FILTER THEORY OF LINEAR OPERATORS WITH SEISMIC APPLICATIONS

by

MARKWICK KERN SMITH, JR.
S.B., Massachusetts Institute of Technology
(1951)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1954

Signature of Author ....
Department of Geology and Geophysics
January 11, 1954

Certified by ....
Thesis Supervisor

Accepted by ....
Chairman, Departmental Committee
on Graduate Students

ABSTRACT

FILTER THEORY OF LINEAR OPERATORS WITH SEISMIC APPLICATIONS

by

Markwick Kern Smith, Jr.

Submitted to the Department of Geology and Geophysics on January 11, 1954, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

The discrete linear operator is discussed as a linear computational technique parallel to electrical filtering. The discrete notation is developed from the continuous linear operator and the special case of the prediction operator is considered. The concept of linear operator dimensions is introduced and the generalized or n-dimensional linear operator is suggested.

The filter characteristics of the linear operator are developed from several different approaches. Both the transfer function and the impulse response are considered. The power transfer function and the transfer function for the prediction operator are also obtained. The concept of two dimensional spectra and filter characteristics is considered.

The restrictions on the transfer function of the linear operator are considered. It is found that the real part of the transfer function is even and the imaginary part is odd. Also, under certain conditions the real part is independent of the imaginary part. The restrictions on the transfer function of a physically realizable, stable network are compared with the transfer function of the linear operator. It is determined that the linear operator and the linear electric filter are essentially equivalent.

Criteria for determining the optimum linear operator or filter are developed from assumptions as to the nature of the input data. Several illustrations are given in a controlled seismic experiment.

Thesis Supervisor: Patrick M. Hurley
Title: Professor of Geology

TABLE OF CONTENTS

Page

Acknowledgments

The author wishes to express his gratitude to Professor P. M. Hurley, Thesis Supervisor, and Professor G. P. Wadsworth of the Department of Mathematics for their help and encouragement in the preparation of this thesis.

A major portion of the work of this thesis was done under the supervision of the Geophysical Analysis Group in the Department of Geology and Geophysics. The author is particularly indebted to E. A. Robinson and S. M. Simpson for their suggestions and criticisms, and to Irene Calnan and Barbara Halpern for their assistance in the computational work and preparation of this manuscript.

The author also wishes to thank Alice Jeghelian for typing the final copy.

I. Introduction

The problem of separating desired information from background interference is of major importance in the field of Geophysics, just as it is in many other fields. If a seismogram contained nothing but distinctly marked reflections, and the desired information consisted of the arrival times of this reflected energy, then the problem would be relatively simple. However, not only is the reflection energy at least partially masked in background interference or noise, but the requirements on the desired information are becoming increasingly complex. For example, in a simple geologic area, where only the subsurface structural features are of interest, the arrival times of reflected energy might be sufficient desired information. As the type of area becomes more complex, the desired information may again be the arrival times of reflected energy, but in this case, only certain reflections will aid in the geologic interpretation of the record, while others will tend to confuse it. In the event one wished to obtain more detailed information on the subsurface stratigraphy, the desired information might consist of the frequency spectrum of certain wavelets in the record.

A corresponding problem occurs in the interpretation of raw data obtained by other geophysical prospecting methods. The desired information, in the case of gravity data, might consist of the residual effect of local ore bodies, while the background interference would be regional and topographic effects plus anomalies due to local density variations.

Desired magnetic data is similarly masked in regional effects, diurnal variations, and fluctuations due to the instability of the instrument.

Many methods, in a wide variety of fields have been devised to separate desired information from background interference. Much work, of a more general nature, has been done on this problem in the fields of Statistics and Electrical Engineering. Recently, there have been a number of papers based on the mathematical theory of smoothing and prediction as developed by Wiener [1/] and Kolmogoroff[2/].

There are three common methods of operating on a set of information or data in an attempt to select the desired information. Though these techniques differ mechanically, they are fundamentally equivalent. The first technique consists of direct observation of the set of information followed by mental analysis. This technique is applicable to information presented in either a discrete or continuous form. The second technique consists of operating on continuous information by means of electrical or mechanical circuits. The third technique consists of operating on discrete information by means of mathematical operations. Though the last two techniques require the information to be presented in a particular form, it is generally possible to transform discrete information into continuous information and vice versa.

Although these three techniques are basically equivalent, they differ in precision, flexibility, speed, ease of design, and stability. For most problems a combination of the first technique and one of the others, proves to be the most efficient system. In certain cases a single technique or a combination of all three will be best. Because the precision, stability, and even the speed are usually greater in the mathematical and electrical or mechanical operations, it is often desirable to put as much of the burden as possible on these techniques. However, flexibility and design restrictions usually necessitate the final stage being direct observation and mental analysis. Further, since the theory of linear systems is more fully developed than nonlinear theory, the ease of design restriction often limits the electrical or mechanical techniques to linear operations. This is also true in the case of the mathematical technique when applied to time series problems. However, this does not imply that nonlinear operations are any the less desirable in general.

In a recent paper Wadsworth et. al.[3/] presented a linear mathematical technique with applications to the detection of reflections on seismic records. Knowledge of this method, which employs the concept of the discrete linear operator, will be assumed for the remainder of the present paper.

