with Fourier transform F(w). From (11) it follows that, in this case, the Fourier transform $A_m(\tilde{w})$ of $a_m(t)$ is given by

$$A_m(w) = \bar{F}(w, m\omega_o) = F(w)\ \frac{2\sin c(w-m\omega_o)}{w-m\omega_o} \tag{29}$$

This leads to the conclusion that

$$A_m(w) = 0 \quad \text{for} \quad w=r\omega_o \qquad r \neq m \tag{30}$$

If f(t) is a stationary stochastic process with power spectrum S(w), then $a_m(t)$ is also a stationary process and its power spectrum $S_m(w)$ is given by [see (15)]

$$S_m(w) = S_{FF}(w, m\omega_o) = S(w)\ \frac{4\sin^2 c(w-m\omega_o)}{(w-m\omega_o)^2} \tag{31}$$

From this it follows that

$$S_m(w) = 0 \quad \text{for} \quad w=r\omega_o \qquad r \neq m \tag{32}$$

We consider now the converse problem: Given a function $a_m(t)$, we wish to find a function $f(t)$ such that the mth sample $F(t, m\omega_o)$ of its running transform equals $a_m(t)$. We assume, of course, that if $a_m(t)$ is a deterministic signal, then it Fourier transform $A_m(w)$ satisfies (30); if it is a random process; then its power spectrum $S_m(w)$ satisfies (32).

To determine f(t), we form the inverse

$$J^{(-1)}(w, m\omega_o) = \frac{w-m\omega_o}{2 \sin c(w-m\omega_o)} \tag{33}$$

of the system $J(w, m\omega_o)$ defined in (12). We maintain that if the input to this system is the given function $a_m(t)$, then the resulting output is the unknown f(t).

Indeed, because of our assumptions, f(t) exists [see also (11) and (15)]. Furthermore, the mth sample $F(t, m\omega_o)$ of its running transform is the output of the filter $J(w, m\omega_o)$ with input f(t), that is, it is the output of the filters $J^{(-1)}(w, m\omega_o)$ and $J(w, m\omega_o)$ is cascade with input $a_m(t)$. Hence,

$$F(t, m\omega_o) = a_m(t)$$

We conclude with the investigation of the uniqueness of the signal f(t) so obtained. From this discussion it will follow that the functions $F(t, m\omega_o)$ are not arbitrary.

Suppose, first, that the running transform $F(t, \omega)$ vanishes identity in t for some $\omega = \omega_c$:

$$F(t, \omega_c) \equiv 0 \tag{34}$$

If f(t) is a deterministic signal with Fourier transform F(w), then (34) yields [see (11)]

$$\bar{F}(w, \omega_c) = F(w) \frac{2 \sin c(w-\omega_c)}{w-\omega_c} \equiv 0 \tag{35}$$

This is an identity in w and can be satisfied only if

$$F(w) = \sum_r{}' b_r \delta(w-\omega_c-r\omega_o) \tag{36}$$

where the prime in the summation indicates that r takes all integer values except r=0. From (36) it follows that

$$f(t) = \frac{1}{2\pi} e^{j\omega_c t} \sum_r{}' b_r e^{jr\omega_o t} \tag{37}$$

Consider, next, two signals $f_1(t)$ and $f_2(t)$ whose running transforms are identical in t for some $\omega = \omega_c$:

$$F_1(t, \omega_c) \equiv F_2(t, \omega_c) \tag{38}$$

With

$$f(t) = f_1(t)-f_2(t) \longmapsto F_1(t, \omega) - F_2(t, \omega) = F(t, \omega)$$

it follows that $F(t, \omega_c) \equiv 0$. Hence [see (37)],

$$f_1(t) = f_2(t) + \frac{1}{2\pi} e^{j\omega_c t} \sum_r{}' b_r e^{jr\omega_o t} \tag{39}$$

This shows that if the running transform $F(t, \omega)$ of $f(t)$ is specified for some $\omega = m\omega_o$ as in (28), then $f(t)$ is unique within a periodic function of period $2c$ whose mth harmonic is zero. Thus, the running transforms of the functions

$$f(t) + \frac{1}{2\pi} \sum_{r \neq m} b_r e^{jr\omega_o t} \tag{40}$$

given by [see (8)]

$$F(t, \omega) + \frac{1}{\pi} \sum_{r \neq m} b_r \frac{\sin c(\omega - r\omega_o)}{\omega - r\omega_o} e^{jr\omega_o t} \tag{41}$$

are identical in t for $\omega = m\omega_o$.

Similar conclusions hold if $f(t)$ is a random process.

3. Filtering

The applications of running transforms in filtering and estimation are based on the following:

If a linear system is time-limited, i.e., if its impulse response $h(t)$ is such that

$$h(t) = 0 \quad \text{for} \quad |t| > c \tag{42}$$

then the response

$$g(t) = \int_{-c}^{c} f(t-\alpha)h(\alpha)\, d\alpha \tag{43}$$

to an arbitrary input f(t) is given by

$$g(t) = \frac{1}{2c} \sum_{m=-\infty}^{\infty} F(t, m\omega_o)\, H(m\omega_o) \tag{44}$$

where

$$H(\omega) = \int_{-c}^{c} h(t) e^{-j\omega t} dt \tag{45}$$

is the system function and $F(t, \omega)$ is the running transform of the input.

Proof. Setting $\tau = -\alpha$ in (22), inserting into (43) and integrating term-wise, we obtain (44).

Equation (44) expresses g(t) in terms of the samples $H(m\omega_o)$ of $H(\omega)$ and the spectral properties of f(t) in the interval (-c, c). It involves an infinite number of terms, however, only a small number of these terms need be considered. Indeed, if f(t) is a deterministic signal, then

$$F(t, m\omega_o) = \int_{-\infty}^{\infty} F(w)\, \frac{\sin c(w - m\omega_o)}{\pi(w - m\omega_o)}\, dw \tag{46}$$

This follows from (11) and the Fourier inversion formula. If f(t) is a random signal, then [see (19)]

$$E\{\, |F(t, m\omega_o)|^2 \} = \frac{2}{\pi} \int_{-\infty}^{\infty} S(w)\, \frac{\sin^2 c(w - m\omega_o)}{(w - m\omega_o)^2}\, dw \tag{47}$$

Both relationships show that, if the bandwidth of f(t) equals σ, then the number of significant terms $F(t, m\omega_o)$ in (44) is of the order of σ/ω_o. This number can be made as small as desired if the constant $c = \pi/\omega_o$ is sufficiently small.

Running low-pass and band-pass filters. A special case of (44) is the low-pass filter obtained with

$$H(m\omega_o) = \begin{cases} 1 & |m| \le M \\ 0 & |m| > M \end{cases} \tag{48}$$

In this case, (44) yields

$$g(t) = \frac{1}{2c} \sum_{m=-M}^{M} F(t, m\omega_o) = \frac{1}{2c} F(t, 0) + \frac{1}{c} \sum_{m=1}^{M} \text{Re } F(t, m\omega_o) \tag{49}$$

Inserting (4) into (49), we obtain

$$g(t) = \frac{1}{2c} \sum_{m=-M}^{M} \int_{-c}^{c} f(t-\alpha)\, e^{jm\omega_o\alpha}\, d\alpha = \frac{1}{2c} \int_{-c}^{c} f(t-\alpha) \sum_{m=-M}^{M} e^{jm\omega_o\alpha}\, d\alpha$$

This shows that the impulse response of the running low-pass filter is given by

$$h_M(t) = \frac{1}{2c} \sum_{m=-M}^{M} e^{jm\omega_o t} = \frac{\sin(M+1/2)\omega_o t}{2c \sin(\omega_o t/2)} \tag{50}$$

for $|t| < c$ and it equals zero for $|t| > c$.

The corresponding frequency response is given by

$$H_M(\omega) = \sum_{m=-M}^{M} \frac{\sin c(\omega - m\omega_o)}{c(\omega - m\omega_o)} \tag{51}$$

In Fig. 4, we plot the functions $h_M(t)$ and $H_M(\omega)$ for a fixed c and for M from 0 to five. The impulse response $h_M(t)$ is zero for $|t| > c$; however, its effective duration is smaller. Using as its measure the width t_c of the main lobe, we find

$$t_c = \frac{2\pi}{(2M+1)\omega_o} = \frac{2c}{2M+1} \tag{52}$$

The frequency response $H_M(\omega)$ is a low-pass curve with cut-off frequency

$$\omega_c = (M+\frac{1}{2})\,\omega_o = \frac{\pi}{t_c} \tag{53}$$

In Fig. 5, we plot $h_M(t)$ and $H_M(\omega)$ for a fixed value of t_c and for M from 0 to five. The bandwidth ω_c of the filter or the duration t_c of its impulse response can, thus, be controlled by varying c or M.

For $M=0$, (49) yields

$$g(t) = \frac{1}{2c}\ F(t, 0) = \frac{1}{2c}\int_{-c}^{c} f(t+\tau)d\tau \tag{54}$$

This is the moving average of $f(t)$ and it has the following familiar applications: Suppose that

$$f(t) = y(t) + n(t) \tag{55}$$

where $y(t)$ is a signal to be estimated and $n(t)$ is noise. The component of $g(t)$ due to the noise decreases if c is large, however, the distortion of $y(t)$ (bias) due to smoothing increases. To minimize the resulting mean-square error, we must choose c adaptively [5] taking into

consideration the local properties of s(t) and n(t). The running low-pass filter permits us to determine the local properties of these signals and to vary the effective smoothing interval by controlling the constant c or the number M. Furthermore, it improves the estimation because it uses not only the average but also the first M harmonics of f(t) in the interval (t-c, t+c).

The running band-pass filter is defined similarly. Its output is given by

$$g(t) = \frac{1}{c} \sum_{m=M_1}^{M_2} \text{Re } F(t, m\omega_o) \tag{56}$$

and its frequency response by

$$H_{M_1M_2}(\omega) = H_{M_2}(\omega) - H_{M_1}(\omega) \tag{57}$$

where $H_M(\omega)$ is the corresponding low-pass filter [see (51)]. The cut-off points of $H_{M_1M_2}(\omega)$ are given by

$$\omega_{c_1} = (M + \tfrac{1}{2})\omega_o \qquad \omega_{c_2} = (M_2 + \tfrac{1}{2})\omega_o \tag{58}$$

and are controlled by varying M_1, M_2 and the interval $c = \pi/\omega_o$.

In Fig. 6, we plot $H_{M_1M_2}(\omega)$ for $M_1 = 3$ and $M_2 = 8$.