The purpose of the present paper is to treat the discrete linear operator as a general linear mathematical technique, and to demonstrate the relationship between this mathematical operation and the corresponding linear electrical operation. The possibility of extending the current linear mathematical or electrical operations to assume a greater share of the work in the interpretation of seismic records will be indicated. Further, methods of determining the linear operator from the input information and from other information and assumptions will be discussed in relation to seismic records. Application of certain of these methods will be demonstrated in a controlled experiment.

The properties of the discrete linear operator will be developed from the definition, but wherever appropriate, the relationship between the continuous and discrete cases will be demonstrated. Terminology and notation will, in general, be explained where it first occurs. Certain standard notations and definitions will be used without explanation in the context, but will usually be found in Appendix I. The term "signal" will be used throughout to mean desired information. The term "noise" will mean background disturbance. The "noise" will in general consist of a predictable element referred to as "interference", and a random element referred to as "random noise". The terms "trace," "input," "data," "perturbed signal", and "information" will refer, in general, to a combination of the signal and noise. It will be assumed throughout most of this paper that the combination of signal and noise is linear.

II. The Generalized Linear Operator

2.1 Review of Information Theory and Linear Operator Concepts

The problem to be considered may be formulated as follows. Given a perturbed signal $x(t)$, which is the sum of a true signal $s(t)$ and a perturbing noise $n(t)$, we wish to operate on $x(t)$ in such a manner as to obtain the best approximation $e(t)$ to the signal $s(t)$. That is

$$x(t) = s(t) + n(t) \qquad 2.1$$

and

$$e(t) = L[x(t)] \qquad 2.2$$

where $e(t)$ is the best approximation to $s(t)$, or more generally $s(t+\alpha)$, and L represents an operation. At this point the criterion of best approximation and the form of the operation are unspecified. For the case considered in this paper, the first restriction put on L is that it must be a linear operation on the available information. Thus if $x(t)$ is known from $-\infty$ to $+\infty$ we may write

$$e(t) = \int_{-\infty}^{\infty} x(t-\tau) \; A(\tau) \; d\tau \qquad 2.3$$

where $A(\tau)$ is a weighting function.

If, on the other hand, we know $x(t)$ only from $-\infty$ to 0, $e(t)$ may be written

$$e(t) = \int_{0}^{\infty} x(t-\tau) \, A(\tau) \, d\tau \qquad 2.4$$

If $A(\tau)$ is independent of time, then equations (2.3) and (2.4) represent a large group of linear operations on the available information that is invariant under a translation in time.

If, as in the case of a seismogram, $x(t)$ is known only between finite limits, but $e(t)$ may be determined from $x(t)$ at times greater than t as well as less than t, then we must use the special case of equation (2.3)

$$e(t) = \int_{-t_M}^{t_n} x(t-\tau)\, A(\tau)\, d\tau \;, \qquad 2.5$$

where t_M and t_n are positive. It may be argued, in the case of a seismogram, that we actually know $x(t)$ back to $t = -\infty$ and $x(t)$ could be recorded to $t = +\infty$. However, later discussion of the application of linear operators will indicate that this information would add very little to the solution of the problem. Further, from a computational point of view we must restrict ourselves to the finite case.

For purposes of computation it is convenient to approximate equation (2.5) by the discrete formula derived in the following manner. Let us approximate the function $x(t)$ by a series of positive and negative rectangular pulses of width h such that the mid point of the top of each pulse lies on the curve $x(t)$. See figure 1. Let the amplitude of each pulse be designated by x_i where i takes on integral values and ih represents discrete values of t at the center of each pulse. Therefore $x_{ih} = x(t)$ at $t = ih$. Equation (2.5) may then be approximated by

$$e(t) \doteq x_{i+m} \int_{(-M-\frac{1}{2})h}^{(-M+\frac{1}{2})h} A(\tau)d\tau + \ldots + x_i \int_{-\frac{h}{2}}^{\frac{h}{2}} A(\tau)d\tau + \ldots + x_{i-n} \int_{(n-\frac{1}{2})h}^{(n+\frac{1}{2})h} A(\tau)d\tau \qquad 2.6$$

where $-Mh$ is the closest integral multiple of h to $-t_M + \frac{1}{2}h$ and nh is the closest integral multiple of h to $t_n - \frac{1}{2}h$. If we then define

$$a_s = \int_{(s-\frac{1}{2})h}^{(s+\frac{1}{2})h} A(\tau)d\tau \qquad 2.7$$

equation 2.6 may be written

$$e(t) \doteq \sum_{s=-M}^{n} a_s x_{i-s}. \qquad 2.8$$

As a final step in passing to the discrete case, let us define the right hand side of the approximation (2.8) equal to e_i. Therefore $e_i \doteq e(t)$ at $ih = t$, where

$$e_i = \sum_{s=-M}^{n} a_s x_{i-s}. \qquad 2.9a$$

Equation (2.9a) is completely equivalent to

$$e_{i+k} = \sum_{s=o}^{m} a'_s \, x_{i-s} \qquad 2.9b$$

where $k = -M$, $a'_s = a_{s-m}$, and $m = n+M$. For the rest of this paper, except where otherwise noted, the term linear operator will refer to the discrete case as defined by equations (2.9). The form (2.9b) will predominate.

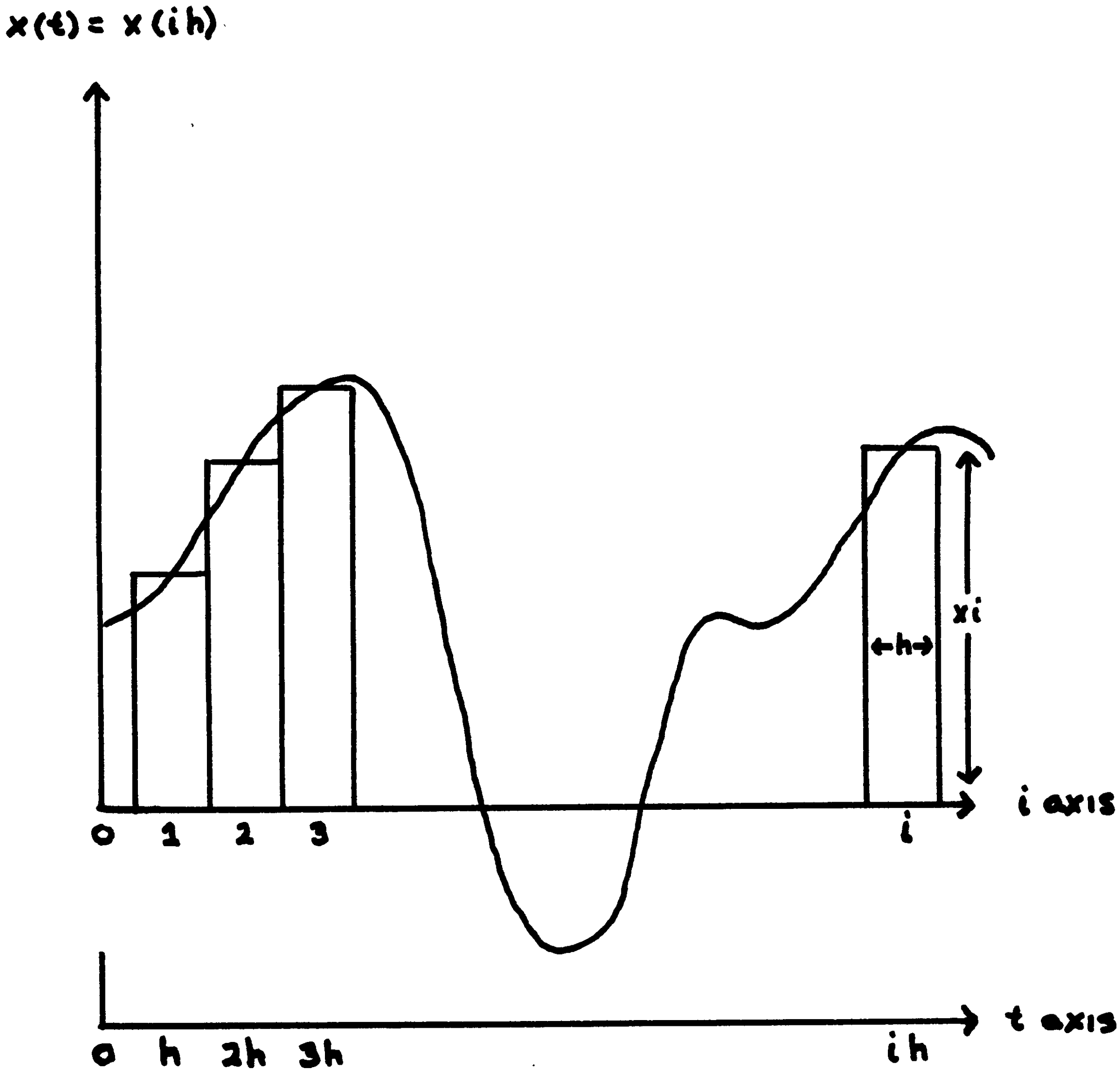


FIGURE 1

2.2 The One Dimensional Linear Operator

The one dimensional linear operator will be defined by equation (2.9b),

$$e_{i+k} = \sum_{s=o}^{m} a_s x_{i-s} \qquad 2.9b$$

where the a_s are known as the operator coefficients, x_i is the information or data, and e_{i+k} is the output or the approximation to the signal. It is often convenient for computational purposes to adjust the mean value of e_{i+k} by the addition of a constant term. Thus a more general form of (2.9b) is

$$e_{i+k} = a + \sum_{s=o}^{m} a_s x_{i-s}. \qquad 2.10$$

For convenience in the following discussion the constant term will be neglected. This corresponds to requiring the mean of all data to be zero.

In Part 2.1 of this section we considered the linear operator as applied to a single seismic trace and we saw that since we must deal with a finite amount of information, our problem is somewhat specialized. Further, the discrete case with which we are now concerned adds special problems of its own. As an example, let the signal that we wish to find be a series of transients representing reflections, and let the noise be the background interference. The signal plus noise, $x(t)$ or x_i, is then the trace. Since the signal must lie within the time limits of the trace data, the e_{i+k} series must necessarily contain $m+1$ less points than the x_i series if $-m \leq k \leq o$, and $m+1+k$ less points than the x_i series if $k > o$ or $k < -m$. Further, if the number of data points on the x_i trace is M and we wish to determine the coefficients a_s from the x_i data, then $2(m+1) \leq M$ for the first case above and $2(m+1)+k \leq M$ for the second case. Thus the maximum value of m is limited by the amount of data M available. In practice it usually turns out that the maximum value of m must be further limited by computational restrictions.