3. Discrete processing

The numerical implimentation of the running Fourier filter involves truncation of the sum in (44) and evaluation of the samples

$$F[n,m] = F(nT, m\omega_o) \tag{59}$$

of the running transform $F(t,\omega)$ in terms of the samples

$$f[n] = f(nT) \tag{60}$$

of f(t). In the above,

$$T = 2c/N \tag{61}$$

is the sampling interval and it is chosen so as to yield a small aliasing error. The truncation and aliasing errors can be determined as usual using the inversion formula [see (11)]

$$F(t,\omega) = \int_{-\infty}^{\infty} F(w)\ \frac{\sin c(w-\omega)}{\pi(w-\omega)}\ e^{jwt}\ dw \tag{62}$$

for deterministic signals and the relationship [see (15)]

$$E\{\,|F(t,\omega)|^2\,\} = \int_{-\infty}^{\infty} \frac{2\sin^2 c(w-\omega)}{\pi(w-\omega)^2}\ S(w)\,dw \tag{63}$$

for random signals.

We present, next, the digital version of running transforms.

The running discrete Fourier series (abbreviation: DFS) of a discrete signal $f[n]$ is the sum [3], [6]

$$F[n,m] = \sum_{k=0}^{N-1} f[n-k]\, w^{km} \qquad w = e^{j2\pi/N} \tag{64}$$

where N is a given integer. For a fixed n, $F[n,m]$ is the DFS in the variable k of the segment $f[n-k]$ of $f[n]$ shown in Fig. 7. Unline (4), we chose the discrete time n not in the center but at the end of the segment shown in Fig. 7. We did so to conform with the usual notations for DFS. If (64) is used as the discrete version of (4), a delay equal to $c = NT/2$ must be allowed.

Inversion formula. From the inversion formula for discrete Fourier series it follows that [3]

$$f[n-k] = \frac{1}{N} \sum_{m=0}^{N-1} F[n,m]\, w^{-km} \qquad 0 \le k < N \tag{65}$$

With $k=0$, this yields

$$f[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[n,m] \tag{66}$$

Filtering. Suppose that the discrete signal $f[n]$ is the input to a non-recursive filter of length N. We maintain that the resulting output

$$g[n] = \sum_{k=0}^{N-1} f[n-k]\, h[k] \tag{67}$$

is given by

$$g[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[n, m]\, H(w^m) \tag{68}$$

where

$$H(z) = \sum_{k=0}^{N-1} h[k]\, z^{-k} \tag{69}$$

is the system function and $F[n, m]$ is the running DFS of the input.

Proof. Inserting (65) into (67) and changing the order of summation, we obtain

$$g[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[n, m] \sum_{k=0}^{N-1} h[k]\, w^{-km} \tag{70}$$

and (68) results because the last sum equals $H(w^m)$.

The filter specified by (68) has been used in a different context as a frequency domain interpolator [7].

Equation (68) is the discrete version of (44) and it forms the basis of running frequency domain filtering. Its implimentation is shown in Fig. 8. It consists of a memory-less system W transforming the N samples of the input $f[n]$ into their DFS, followed by N weights $H(w^m)$ and an adder. The discrete version of the running low-pass and band-pass filters discussed earlier results if $H(w^m)$ takes the values zero or one.

The system of Fig. 8 can be used in a variety of applications. As an illustration, we present briefly the frequency domain version of the Wiener filter stressing the advantages over the familiar approaches.

The problem under consideration is the estimation of a discrete process $y[n]$ in terms of the N most recent samples $x[n-k]$ of another process $x[n]$ (data). The usual approach to this problem involves the determination of the N weights a_k such that if the sum

$$\hat{y}[n] = \sum_{k=0}^{N-1} a_k \, x[n-k] = X^T[n] \, A \tag{71}$$

is used as the estimate of $y[n]$, the resulting MS error

$$E\{ |y[n] - \hat{y}[n]|^2 \}$$

is minimum. As, we know (orthogonality principle), the optimum vector A is given by

$$A = R^{-1} \Gamma \tag{72}$$

where

$$R = E\{ X[n] \, X^T[n] \}$$

is the correlation matrix of the data vector $X[n]$ and

$$\Gamma = E\{ X[n] \, y[n] \}$$

is the cross-correlation vector between the data vector and the signal y[n] to be estimated.

This approach has the disadvantage that it involves the inversion of the matrix R whose order N is,in general, large. This problem is particularly severe, if R is not known in advance but it is estimated for each n. In this case, R depends on n(adaptive Wiener filter) and the inversion must be performed for each n.

In the frequency domain version of the Wiener filter, the unknown signal y[n] is estimated by the sum

$$\hat{y}[n] = \sum_{m=-M}^{M} b_m \, F[n, m] \tag{73}$$

where F[n, m] is the running DFS of the data sequence x[n]. If the optimality criterion is the minimization of the MS error, then the vector B of the unknown weights b_m is given by

$$B = C^{-1} D \tag{74}$$

where C is a matrix of order 2M+1 with elements

$$E \{ F[n, m_i] \, F^{*} [n, m_j] \}$$

and D is a vector with elements

$$E \{ F[n, m] \, y[n] \}$$

In this approach, the required value of M for nearly optimal estimation is small compared to N. Furthermore, the matrix C is nearly diagonal and can be easily inverted. This follows from the discrete version of (17) but the details are omitted.

We mention without elaboration that the frequency domain Wiener filter (73) properly modified can be used in the estimation of non-stationary processes.

The DFS analyzer. The implimentation of the frequency domain filter of Fig. 8 involves the evaluation of the running DFS $F[n, m]$ of the input. It appears from (64) that the required number of multiplications equal N-1 for each m. However, as we show next, this number can be reduced to a single multiplication if $F[n, m]$ is evaluated recursively. Indeed, from (64) it follows that

$$F[n, m] = w^m F[n-1, m] + f[n] - f[n-N] \tag{75}$$

This can be realized with a first order system if we have access to $f[n]$ and $f[n-N]$. The realization of the component W of the running filter of Fig. 8 requires 2M+1 such systems.

4. Two-dimensional signals

The preceding analysis can be extended to two-dimensional signals. We conclude with a brief outline of typical results, omitting all proofs.

The running Fourier transform of a signal $f(x, y)$ is the integral

$$F(x, y; u, v) = \int_{-a}^{a} \int_{-b}^{b} f(x+\xi,\ y+\eta) e^{-j(u\xi + v\eta)}\, d\xi\, d\eta \tag{76}$$

where a and b are given constants.

Inversion formula. It follows as in (20), that

$$f(x, y) = \frac{1}{4ab} \sum_{m_1 m_2 = -\infty}^{\infty} F(x, y; m_1 u_o, m_2 v_o) \qquad u_o = \frac{\pi}{a} \quad v_o = \frac{\pi}{b} \tag{77}$$

Filtering. If a two-dimensional linear system is space-limited, i.e., if its point spread $h(x, y)$ is such that

$$h(x, y) = 0 \qquad \text{for} \quad |x| > a\,, \quad |y| > b \tag{78}$$

then its response

$$g(x, y) = \int_{-a}^{a} \int_{-b}^{b} f(x-\xi,\ y-\eta)\, d\xi\, d\eta \tag{79}$$

to an arbitrary input $f(x, y)$ is given by

$$g(x, y) = \frac{1}{4ab} \sum_{m_1, m_2 = -\infty}^{\infty} F(x, y; m_1 u_o, m_2 v_o)\, H(m_1 u_o, m_2 v_o) \tag{80}$$

where $H(u, v)$ is the system function and $F(x, y; u, v)$ is the running transform of the input.

Discrete signals. The running DFS of a two-dimensional discrete signal $f[n_1, n_2]$ is the sum

$$F[n_1, n_2; m_1, m_2] = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} f[n_1-k_1,\ n_2-k_2]\, w_1^{k_1 m_1}\, w_2^{k_2 m_2} \tag{81}$$

where

$$w_1 = e^{j2\pi/N_1}, \quad w_2 = e^{j2\pi/N_2}$$

Inversion formula . It follows as in (65) that

$$f[n_1, n_2] = \frac{1}{N_1 N_2} \sum_{m_1=0}^{N_1-1} \sum_{m_2=0}^{N_2-1} F[n_1, n_2; m_1, m_2] \tag{82}$$

Filtering. If the signal $f[n_1, n_2]$ is the input to a non-recursive two-dimensional filter, then the resulting output

$$g[n_1, n_2] = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} f[n_1-k_1, n_2-k_2]\, h[k_1, k_2] \tag{83}$$

is given by

$$g[n_1, n_2] = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F[n_1, n_2; m_1, m_2]\, H(w_1^{m_1}, w_2^{m_2})$$

where

$$H(z_1, z_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} h[k_1, k_2]\, z_1^{-k_1} z_2^{-k_2}$$

is the system function and $F[n_1, n_2; m_1, m_2]$ is the running DFS of the input.

References

[1] A. Papoulis, The Fourier Integral and Its Applications, McGraw-Hill, New York, 1962.

[2] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

[3] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.

[4] C.H. Page, "Instantaneous Power Spectra", J.Appl.Phys., 1952, pp.103-106.

[5] A. Papoulis, "Two-to-One Rule in Data Smoothing", IEEE Trans. Inform. Theory, vol.IT-23, pp. September 1977.

[6] J.B. Allen and L.R. Rabiner, "A Unified Theory of Short-Time Spectrum Analysis and Synthesis", Proc. IEEE, November 1977.

[7] L.R. Rabiner and R.W. Shafer, "Recursive and Nonrecursive Realizations of Digital Filters Designed by Frequency Sampling Techniques, IEEE Trans. on Audio and Electroacoustics, vol.AU-19, no.3, pp.200-207, September 1971.

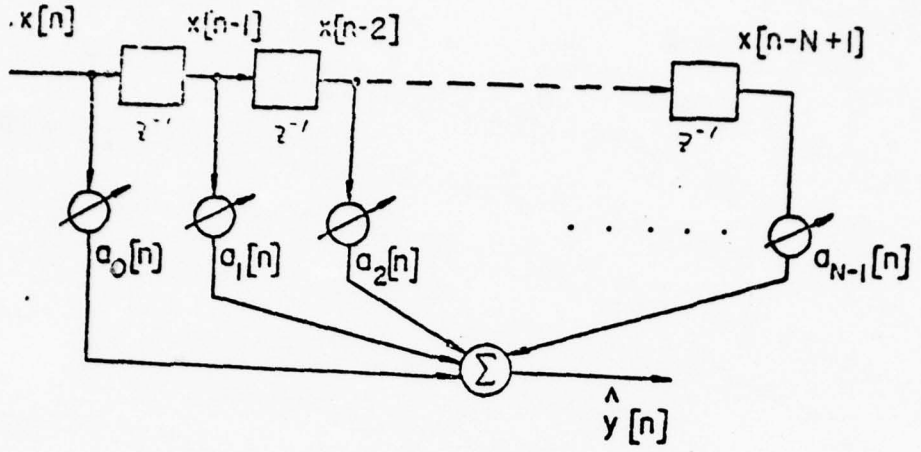


Fig. 1

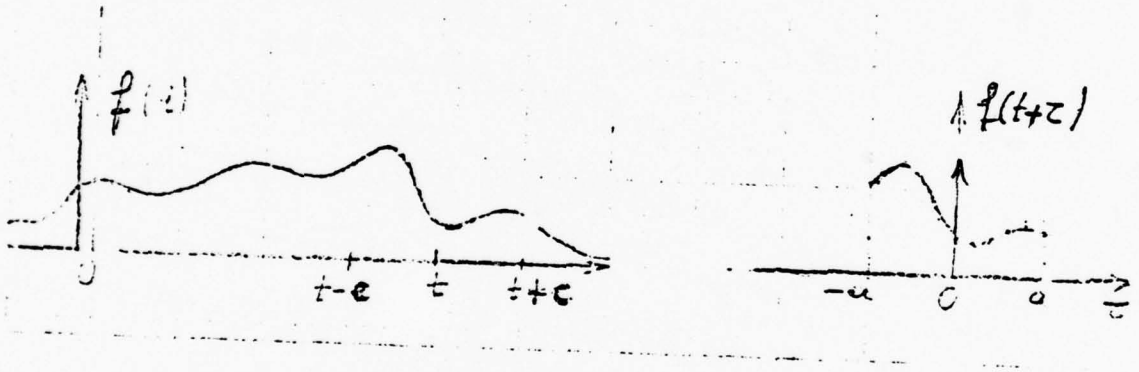


Fig. 2

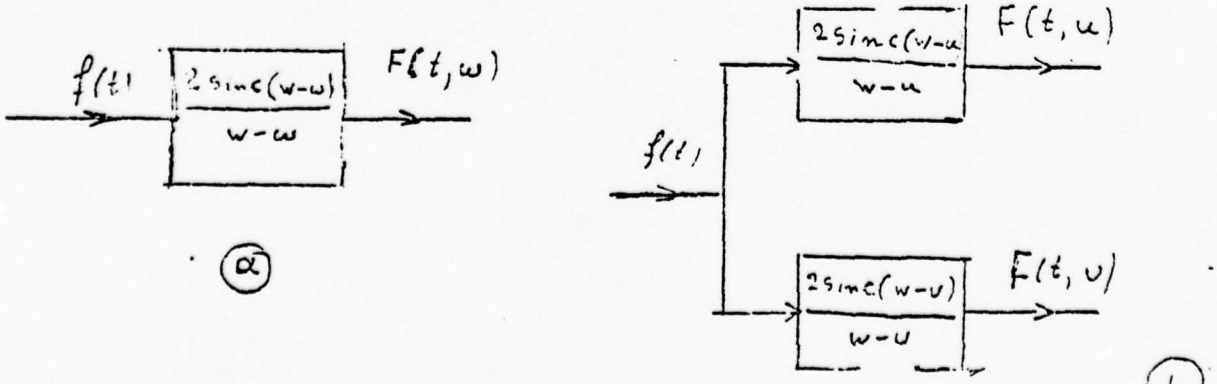
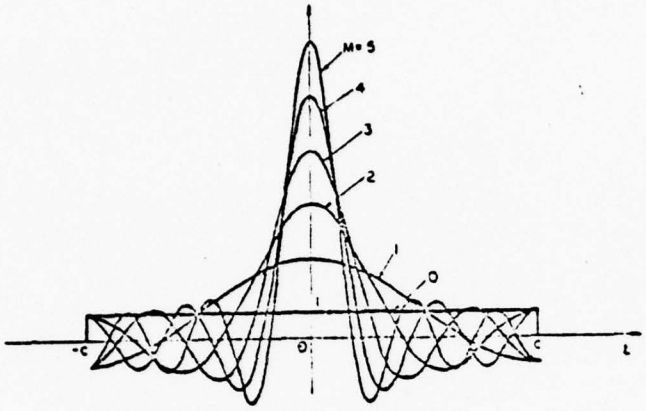
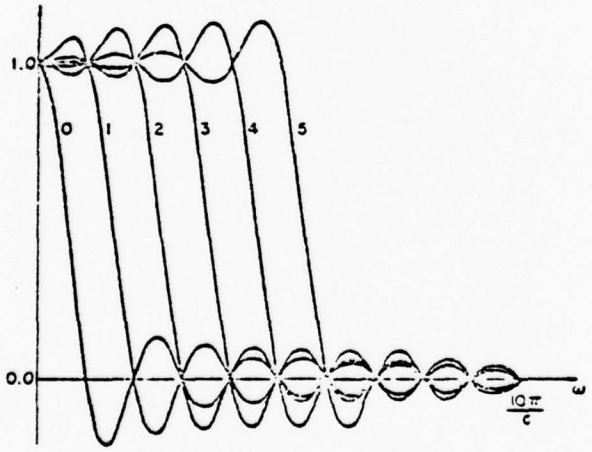


Fig. 3

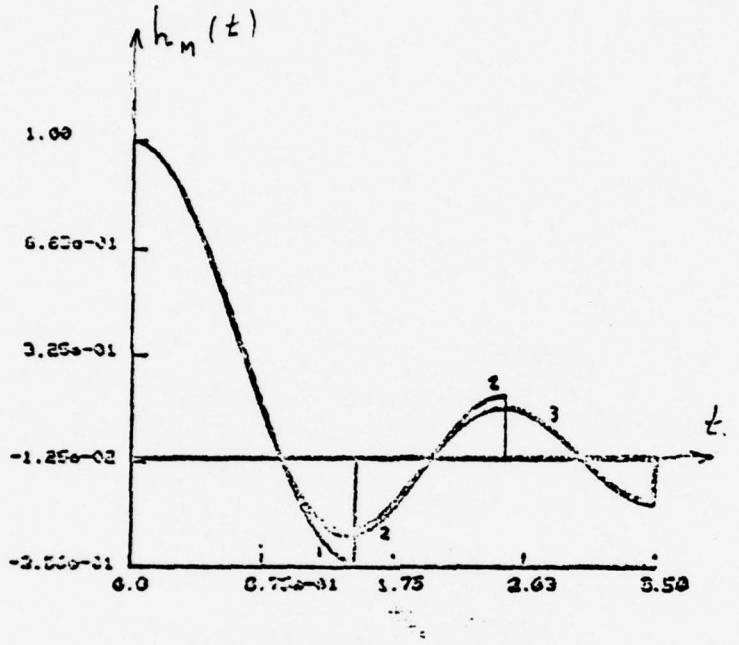


(a) Impulse response.

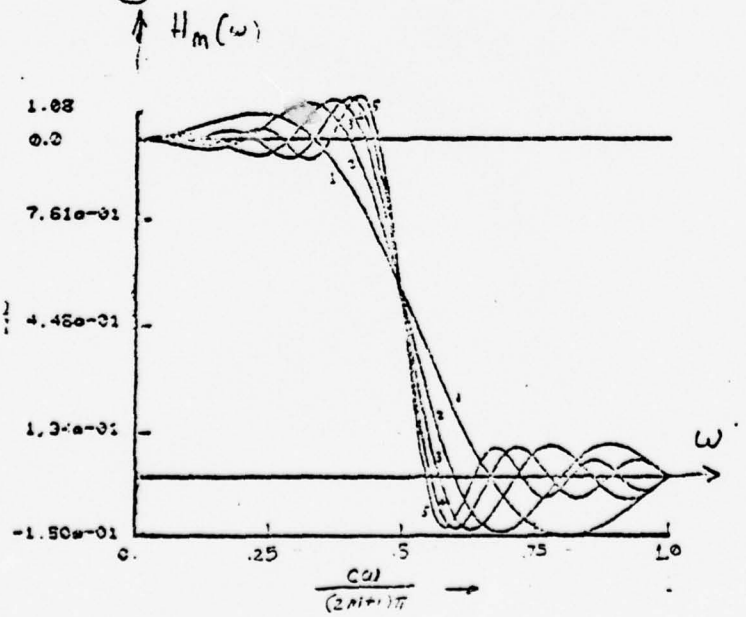


(b) Frequency response.

Fig. 4. Running low-pass filter.


$h_m(t)$
1.00
6.63e-01
3.25e-01
-1.25e-02
-3.50e-01
0.0
1.75
2.63
3.50
t
1
2
3

(a)


$H_m(\omega)$
1.08
0.0
7.61e-01
4.48e-01
1.34e-01
-1.50e-01
0
.25
.5
.75
1.0
$\frac{\omega}{(2m+1)\pi}$
ω

(b)