A similar problem may be stated as follows. Although the approximation to the signal e_{i+k} tends to improve, by certain criterion, as m increases, in some cases the arbitrariness of the phase relation between the signal and the approximation to the signal also tends to increase. Thus from the point of view of discrimination of reflections with respect to time, we do not want m to be too great. We therefore have the dilemma that an increase in the value of m will in general improve the frequency and relative phase characteristics of the approximate signal, but will increase the arbitrariness of the absolute time relationship. We will see later how this situation may be improved with the introduction of another dimension in the operator.

Let us now return to the one dimensional operator form, equation (2.9b), and again apply it to a seismogram, but from a different point of view. Assume that the seismogram contains M traces where M is a reasonably large number. If we now select one reading from each trace, say on a line perpendicular to the time axis, we may consider this set of data as a series x_i where the index i now refers to the ensemble of traces. To what the approximation e_{i+k} might now refer will be left unconsidered.

While still on the subject of one dimensional linear operators, we may introduce an additional concept for those who think in terms of correlations. If we think of the series of coefficients as a transient, then

$$e_{i+k} = \sum_{s=o}^{m} a_s x_{i-s} \qquad 2.9b$$

is nothing more than an approximation to the cross-correlation function of the series a_s and x_{-s} or the convolution of the series a_s and x_s. This last concept will enter into the discussion of filter characteristics.

2.3 The Two Dimensional Linear Operator

We have now considered applying the one dimensional linear operator along the time axis and perpendicular to the time axis. Let us combine these two approaches and form the two dimensional linear operator

$$e_{i_1+k_1,\, i_2+k_2} = \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1 s_2} \; x_{i_1-s_1,\, i_2-s_2}. \qquad 2.10$$

The notation is similar to that of the one dimensional linear operator except that the indices with the subscript 1 refer to the time axis and the indices with the subscript 2 refer to the ensemble axis. By this technique we have introduced more information into the linear operator with the probability of decreasing or removing some of the limitations or drawbacks encountered in the one dimensional case. From another point of view, the statistical sample, from which we determine the operator, has been increased and we might thus expect greater stability in the resulting mechanism. All the above improvements of course depend upon the assumption that the traces are physically related or correlated.

2.4 The Generalized Linear Operator

The Generalized Linear Operator or n-dimensional linear operator is nothing more than a generalization of the previous concepts to n dimensional space. That is, instead of attempting to determine the signal from one record, we might introduce a third dimension by adding a set of records from the same area to our linear operator. A fourth dimension might consist of a group of sets from areas of the same type, and so on. Although the practical manipulation of an n-dimensional operator would probably be too unwieldly, the concept of combining all pertinent information in a mathematical operation to determine the signal is intriguing.

2.5 The Relationship Between the Two Dimensional Linear Operator and the Prediction Operator Using More than One Trace

The prediction linear operator for two traces as defined in Reference 3 is

$$\hat{x}_{i+k'} = c + \sum_{s'=0}^{M} (a'_{s'} x_{i-s'} + b'_{s'} y_{i-s'}) \qquad 2.11$$

where x_i is the trace to be predicted and y_i is the second trace. In this case, $\hat{x}_{i+k'}$ represents an approximation to the x_i trace at a phase lead k'. However, the approximation to the signal which we wish to find is given by the error series

$$e_{i+k'} = x_{i+k'} - \hat{x}_{i+k'} \qquad 2.12$$

or

$$e_{i+k'} = x_{I+k'} - c - \sum_{s'=o}^{M} a'_{s'} x_{i-s'} - \sum_{s'=o}^{M} b'_{s'} y_{i-s'} \qquad 2.13$$

Now the two dimensional linear operator with a constant term added is

$$e_{i_1+k_1,\ i_2+k_2} = a + \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1} a_{s_2} x_{i_1-s_1,\ i_2-s_2} \qquad 2.14$$

which, if we are considering only two traces, may be written as

$$e_{i_1+k_1,\ i_2+k_2} = a + \sum_{s_1=o}^{m_1} a_{s_1,o}\, x_{i_1-s_1,i_2} + \sum_{s_1=o}^{m_1} a_{s_1,1}\, x_{i_1-s_1,i_2-1} \qquad 2.15$$

Further, if we plan to use this operator to predict only the first trace , then we can simplify the notation by letting $x_{i_1-s_1,i_2} = x_{i-s}$, representing the first trace, $x_{i_1-s_1,\ i_2-1} = y_{i-s}$, representing the second trace, and $e_{i_1+k_1,\ i_2+k_2} = e_{i+k}$, since the index i_2 is constant and $k_2 = 0$. Also, we can let $a_{s_1,o} = a_s$ and $a_{s_1,1} = b_s$. Therefore equation (2.15) simplifies to

$$e_{i+k} = a + \sum_{s=o}^{m} a_s x_{i-s} + \sum_{s=o}^{m} b_s y_{i-s}. \qquad 2.16$$

If we require

$$k = o$$
$$a = -c$$
$$a_o = 1$$
$$a_s = -a'_{s'-k'} \quad \text{for } s \geq k'$$
$$a_s = o \quad \text{for } s < k'$$
$$b_s = o \quad \text{for } s < k'$$
$$b_s = -b'_{s'-k'} \quad \text{for } s \geq k'$$
$$m = M + k'$$
$$s = s' - k'$$

then equation (2.16) may be written as

$$e_i = x_i - c - \sum_{s'=o}^{M} a'_{s'} x_{i-s'-k'} - \sum_{s'=o}^{M} b'_{s'} y_{i-s'-k'} \qquad 2.17$$

which is equivalent to equation (2.13)

$$e_{i+k'} = x_{i+k'} - c - \sum_{s'=o}^{M} a'_{s'} x_{i-s'} - \sum_{s'=o}^{M} b'_{s'} y_{i-s'} \qquad 2.13$$

We therefore see that the equation for the error series using the prediction operator is equivalent to the two dimensional linear operator plus a set of restrictions. It is, thus, a special case of the two dimensional operator. We might therefore expect that the two dimensional operator would be more flexible and would, in general, provide a better approximation to the signal.