Fig. 5

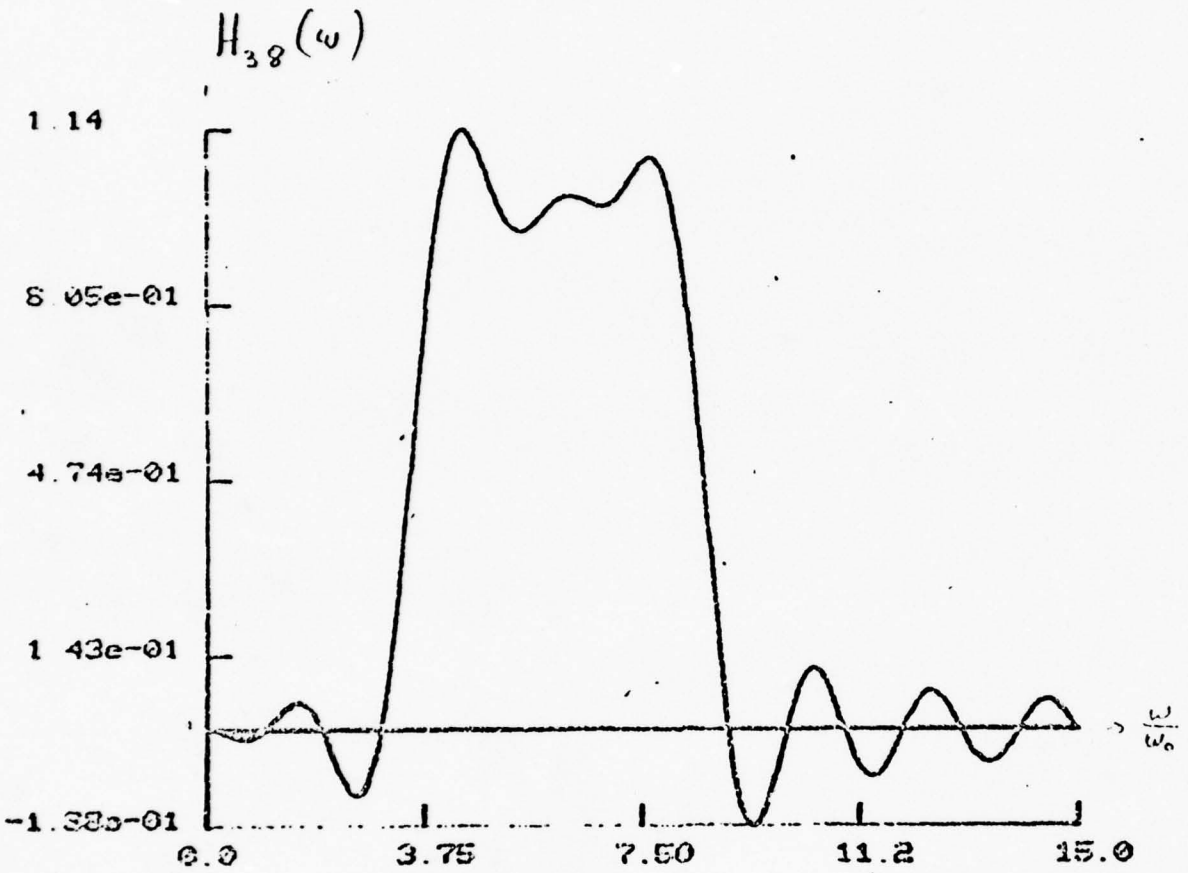

$H_{38}(\omega)$
1.14
8.05e-01
4.74e-01
1.43e-01
-1.33e-01
0.0
3.75
7.50
11.2
15.0
$\frac{\omega}{\omega_0}$

Fig. 6

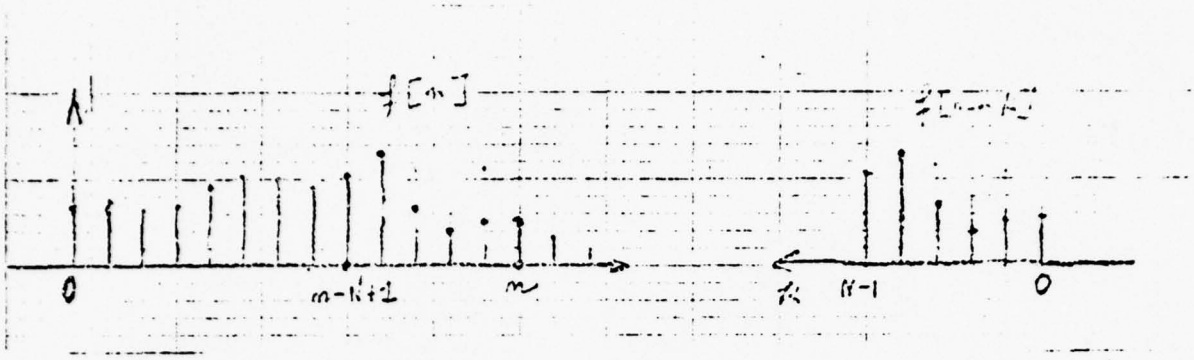


Fig. 7

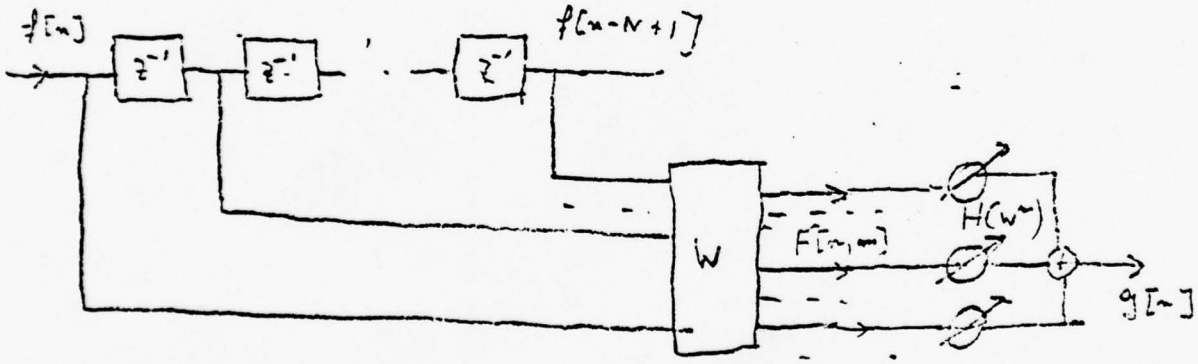


Fig. 8

Section 7

Adaptive Frequency Domain Estimators

7.1. Adaptive and recursive MS estimators

Adaptive estimators are of interest because they require no knowledge of prior statistics and can be used to process stationary and non-stationary signals [1], [2]. Recursion is desirable because it simplifies the required arithmetic operations. In this paper, we develop a new estimator that combines these advantages with the advantages of frequency domain processing.

To place our development into familiar perspective, we reexamine two basic methods of adaptive filtering using as illustration the estimation of a discrete process $y[n]$ in terms of the output

$$\hat{y}[n] = \sum_{k=0}^{N-1} a_k[n]\, x[n-k] \tag{1}$$

of a non-recursive time-varying system whose input is the data process $x[n]$. In vector notation,

$$\hat{y}[n] = X^T[n]\, A[n] \tag{2}$$

where $A[n]$ is the vector whose components are the N weights $a_k[n]$ and $X^T[n]$ is the transpose of the data vector $X[n]$ consisting of the N most recent samples $x[n-k]$ of $x[n]$. The resulting estimation error is given by

$$e[n] = y[n] - \hat{y}[n] = y[n] - X^T[n]\, A[n] \tag{3}$$

and our objective is to determine the vector A[n] so as to minimize in some sense this error.

The adaptive Wiener filter. If the processes x[n] and y[n] are jointly stationary and the optimality criterion is the minimization of the MS value of e[n], then the optimum A[n] is the solution of the system (orthogonality principle [3])

$$RA = \Gamma \tag{4}$$

where R is the correlation matrix of the data vector and Γ is the cross-correlation vector between the data and the signal y[n] to be estimated:

$$R = E\{ X[n] X^T[n] \} \qquad \Gamma = E\{ X[n] y[n] \} \tag{5}$$

The solution of (4), yields the familiar Wiener filter

$$A = R^{-1} \Gamma \tag{6}$$

This filter is optimum but it requires prior knowledge of the matrix R and the vector Γ.

For the estimation of a process with unknown statistics we can use a vector A[n] such that

$$\hat{R}[n] A[n] = \hat{\Gamma}[n] \tag{7}$$

where

$$\hat{R}[n] = (1-\alpha) \sum_{k=1}^{n-1} \alpha^k X[n-k]\, X^T[n-k]$$

$$0 < \alpha < 1 \qquad (8)$$

$$\hat{\Gamma}[n] = (1-\alpha) \sum_{k=1}^{n-1} \alpha^k X[n-k]\, y[n-k]$$

The solution of (7)

$$A[n] = \hat{R}^{-1}[n]\, \hat{\Gamma}[n] \qquad (9)$$

is the adaptive Wiener filter.

It can be shown (Appendix A), that under general conditions, the matrix $\hat{R}[n]$ and the vector $\hat{\Gamma}[n]$ tend to R and Γ respectively as $n \to \infty$ and $\alpha \to 1$. Hence, for sufficiently large n and for α close to one, the filter specified by A[n] is nearly optimum.

We note that the adaptive Wiener filter can also be used in the estimation of non-stationary processes provided that their statistics vary sufficiently slowly (see Appendix A).

For a given n, the output $\hat{y}[n]$ of the above filter is not optimum in the mean-square sense. It can be shown, however, that, if the optimality criterior is the minimization not of the mean-square value of the error but the minimization of its weighted time-average

$$\sum_{k=1}^{n-1} \alpha^k \, e^2[n-k] = \sum_{k=1}^{n-1} \alpha^k \left(y[n-k] - X^T[n-k]\, A[n-k] \right)^2 \qquad (10)$$

then $\hat{y}[n]$ is optimum for any n.

The realization of the vector $A[n]$ involves the evaluation of the vector $\hat{\Gamma}[n]$ and the matrix $\hat{R}[n]$ and its inverse for each n. The underlying computations can, however, be simplified if they are carried out recursively. Indeed, from (8) it follows readily that

$$\begin{aligned} \hat{R}[n+1] &= \alpha \hat{R}[n] + \hat{\alpha}(1-\alpha)\, X[n]\, X^T[n] \\ \hat{\Gamma}[n+1] &= \alpha \hat{\Gamma}[n] + \hat{\alpha}(1-\alpha)\, X[n]\, y[n] \end{aligned} \qquad (11)$$

To obtain a recursion for the solution $A[n]$ of (7), we note with (11) and (7) that

$$A[n+1] = \hat{R}^{-1}[n+1]\, \hat{\Gamma}[n+1] = \hat{R}^{-1}[n+1] \left\{ \alpha \hat{\Gamma}[n] + \hat{\alpha}(1-\alpha)\, X[n]\, y[n] \right\}$$

$$= \alpha \hat{R}^{-1}[n+1]\, \hat{R}[n]\, A[n] + \hat{\alpha}(1-\alpha)\, \hat{R}^{-1}[n+1]\, X[n]\, y[n]$$

$$= \hat{R}^{-1}[n+1] \left\{ \hat{R}[n+1] - \hat{\alpha}(1-\alpha)\, X[n]\, X^T[n] \right\} A[n]$$

$$+ \hat{\alpha}(1-\alpha)\, \hat{R}^{-1}[n+1]\, X[n]\, y[n]$$

and since

$$\hat{y}[n] = X^T[n]\, A[n] \qquad e[n] = y[n] - \hat{y}[n]$$

we conclude that

$$A[n+1] = A[n] + \alpha(1-\alpha)\, \hat{R}^{-1}[n+1]\, e[n]\, X[n] \tag{12}$$

We can similarly derive a recursion equation for the inverse matrix $\hat{R}^{-1}[n+1]$.

The Widrow filter. A simplified form of (12) is the algorithm

$$A[n+1] = A[n] + \mu\, e[n]\, X[n] \tag{13}$$

where μ is some constant that determines the time of adaptation. This is the basis of the Widrow filter [4], [5] and it is justified as an instantaneous version of gradient seeking least MS techniques [5] .

Unlike the adaptive Wiener filter, the Widrow filter is not in general optimum even asymptotically [6] , [7] . However, it is simple and in a number of applications it performs well [8] .