Examining equation (2.13) it can be seen that since the output is e_{i+k} rather than $\hat{x}_{i+k}$ the concept of prediction does not really enter the problem, and that if we had applied another set of restrictions to equation (2.16) we might have obtained the error series equation corresponding to the so-called interpolation operator.

III. Filter Characteristics

3.1 Fourier Transform Concepts

Consider the Fourier integral representation of the function $x(t)$

$$x(t) = \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t} d\omega \qquad 3.1a$$

and the Fourier transform of $x(t)$

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt \qquad 3.1b$$

The variable t is regarded as the independent variable in a certain region designated as the time domain, whereas the variable ω represents an independent variable in a corresponding region known as the frequency domain. The function $x(t)$ represents a specification of some desired function in the time domain, and the corresponding function $X(\omega)$ is regarded as the specification of that same function in the frequency domain. In other words, the function $X(\omega)$ is in every respect just as complete and specific a representation of the desired function as is $x(t)$.

Though the above statement is true, certain functions are more readily interpretable in one domain than in the other. The possibility that this might be the case with the linear operator motivates the following development. Further, since $x(t)$ in equation (2.1) represents a physical phenomenon, it conforms to the requirements of representation by a Fourier integral.

3.2 Filter Characteristics of the Continuous Linear Operator

Let us consider the continuous linear operator corresponding to equation (2.9b)

$$e(t+t_o) = \int_0^{t_m} x(t-\tau)\, A(\tau)\, d\tau . \qquad 3.2$$

We then write

$$x(t) = \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t} d\omega, \qquad 3.3a$$

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt , \qquad 3.3b$$

$$e(t) = \int_{-\infty}^{\infty} E(\omega)\, e^{j\omega t} d\omega , \qquad 3.4a$$

and

$$E(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e(t)\, e^{-j\omega t} dt \qquad 3.4b$$

or for the particular operations that we wish to perform,

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t-\tau)\, e^{-j\omega(t-\tau)} dt \qquad 3.5a$$

and

$$E(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e(t+t_o)\, e^{-j\omega(t+t_o)} dt \qquad 3.5b$$

Multiplying both sides of equation (3.2) by $e^{-j\omega(t+t_o)}$, we obtain

$$e(t+t_o)\, e^{-j\omega(t+t_o)} = \int_0^{\infty} x(t-\tau)\, e^{-j\omega(t-\tau)}\, A(\tau)\, e^{-j\omega(\tau+t_o)} d\tau$$

3.6

and if we integrate both sides of (3.6) with respect to t from $t = -\infty$ to $t = +\infty$, we obtain, by equations (3.5),

$$E(\omega) = X(\omega) \int_0^{\infty} A(\tau)\, e^{-j\omega(\tau + t_o)}\, d\tau \qquad 3.7$$

Now $E(\omega)$ represents the amplitude and phase spectra of the $e(t+t_o)$ series and $X(\omega)$ represents the amplitude and phase spectra of the $x(t)$ series. Thus if $x(t)$ is regarded as the input, and $e(t+t_o)$ is regarded as the output, then

$$H(\omega) = \int_0^{\infty} A(\tau)\, e^{-j\omega(\tau + t_o)}\, d\tau \qquad 3.8$$

represents the amplitude and phase transfer function or filter characteristic. An entirely similar derivation may be made using the concept of the unit impulse response.

3.3 Filter Characteristics of the One Dimensional Linear Operator

Now let us consider a parallel development of the above expression for the filter characteristics in the case of the one dimensional linear operator

$$e_{i+k} = \sum_{s=o}^{m} a_s x_{i-s} \qquad 2.9b$$

Let the spectrum of the x_i series in the interval A to B be represented by

$$X(\omega) = \sum_{i=A}^{B} x_i e^{-j\omega hi} \qquad 3.9$$

and let the spectrum of the e_i series in the same interval be defined as

$$E(\omega) = \sum_{i=A}^{B} e_i e^{-j\omega h i} \qquad 3.10$$

or in more convenient form

$$X(\omega) = \sum_{i=A+s}^{B+s} x_{i-s} e^{-j\omega h(i-s)} \qquad 3.11a$$

and

$$E(\omega) = \sum_{i=A-k}^{B-k} e_{i+k} e^{-j\omega h(i+k)} \qquad 3.11b$$

Multiplying both sides of equation (2.9b) by $e^{-j\omega h(i+k)}$, we obtain

$$e_{i+k} e^{-j\omega h(i+k)} = \sum_{s=o}^{m} x_{i-s} e^{-j\omega h(i-s)} a_s e^{-j\omega h(s+k)}.$$

If we now sum both sides over i from $i = A$ to $i = B$, we get

$$\sum_{i=A}^{B} e_{i+k} e^{-j\omega h(i+k)} = \sum_{s=o}^{m} a_s e^{-j\omega h(s+k)} \sum_{i=A}^{B} x_{i-s} e^{-j\omega h(i-s)} \qquad 3.12$$