2. The adaptive frequency domain filter

Our objective is to improve the design of adaptive filters in the following respects:

1. Reduction of the number of adaptively controlled parameters.

In (1), the number of unknown parameters equals the length N of the memory of the filter. This constraint complicates in many cases unnecessarily

the design of the filter. We mention as a simple illustration the non-recursive realization of a one-pole system. In this case, the coefficients $a_k[n]$ approach a geometric progression z_o^k involving a single parameter. However, if z_o is close to one, the required length N of the non-recursive approximation is large.

2. Data vector with nearly diagonal correlation matrix.

The correlation matrix of the data can be simply inverted if it is diagonal. This is not, in general, true for the estimator in (1).

3. Design based on the local spectral properties of the data.

In a number of applications, the design of an estimator is simplified if its adaptively controlled parameters are directly related to the frequency domain characteristics of the data. For example, elimination of high frequency noise or low frequency clutter is simply achieved in the frequency domain. Again, this cannot be done with the estimator in (1) because each frequency component of the output depends on all the components $a_k[n]$ of the vector $A[n]$. Frequency domain processing, however, based on traditional methods, involves global spectral properties and is, therefore, unsuitable for local processing. To eliminate this difficulty, we shall introduce the concept of running transforms.

As a preparation, we impose the constraint that the vector $A[n]$ can be written as a product

$$A[n] = WB[n] \tag{14}$$

where $B[n]$ is a vector with 2M+1 components $b_m[n]$ and W is a 2M+1 by N matrix to be determined.

Inserting (14) into (2), we obtain

$$\hat{y}[n] = X^T[n]\, W\, B[n] \tag{15}$$

With

$$P[n] = W^T\, X[n] \tag{16}$$

(15) yields

$$\hat{y}[n] = P^T[n]\, B[n] \tag{17}$$

The matrix W must be so chosen as to lead to a small value of M without significant increase of the estimation error and to a data vector P[n] whose covariance matrix is nearly diagonal. These requirements can be satisfied if we use the 2M+1 significant components of the Karhunen-Loève expansion of X[n]. However, the underlying transformation matrix depends on the statistics of X[n] which we assumed unknown. Furthermore, in this approach the spectral properties of x[n] are not utilized directly.

In the course of this investigation, it will be shown that a matrix W whose elements w_{km} are the roots of unity:

$$w_{km} = w^{km} \qquad \text{where} \qquad w = e^{j2\pi/N} \tag{18}$$

has the desired properties.

The elements of the resulting data vector $P[n]$ will be denoted by $X[n, m]$. Inserting (18) into (16), we obtain

$$X[n, m] = \sum_{k=0}^{N-1} x[n-k]\, w^{km} \tag{19}$$

The sequence $X[n, m]$ so formed is the mth discrete Fourier series [9] (abbreviation: DFS) of the N most recent values of the process $x[n]$. It describes thus, the local spectral properties of $x[n]$.

The output $\hat{y}[n]$ of the resulting estimator is given by [see (17)]

$$\hat{y}[n] = \sum_{m=-M}^{M} X[n, m]\, b_m[n] \tag{20}$$

and it defines the proposed frequency domain filter. This filter consists of a memory-less system (Fig. 1) transforming the data vector $X[n]$ into the DFS vector $X[n, m]$ followed by $2M+1$ weights $b_m[n]$ and an adder. The weights $b_m[n]$ can be adaptively controlled to minimize in some sense the estimation error.

The advantage of this filter is the fact that we are operating directly on the local frequency components of the input. Furthermore, as we shall show [see (41) below], a significant data reduction results because in many cases, the required value of M is much less than N. Finally, the correlation matrix of $X[n, m]$ is nearly diagonal [see (42) below].

The disadvantage of the frequency domain filter is the need to compute the sum in (19). This, however, is not a serious problem. With the available FFT algorithms, the computations can be carried out routinely.

Furthermore, it is possible to compute $X[n,m]$ recursively using one multiplication only. Indeed, from (19) it follows readily that

$$X[n,m] = w^m X[n-1,m] + x[n] - x[n-N] \tag{21}$$

We discuss next the frequency domain form of the estimators introduced in Sec. 1 .

The frequency domain Wiener filter. We wish to determine the $2M+1$ components $b_m[n]$ of the vector $B[n]$ such that if the sum $\hat{y}[n]$ in (20) is used as the estimate of a process $y[n]$, the resulting MS error is minimum.

If the processes $x[n]$ and $y[n]$ are jointly stationary with known second order moments, then the optimum $B[n]$ is independent of n and it is given by (orthogonality principle)

$$B = R^{-1} \Gamma \tag{22}$$

where R is the correlation matrix of the running DFS $X[n,m]$ of the data $x[n]$, and Γ is the cross-correlation vector between $X[n,m]$ and the signal $y[n]$ to be estimated. Thus, the elements of R and Γ are independent of n and are given by

$$E\{ X[n,m] X^*[n,k] \} \tag{23}$$

and

$$E\{X[n,m]\, y[n]\} \tag{24}$$

respectively.

In the next section, we express R and Γ in terms of the moments of x[n] and y[n].

We note that, if 2M+1 =N , then the estimator (20) obtained with the vector B in (22) is identical with the estimator (1) obtained with the vector A in (6). As M decreases, the design is simplified without appreciable increase in the resulting M.S. error. In fact, if the process to be estimated is narrow-band, then a nearly optimum estimator requires a small number of terms in (20) [see (41) below]. No such simplification is possible with the estimator in (1).

We now assume that no prior statistics are known. In this case, the estimator is a time-varying system and the vector B[n] is given by

$$B[n] = \hat{R}^{-1}[n]\, \hat{\Gamma}[n] \tag{25}$$

In the above, $\hat{R}[n]$ is a matrix whose elements $r_{mk}[n]$ are computed from the recursion equation [see (11)]

$$r_{mk}[n+1] = \alpha r_{mk}[n] + (1-\alpha)\, X[n,m]\, X^*[n,k] \tag{26}$$

and $\hat{\Gamma}[n]$ is a vector whose elements $\gamma_m[n]$ are computed from the recursion equation

$$\gamma_m[n+1] = \alpha\gamma_m[n] + \overset{\frown}{(1-\alpha)}\, X[n,m]\, y[n] \tag{27}$$

Reasoning as in (12), we can show that the vector $B[n]$ satisfies the recursion equation

$$B[n+1] = B[n] + \overset{\frown}{\alpha}(1-\alpha)\, \hat{R}^{-1}[n+1]\, e[n]\, D[n] \tag{28}$$

where $e[n]$ is the estimation error and $D[n]$ is a vector with components $X[n,m]$.

The frequency domain Widrow filter. In this filter, the components $b_m[n]$ of $B[n]$ are such that

$$b_m[n+1] = b_m[n] + \mu_m e[n]\, X[n,m] \tag{29}$$

as in (13). Unlike (13), the coefficient μ_m in (29) is different for each m. This is so because it might be desirable to have a different adaptation time for each frequency component. We mention as example, the use of this filter for speech processing [10].

3. The running DFS

The running DFS of a discrete signal $f[n]$ is the sum

$$F[n,m] = \sum_{k=0}^{N-1} f[n-k]\, w^{km} \tag{30}$$

where N is a given integer. For a given n, F[n, m] is the DFS in the variable k of the segment f[n-k] of f[n] shown in Fig. 2.

Properties. Suppose, first, that f[n] is a deterministic signal with z-transform

$$F(z) = \sum_{n=-\infty}^{\infty} f[n]\, z^{-n} \tag{31}$$

In this case, the sequence F[n, m] has a z-transform in the variable n:

$$F(z,m) = \sum_{n=-\infty}^{\infty} F[n,m]\, z^{-n} \tag{32}$$

From (30) and (32) it follows that

$$\overline{F}(z,m) = \sum_{k=0}^{N-1} z^{-k}\, F(z) w^{km} = F(z)\, \frac{1-(w^m/z)^N}{1-w^m/z} \tag{33}$$

because the z-transform of f[n-k] equals $z^{-k}\, F(z)$.

This shows that the sequence F[n, m] is the output of a linear system with input f[n] and system function

$$J(z,m) = \frac{1-z^{-N}}{1-w^m/z} \tag{34}$$

(Fig. 3a) because $w^N = 1$.

We, next, assume that f[n] is a stationary random process with autocorrelation

$$R[k] = E\{ f[n+k]\, f[n] \} \tag{35}$$

and power spectrum

$$S(\omega) = \sum_{k=-\infty}^{\infty} R[k]\, e^{-jk\omega T} \tag{36}$$

In this case, F[n, m] is also a random process, stationary in the variable n and non-stationary in the variable m. We shall determine its autocorrelation

$$R_{FF}[k, m] = E\{ F[n+k, m]\, F^{*}[n, m] \} \tag{37}$$

and power spectrum

$$S_{FF}(\omega, m) = \sum_{k=-\infty}^{\infty} R_{FF}[k, m]\, e^{-jk\omega T} \tag{38}$$

Since F[n, m] is the output of the filter of Fig. 3a with input f[n], it follows that [9]

$$S_{FF}(\omega, m) = S(\omega) \left| J(e^{j\omega T}, m) \right|^2 = S(\omega) \frac{\sin^2 \frac{N\omega T}{2}}{\sin^2 (\frac{\omega T}{2} - \frac{\pi m}{N})} \tag{39}$$

The inverse transform of the above yields

$$R_{FF}[k,m] = \sum_{r=-N+1}^{N-1} (N-|r|)\, R[r+k]\, w^{rm} \tag{40}$$

Equation (39) justifies our claim that frequency domain filtering leads to data reduction. Indeed, applying the inversion formula, we obtain

$$E\{\,|F[n,m]|^2\} = \frac{1}{2\pi}\int_{-\sigma}^{\sigma} S(\omega)\,\frac{\sin^2\frac{N\omega T}{2}}{\sin^2\left(\frac{\omega T}{2}-\frac{\pi m}{N}\right)}\,d\omega \tag{41}$$

where

$$\sigma = \frac{\pi}{T}$$

This shows that, if the bandwidth of $S(\omega)$ equals B, then $E\{|F[n,m]|^2\}$ takes significant values only for m of the order of BN/σ (Fig. 4). If, as it is often the case, $f[n]$ equals the samples $f(nT)$ of a continuous-time process $f(t)$ with power spectrum $S_c(\omega)$ and T is sufficiently small, then

$$S(\omega) = S_c(\omega) \qquad \text{for} \qquad |\omega| < \sigma = \pi/T$$

(no aliasing). If, therefore, T is sufficiently small, then the number of the significant components of $F[n,m]$ is small.

We shall next show that

$$E\{F[n,m_1]F^*[n,m_2]\}$$

$$= \frac{w^{(m_2-m_1)/2}}{2\pi} \int_{-\sigma}^{\sigma} S(\omega) \frac{\sin\frac{N\omega T}{2}}{\sin\left(\frac{\omega T}{2}-\frac{\pi m_1}{N}\right)} \frac{\sin\frac{N\omega T}{2}}{\sin\left(\frac{\omega T}{2}-\frac{\pi m_2}{N}\right)} d\omega \qquad (42)$$

As we have shown, the processes $F[n,m_1]$ and $F[n,m_2]$ are the outputs of the two systems of Fig. 3b with common input $f[n]$ and frequency responses

$$\frac{1-e^{-jN\omega T}}{1-w^{m_1}e^{-j\omega T}} \quad \text{and} \quad \frac{1-e^{-jN\omega T}}{1-w^{m_2}e^{-j\omega T}} \qquad (43)$$

respectively. Therefore, the cross-power spectrum of the processes $F[n,m_1]$ and $F[n,m_2]$ equals $S(\omega)$ multiplied by the product of the first fraction in (43) times the conjugate of the second. Equation (42) follows, thus, from the inversion formula.

It is easy to see that the fraction in the integrand of (42) decreases rapidly as $|m_1-m_2|$ increases. This justifies our second claim that the processes $F[n,m_1]$ and $F[n,m_2]$ are loosely correlated if $|m_1-m_2|$ is large compared to one.

Filtering. We show next that an arbitrary non-recursive filter of length N can be simply realized in terms of the running DFS $F[n,m]$ of the input $f[n]$. As a preparation, we establish, first, the inversion formula for running transforms.

We maintain that

$$f[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[n, m] \tag{44}$$

Proof. Since $F[n, m]$ is the DFS of the segment $f[n-k]$ of $f[n]$, it follows from the inversion formula for DFS that [9]

$$f[n-k] = \frac{1}{N} \sum_{m=0}^{N-1} F[n, m]\, w^{-km} \qquad 0 \le k \le N-1 \tag{45}$$

and (44) results with $k=0$.

We now assume that $f[n]$ is the input to a non-recursive filter with delta response $h[n]$ and system function

$$H(z) = \sum_{n=0}^{N-1} h[n]\, z^{-n} \tag{46}$$

We shall express the resulting response

$$g[n] = \sum_{k=0}^{N-1} f[n-k]\, h[k] \tag{47}$$

in terms of $F[n, m]$. Inserting (45) into (47) and changing the order of summation, we obtain

$$g[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[n, m] \sum_{k=0}^{N-1} h[k]\, w^{-km}$$

Hence [see (46)]

$$g[n] = \frac{1}{N} \sum_{m=0}^{N-1} H(w^m)\, F[n, m] \tag{48}$$

This result shows that an arbitrary filter with system function $L(z)$ can be approximated by a non-recursive filter $H(z)$ such that

$$H(e^{j\omega T}) = L(e^{j\omega T}) \text{ for } \qquad \omega = 2\pi m/N$$

It is, thus, an interpolating filter [12], [13], interpolating the desired frequency response at N equidistant points.

The running low-pass filter. A special case of (44) leads to a simple design of a low-pass, high-pass, or band-pass filter. We shall discuss only the low-pass case. The result is a refinement of a familiar method of filtering (smoothing) used to estimate a slowly varying signal contaminated by rapidly varying noise [14].

We first observe that, since

$$H(w^{m+N}) = H(w^m) \qquad\qquad F[n, m+N] = F[n, m]$$

the limits in (48) can be changed appropriately.

The running low-pass filter is a non-recursive filter such that

$$H(w^m) = \begin{cases} 1 & |m| \le M \\ 0 & \text{otherwise} \end{cases} \tag{49}$$

We shall determine its delta response $h[n]$ and frequency response $H(e^{j\omega T})$. Inserting (49) into (48), we obtain

$$g[n] = \frac{1}{N} \sum_{m=-M}^{M} F[n,m] = \frac{1}{N} F[n,0] + \frac{2}{N} \sum_{m=1}^{M} \text{Re } F[n,m] \tag{50}$$

From the above and (30) it follows that

$$g[n] = \frac{1}{N} \sum_{m=-M}^{M} \sum_{k=0}^{N-1} f[n-k]\, w^{km} = \frac{1}{N} \sum_{k=0}^{N-1} f[n-k] \sum_{m=-M}^{M} w^{km}$$

Hence,

$$h[n] = \frac{1}{N} \sum_{m=-M}^{M} w^{nm} = \frac{w^{(M+1)n} - w^{-Mn}}{N(1-w^n)} = \frac{\sin(M+\frac{1}{2})\frac{2\pi n}{N}}{N \sin\frac{2\pi n}{N}} \tag{51}$$

for n between 0 and N-1 and $h[n] = 0$ otherwise.

The system function is given by

$$H(z) = \sum_{n=0}^{N-1} \left(\frac{1}{N} \sum_{m=-M}^{M} w^{nm} \right) z^{-n} = \frac{1}{N} \sum_{m=-M}^{M} \frac{1-z^{-N}}{1-w^m/z} \tag{52}$$

because $w^N = 1$.

As an illustration of the running low-pass filter, we mention the adaptive estimation of a process $y[n]$ in terms of the data

$$f[n] = y[n] + \nu[n] \tag{53}$$

containing the noise process $\nu[n]$ which we assume white. If we use as an

estimate of $y[n]$ the output $g[n]$ of the low-pass filter with $M=0$, we obtain [see (50)]

$$g[n] = \frac{1}{N} F[n, 0] = \frac{1}{N} \sum_{k=0}^{N-1} f[n-k] \tag{54}$$

This is the average of the N most recent values of the data. The resulting MS estimation error depends on N. The component of $g[n]$ due to the noise decreases if N is large, however, the distortion of $y[n]$ (bias) due to smoothing increases. To minimize the MS error, we must vary N adaptively [14] taking into consideration the local properties of $y[n]$ and $\nu[n]$. The running low-pass filter permits us to determine these properties and to very the effective width of $h[n]$ by varying M. Furthermore, it improves the estimation because it uses not only the average, but also the first M harmonics of the segment $f[n-k]$ of $f[n]$.

7.4 Numerical Results

We apply the adaptive frequency domain filter to the problem of estimating a signal $y[n]$ in terms of the data

$$x[n] = y[n] + \nu[n]$$

shown in Fig. 5a. The unknown signal $y[n]$ is generated as the output of a fifth order low-pass digital filter (Butterworth) driven by a white noise process $\zeta[n]$ (Fig. 5b). The noise component $\nu[n]$ is white and the signal-to-noise ratio measured as the ratio of the corresponding energies in the observation interval equals 2DB.

The output $\hat{y}[n]$ of the estimator of $y[n]$ is the sum

$$\hat{y}[n] = \sum_{m=-M}^{M} X[n, m]\, b_m[n] \tag{55}$$

as in (20) where $X[n, m]$ is the mth running DFS obtained with FFT size $N=32$. The value of M is determined by establishing a threshold level T and accepting only the components of $X[n, m]$ that exceed T. The value of T is chosen either by prior knowledge of the r.m.s. level of the noise $\nu[n]$ or by a measurement of this level. Since $\nu[n]$ is white by assumption, it suffices to measure the components of $X[n, m]$ for sufficiently large m. The results of the measurement are shown in Fig. 5.

In Fig. 7 we show the value of M so obtained. As we see from the figure, after some initial interval, M varies between 1 and 4.

The coefficients $b_m[n]$ are adaptively controlled as in (29) and the output $\hat{y}[n]$ of the filter is shown in Fig. 8.

For comparison, we show in Fig. 9 the estimate obtained with the time-domain widrow filter (13). As we see from the figure, the estimate is not as accurate as the estimate of Fig. 8 although the number of adaptively controlled coefficients is considerably larger.

Appendix A

We wish to estimate the mean of a discrete random process $x[n]$ in terms of the weighted average

$$\bar{x}[n] = (1-\alpha) \sum_{k=0}^{n-1} \alpha^k x[n-k] \qquad 0 < \alpha < 1 \qquad \text{(A-1)}$$

of its past values $x[n-k]$.

We assume, first, that the process $x[n]$ is stationary with mean η and autocorrelation $R[k]$:

$$E\{ x[n] \} = \eta \qquad E\{ x[n+k]\, x[n] \} = R[k] \qquad \text{(A-2)}$$

From (A-1) it follows readily that

$$E\{ \bar{x}[n] \} = (1-\alpha^n)\, \eta \qquad \text{(A-3)}$$

$$E\{(\bar{x}[n] - \eta)^2\} = \frac{1-\alpha}{1+\alpha} \sum_{k=-n+1}^{n-1} \alpha^{|k|} \left[1-\alpha^{2(n-|k|)}\right] R[k] \qquad \text{(A-4)}$$

Equations (A-3) and (A-4) show that, as $n \to \infty$, the mean of $\bar{x}[n]$ tends to the mean of $x[n]$ and the variance of $\bar{x}[n]$ tends to

$$\frac{1-\alpha}{1+\alpha} \sum_{k=-\infty}^{\infty} \alpha^{|k|} R[k] \qquad \text{(A-5)}$$

Therefore, if $R[k]$ is absolutely summable, i.e., if

$$I = \sum_{k=-\infty}^{\infty} |R[k]| < \infty \qquad \text{(A-6)}$$

then the variance of $\bar{x}[n]$ tends to zero as $n \to \infty$ and $\alpha \to 1$.

From the above it follows that, if the constant α and the integer n_o are such that

$$(1-\alpha)\, I << 2\eta^2 \quad \text{and} \quad \alpha^{n_o} << 1 \qquad \text{(A-7)}$$

then the mean of $\bar{x}[n]$ is nearly η and its variance is small compared to η^2. Hence, $\bar{x}[n]$ is a satisfactory estimator of the mean of $x[n]$ provided that n is larger than n_o.

We note that the time-average $\bar{x}[n]$ can be determined recursively:

$$\bar{x}[n+1] = \alpha\, \bar{x}[n] + (1-\alpha)\, x[n] \qquad \text{(A-8)}$$

with the initial condition $\bar{x}[1] = x[1]$. This follows readily from (A-1).

We conclude with the observation that $\bar{x}[n]$ can be used to estimate the mean

$$E\{ x[n]\} = \eta[n]$$

of a non-stationary process provided that $\eta[n]$ is nearly constant in any interval of length n_o.

References

[1] R.W. Lucky, "Techniques for Adaptive Equalization of Digital Communication Systems", Bell Syst. Tech. J., vol. 48, pp. 255-286, Feb. 1966.

[2] J.G. Proakis and J.H. Miller, An Adaptive Receiver for Digital Signaling through Channels with Intersymbol Interference", IEEE Trans. Inform. Theory, vol. IT-15, pp. 484-497, July 1969.