If the spectra of e_i and x_i are constant with respect to time or if the spectra vary slowly and $|B-A| \gg m+k$, if $k > o$ or $|B-A| \gg m$, if $-m \leq k \leq o$, then we may neglect the s and k dependence of the limits and, using equations (3.11), obtain

$$E(\omega) = X(\omega) \sum_{s=o}^{m} a_s e^{-j\omega h(s+k)} \qquad 3.13$$

or $$H(\omega) = \frac{E(\omega)}{X(\omega)} \qquad 3.14$$

where

$$H(\omega) = \sum_{s=0}^{m} a_s e^{-j\omega h(s+k)}. \qquad 3.15$$

A more direct method consists of the following. Let the input information x_i be data from a sine wave of angular frequency ω. Since the system is linear the output will be a single frequency sine wave of the same frequency but, in general, different phase and amplitude. The complex notation for a sine wave, $x(t) = e^{j\omega t}$ or in this case $x_i = e^{j\omega hi}$, is convenient. Using equations (2.9) we see that if the input is x_i, then the output e_i will be given by

$$e_i = \sum_{s=0}^{m} a_s e^{j\omega h(i-s-k)}$$

and the output over the input, by

$$H(\omega) = \frac{e_i}{x_i} = \sum_{s=0}^{m} a_s e^{-j\omega h(s+k)} \qquad 3.16$$

This equation is identical with (3.15).

The characteristics of a filter may also be expressed in terms of its unit impulse response. If $K(t)$ represents the unit impulse response of some filter system, then the output $e(t)$ may be expressed in terms of the input $x(t)$, and the impulse response $K(t)$ in the following manner. Let the input $x(t)$ be represented by a series of rectangular pulses of width h as was done in Section II, figure 1.

Then $$e(t) = \lim_{h \to \infty} \sum_{\tau = nh}^{t} x(\tau)\, hK(t-\tau).$$

3.17

This expression corresponds to the convolution integral. Let $sh = t-\tau$, then

$$e(t) = \lim_{h \to o} \sum_{s=o}^{n+t/h} x(t-sh)\, hK(sh)$$

3.18

If K(sh) is zero outside the interval $x=o$ to $s=m$, then (3.18) may be written

$$e(t) = \lim_{h \to o} \sum_{s=o}^{n} x_{1-s} hK(sh) \qquad 3.19$$

But (3.19) is equivalent to equation (2.8) of Section II.

$$e(t) \doteq \sum_{s=-M}^{n} a_s x_{1-s} \qquad 2.8$$

Therefore $hK(sh) = a_s$ in the special form of the linear operator with $M = 0$, and the a_s series thus represent h times the impulse response of this linear operator. The transfer function of this linear operator may now be obtained by the formula

$$H(\omega) = \int_{o}^{\infty} K(t)\, e^{-j\omega t} dt \qquad 3.20$$

and by its approximation

$$H(\omega) = \sum_{s=o}^{m} hK(sh)\, e^{-j\omega hs} = \sum_{s=o}^{m} a_s e^{-j\omega hs}$$

3.21

The reason that (3.19) corresponds to the restricted linear operator with $M = 0$, will be discussed later. It turns out that a more general definition of $K(t)$ admits linear operators with $M \neq 0$.

We therefore have two approaches to the consideration of the transfer properties of the linear operator. As might have been suspected, one approach is the Fourier Transform of the other. Another function which will be useful in our later work is the power transfer characteristic

$$\overline{H}(\omega)\, H(\omega) = \sum_{\tau=-m}^{m} \phi_{aa}(\tau)\, e^{-j\omega h\tau} = \phi_{aa}(0) + 2 \sum_{\tau=1}^{m} \phi_{aa}(\tau) \cos \omega h\tau \qquad 3.21$$

where $\phi_{aa}(\tau) = \sum_{s=0}^{m} a_s a_{s+\tau}$ (see Appendix II).

3.4 Filter Characteristics of the Prediction Operator

Although, as we have seen, the error series expression for the prediction operator is a special case of the two dimensional linear operator, it is convenient for some purposes to develop the expression for the transfer function in the notation of the prediction operator. Equation (2.13) gives the error series, in the case of a two trace prediction operator, as

$$e_{i+k'} = x_{i+k'} - c - \sum_{s=0}^{M} a'_s\, x_{i-s} - \sum_{s=0}^{M} b'_s\, y_{i-s} \qquad 2.13$$

where as usual we will consider $c = 0$. If, again, the complex spectrum of the e_i series is $E(\omega)$, that of the x_i series is $X(\omega)$ and that of the y_i series is $Y(\omega)$, then

$$E(\omega) = X(\omega)\ [1-A(\omega)] - Y(\omega)\ B(\omega) \qquad 3.22$$

where

$$A(\omega) = \sum_{s=0}^{M} a'_s\ e^{-j\omega h(s+k')} \qquad 3.23$$

and

$$B(\omega) = \sum_{s=0}^{M} b'_s\ e^{-j\omega h(s+k')} \qquad 3.24$$

The power transfer characteristics are, in this case,

$$(1-\bar{A})\ (1-A) = 1 - 2 \sum_{s=0}^{M} a'_s \cos \omega h(s+k') + \sum_{\tau=-M}^{M} \phi_{aa}(\tau) \cos \omega h\tau \qquad 3.25$$