[3] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

[4] B. Widrow and M.E. Hoff, "Adaptive Switching Circuits", 1960 WESCON Conf. Rev. Pt. 4, pp. 96-140.

[5] B. Widrow et al, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", Proceedings of the IEEE, vol. 64, no. 8, pp. 1151-1162, 1976.

[6] L.D. Davisson, "Steady-State Error in Adaptive Mean-Square Minimization", IEEE Trans. Inform. Theory, vol. IT-19, pp. 382-385, July 1970.

[7] Jac-Kyoon Kim and L.D. Davisson, "Adaptive Linear Estimation for Stationary M-Dependent Processes", IEEE Trans. Inform. Theory, vol. IT-21, no. 1, pp. 23-31, January 1975.

[8] L.J. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency", IEEE Trans. on Acoustics, vol. ASSP-28, no. 2, pp. 207-222, 1975.

[9] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.

[10] J.B. Allen, D.A. Berkley and J. Blauert, "Multimicrophone Signal Processing Technique to Remove Room Reverbarations from Speech Signals", J. Acoust. Soc. Am., vol. 62, no. 4, pp. 912-915, 1977.

[11] J.B. Allen and L.R. Rabiner, "A Unified Theory of Short-Time Spectrum Analysis and Synthesis", Proc. IEEE, November 1977.

[12] L.R. Rabiner and R.W. Shafer, "Recursive and Nonrecursive Realizations of Digital Filters Designed by Frequency Sampling Techniques, IEEE Trans. on Audio and Electroacoustics, vol. AU-19, no. 3, pp. 200-207, September 1971.

[13] A.V. Oppenheim and R.W. Shafer, Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[14] A. Papoulis, "Two-to-One Rule in Data Smoothing", IEEE Trans. Inform. Theory, vol. IT-23, pp. 631-633, September 1977.

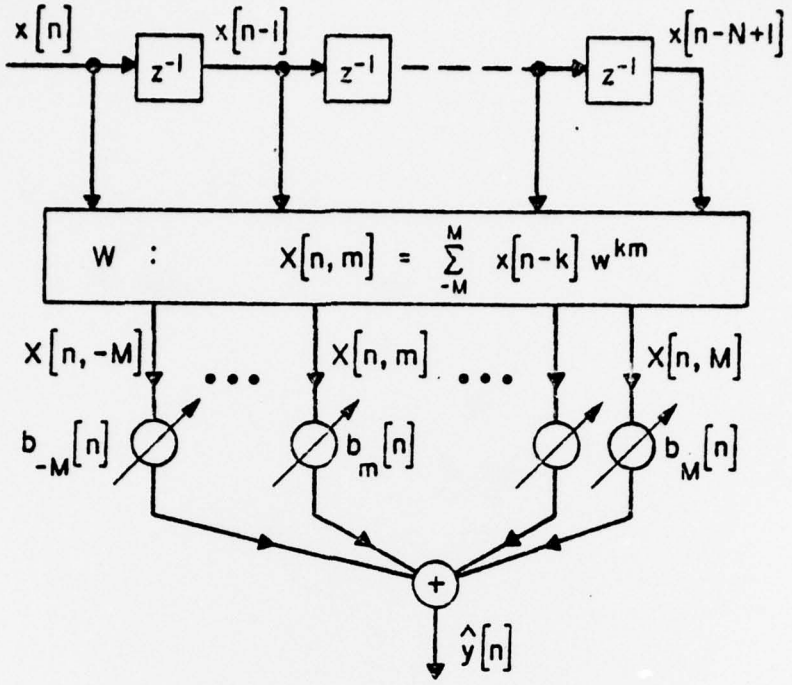


Figure 1

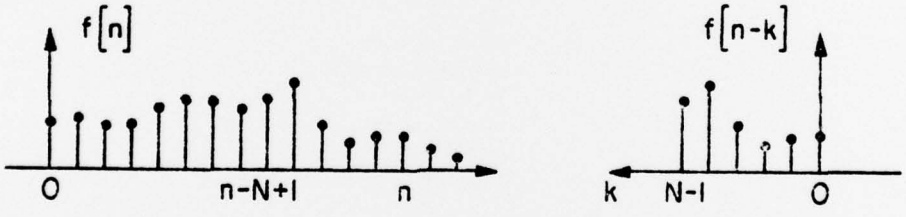


Figure 2

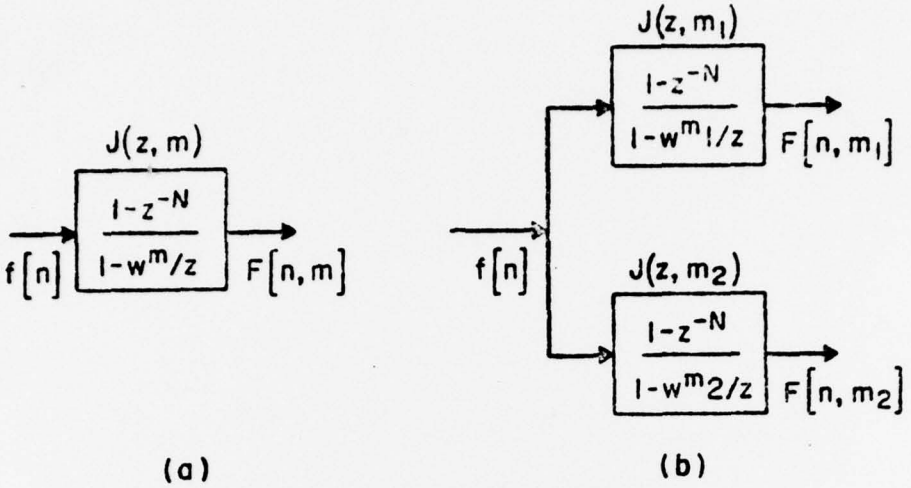


Figure 3

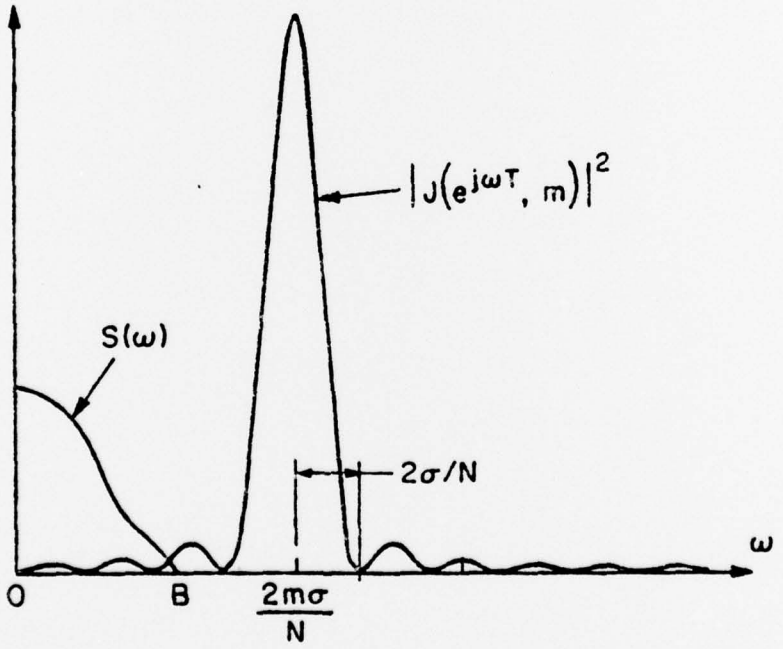


Figure 4

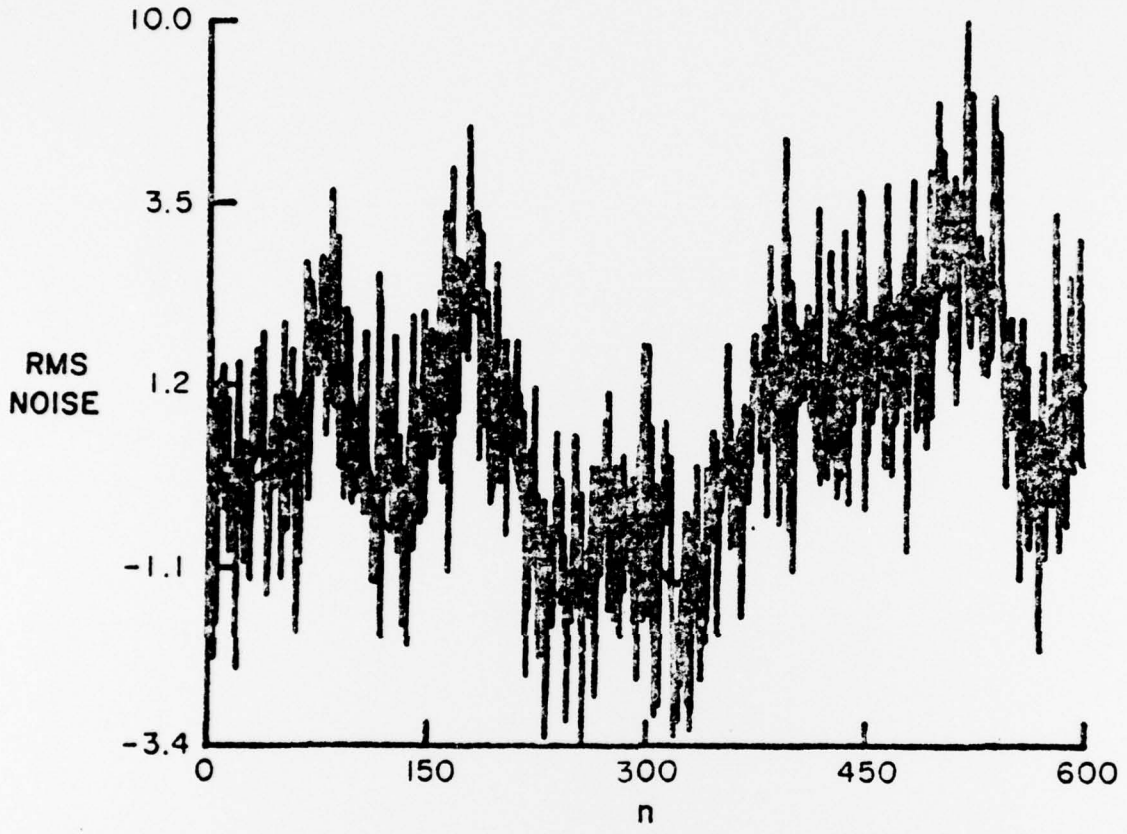


Figure 5(a)

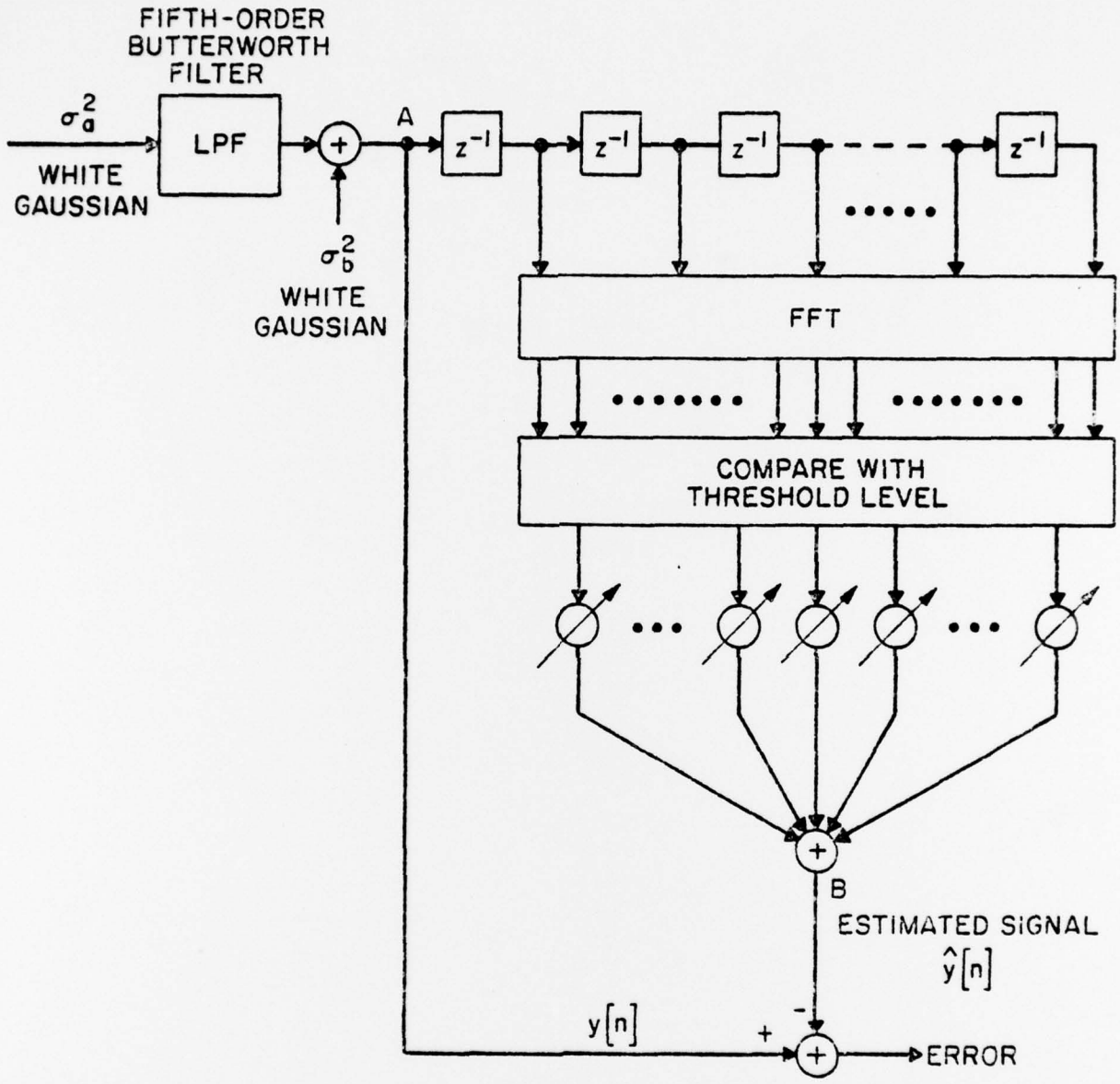

FIFTH-ORDER
BUTTERWORTH
FILTER
σ_a^2
WHITE
GAUSSIAN
LPF
A
z^{-1}
z^{-1}
z^{-1}
z^{-1}
σ_b^2
WHITE
GAUSSIAN
FFT
COMPARE WITH
THRESHOLD LEVEL
B
ESTIMATED SIGNAL
$\hat{y}[n]$
$y[n]$
+
-
ERROR

Figure 5(b)

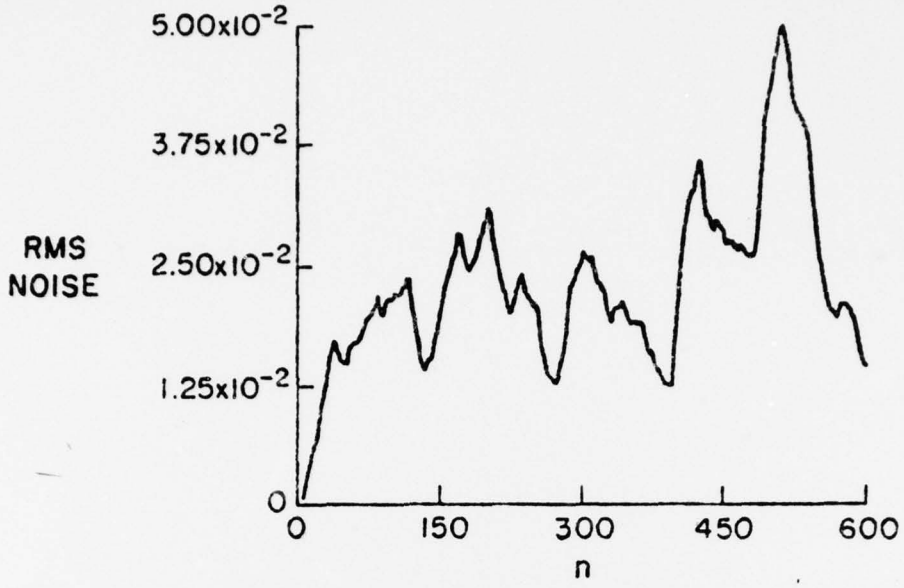


Figure 6

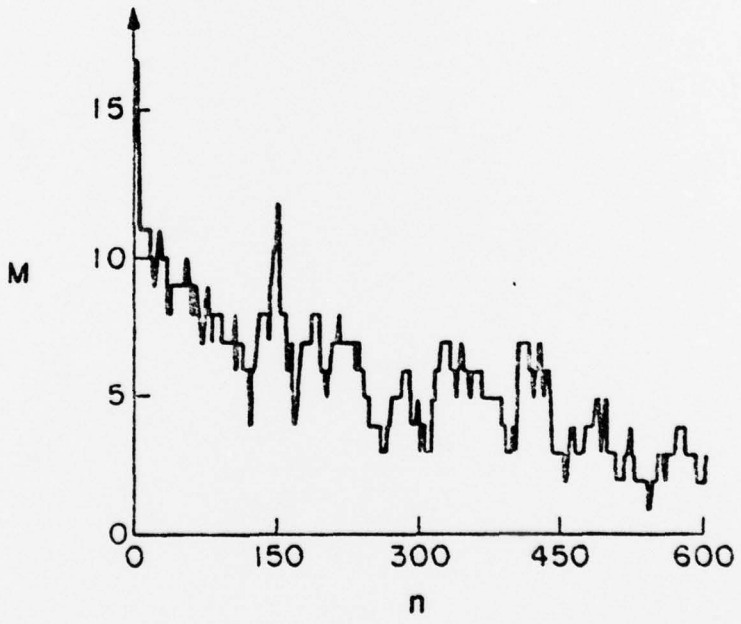


Figure 7

M ADAPTIVELY CONTROLLED

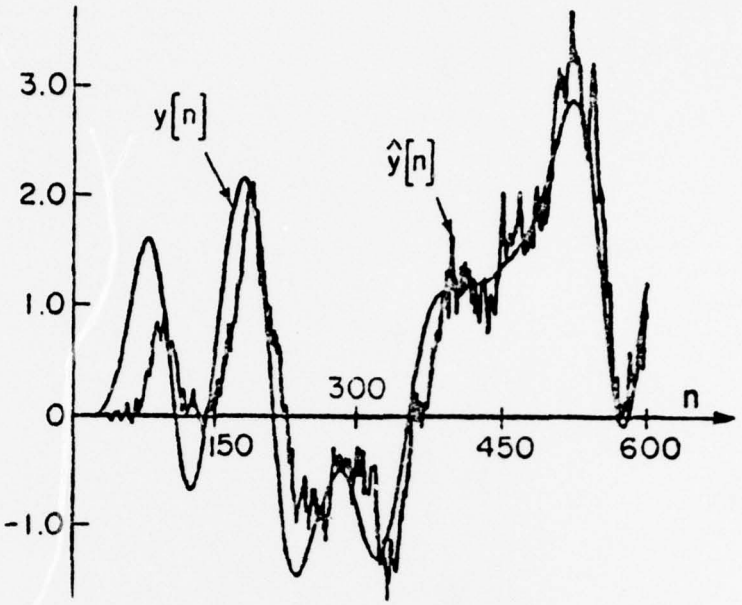


Figure 8

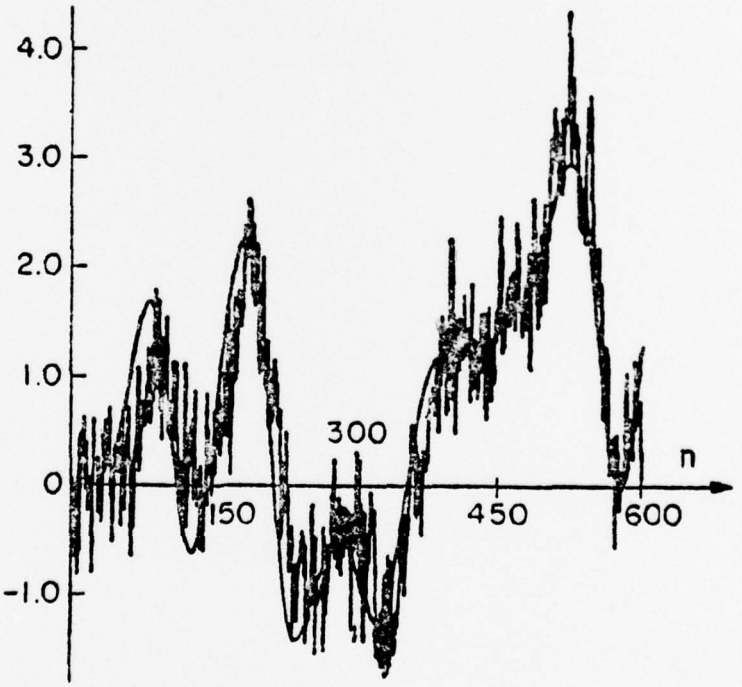


Figure 9