where

$$\phi_{aa}(\tau) = \sum_{s=0}^{M} a'_s\ a'_{s+\tau}, \qquad 3.26$$

$$\bar{B}B = \sum_{\tau=-M}^{M} \phi_{bb}(\tau) \cos \omega h\tau \qquad 3.27$$

where

$$\phi_{bb}(\tau) = \sum_{s=0}^{M} b'_s\ b'_{s+\tau}, \qquad 3.28$$

and

$$\bar{B}(1-A) = -\sum_{s=0}^{M} b'_s\ e^{j\omega h(s+k')} + \sum_{\tau=-M}^{M} \phi_{ab}(\tau)\ e^{j\omega h\tau} \qquad 3.29$$

where

$$\phi_{ab}(\tau) = \sum_{s=0}^{M} a'_s\ b'_{s+\tau} \qquad 3.30$$

Then $\overline{EE} = \overline{XX}\ (1-\overline{A})\ (1-A) + 2Re\ [\overline{XYB}(1-A)] + \overline{YY}\ \overline{BB}$ 3.31

The extension to more than two traces is obvious.

3.5 Significance of and Restrictions on the One Dimensional Linear Operator Filter Characteristics

The physical significance of $H(\omega)$ in equations (3.14) and (3.15) is the same as that of the transfer function of an electric wave filter. That is, for a given amplitude and phase spectrum, $X(\omega)$, of the input, the output amplitude and phase spectrum, $E(\omega)$, will be given by

$$E(\omega) = H(\omega)\ X(\omega).$$

Because of the computational procedure, all functions of ω will have a period of $\omega = \frac{2\pi}{h}$. Whether or not $H(\omega)$, as given by equation (3.15)

$$H(\omega) = \sum_{s=o}^{m} a_s e^{-j\omega h(s+k)}, \qquad 3.15$$

is realizable in terms of real, linear, electric circuit components, excluding the peculiarity of periodicity, is a problem which will be discussed later.

In general $H(\omega)$ will be a complex function of frequency. Since the coefficients a_s must be real, we can write

$$H(\omega) = Re[H(\omega)] + jIm[H(\omega)] = \sum_{s=o}^{m} a_s \cos \omega h(k+s) - j\sum_{s=o}^{m} a_s \sin \omega h(k+s)$$

3.31

Thus the real part of $H(\omega)$ must be an even function of ω and the imaginary part must be an odd function of ω. If we write $H(\omega)$ in polar form

$$H(\omega) = |H(\omega)| \; e^{-j\theta(\omega)} \qquad 3.32$$

then, since we may factor out $e^{-j\omega hk}$ from the sum in (3.15), we may write

$$|H(\omega)| \; e^{-j\theta(\omega)} = e^{-j\omega hk} \sum_{s=o}^{m} a_s e^{-j\omega hs} \qquad 3.33$$

or

$$|H(\omega)| \; e^{-j[\theta(\omega)-\omega hk]} = \sum_{s=o}^{m} a_s e^{-j\omega hs} \qquad 3.34$$

where $|H(\omega)|^2 = \bar{H}(\omega) H(\omega)$ of equation (3.21),

$$\text{and} \quad \tan[\theta(\omega) - \omega hk] = \frac{\sum_{s=o}^{m} a_s \sin \omega hs}{\sum_{s=o}^{m} a_s \cos \omega hs} \; . \qquad 3.35$$

If equation (3.31) is examined, it can be seen that the imaginary part of $H(\omega)$, $Im[H(\omega)]$, may be made equal to zero for all frequencies only if all the coefficients a_s are equal to zero (the trivial case) or if m is an even integer, $k = -\frac{m}{2}$ and the coefficients are symmetric about $a_{\frac{m}{2}}$. If this is the case, then $\theta(\omega) = n\pi$ where $n=o, \pm 1, \pm 2, \ldots$. Similarly, if m is an even integer, $k = -\frac{m}{2}$, and the coefficients a_s are antisymmetric about $a_{\frac{m}{2}}$, then the real part of $H(\omega)$, $Re[H(\omega)]$, is equal to zero and $\theta(\omega) = n\pi + \frac{\pi}{2}$ where $n = 0, \pm 1, \pm 2, \ldots$.

If we consider m to be an even integer, we can analyse the coefficient series a_s into its symmetric and antisymmetric components about

$a_{\frac{m}{2}}$. Let A_s represent the symmetric component and B_s represent the antisymmetric component. Then

$$A_s = \frac{1}{2}\,[a_s + a_{m-s}] \quad s = 0, \pm 1 \;\ldots\ldots\; \frac{m}{2} \tag{3.36}$$

$$B_s = \frac{1}{2}\,[a_s - a_{m-s}] \quad s = 0, \pm 1 \;\ldots\ldots\; \frac{m}{2} \tag{3.37}$$

and $a_s = A_s + B_s$. 3.38

If we substitute (3.38) into (3.31), where $k = -\frac{m}{2}$, we obtain

$$H(\omega) = \sum_{s=o}^{m} A_s \cos \omega h(s-\tfrac{m}{2}) - j \sum_{s=o}^{m} A_s \sin \omega h(s-\tfrac{m}{2}) + + \sum_{s=o}^{m} B_s \cos \omega h(s-\tfrac{m}{2}) - j \sum_{s=o}^{m} B_s \sin \omega h(s-\tfrac{m}{2}) \tag{3.39}$$

We can therefore consider $H(\omega)$ in (3.39) as the sum of two filter characteristics

$$H_A(\omega) = \sum_{s=o}^{m} A_s \cos \omega h(s-\tfrac{m}{2}) - j \sum_{s=o}^{m} A_s \sin \omega h(s-\tfrac{m}{2}) \tag{3.40}$$

and

$$H_B(\omega) = \sum_{s=o}^{m} B_s \cos \omega h(s-\tfrac{m}{2}) - j \sum_{s=o}^{m} B_s \sin \omega h(s-\tfrac{m}{2}) \tag{3.41}$$

where

$$H_A(\omega) = |H_A(\omega)| \quad e^{-jn\pi} = \sum_{s=o}^{m} A_s \, e^{-j\omega h(s-\frac{m}{2})} \tag{3.42}$$

and

$$H_B(\omega) = |H_B(\omega)| \quad e^{-j(\frac{\pi}{2} + n\pi)} = \sum_{s=o}^{m} B_s e^{-j\omega h(s-\frac{m}{2})} \tag{3.43}$$

This means that only the symmetric part of the operator coefficients a_s contributes to the real part of $H(\omega)$, and only the antisymmetric part contributes to the imaginary part when $k = -\frac{m}{2}$ and m is an even integer. This result may also be demonstrated in the following manner.

$$H(\omega) = \text{Re}\,[H(\omega)] + j\text{Im}[H(\omega)] = \sum_{s=o}^{m} a_s \cos \omega h(s-\frac{m}{2}) - j \sum_{s=o}^{m} a_s \sin \omega h(s-\frac{m}{2}) \qquad 3.44$$

Therefore

$$\text{Re}[H(\omega)] = a_{\frac{m}{2}} + \sum_{s=o}^{\frac{m}{2}-1} (a_s + a_{m-s}) \cos \omega h(s-\frac{m}{2}), \qquad 3.45$$

which is a Fourier cosine series with $\frac{m}{2}+1$ terms. Similarly

$$\text{Im}[H(\omega)] = -\sum_{s=o}^{\frac{m}{2}-1} (a_s - a_{m-s}) \sin \omega h(s-\frac{m}{2}) \qquad 3.46$$

which is a Fourier sine series with $\frac{m}{2}$ terms. Therefore, the coefficients in (3.45) are given by

$$(a_s + a_{m-s}) = \frac{1}{\pi} \int_{-\pi}^{\pi} \text{Re}[H(\omega)] \cos \omega h(s-\frac{m}{2})\, d\omega \qquad 3.47$$

and the coefficients in (3.46) are given by

$$(a_s - a_{m-s}) = -\frac{1}{\pi} \int_{-\pi}^{\pi} \text{Im}[H(\omega)] \sin \omega h(s-\frac{m}{2})\, d\omega \qquad 3.48$$

It is seen that these coefficients are the definitions of twice the symmetric and antisymetric components as given in equations (3.36) and (3.37). Thus it is possible to derive an operator using equations (3.47) and (3.48) if we are given any filter characteristics whose real part is representable by a Fourier cosine series, and whose imaginary part is representable by a Fourier sine series of one less term and of the same period. The resulting operator will have a value of $k = -\frac{m}{2}$ and if the Fourier cosine series contained M terms, then the operator will have $m = 2(M-1)$.

Now let us attempt to examine the physical significance of this result. Again we return to equation (3.15)

$$H(\omega) = \sum_{s=0}^{m} a_s e^{-j\omega h(s+k)} \qquad 3.15$$

If we let $p = j\omega$ where $p = p_R + jp_I$ is a complex number representing the complex frequency plane (Fig. 2), we then have

$$H(p) = \sum_{s=0}^{m} a_s e^{-p_R h(s+k)} e^{-jp_I h(s+k)} \qquad 3.49$$

It is seen that this function is analytic except at $p = \pm \infty$. If $k \leq 0$ there is a pole at $+\infty$. Therefore $Re[H(p)] = H_R$ and $Im[H(p)] = H_I$ are conjugate harmonic functions in the finite plane and they satisfy Laplace's equation [4]

$$\frac{\partial^2 F}{\partial p_R^2} + \frac{\partial^2 F}{\partial p_I^2} = 0. \qquad 3.50$$

We therefore know that if, for instance, H_R is known for all values of p, then H_I may be determined by the equation

$$H_I = \int_{o,o}^{p_R p_I} \left(-\frac{\partial H_R}{\partial p_I}\, dp_R + \frac{\partial H_R}{\partial p_R}\, dp_I\right) + c \qquad 3.51$$

To determine the arbitrary constant c, let us evaluate the integral.

Now

$$H_R = \sum_{s=o}^{m} a_s e^{-p_R h(s+k)} \cos p_I (s+k) h \qquad 3.52$$

Therefore

$$\frac{\partial H_R}{\partial p_I} = - \sum_{s=o}^{m} a_s (s+k)\, h e^{-p_R h(s+k)} \sin p_I h(s+k) \qquad 3.53$$

and

$$\frac{\partial H_R}{\partial p_R} = - \sum_{s=o}^{m} a_s (s+k)\, h e^{-p_R h(s+k)} \cos p_I h(s+k) \qquad 3.54$$

If we break the line integral in equation (3.51) into two parts

$$H_I = \int_{o,o}^{p_R,0} + \int_{p_R,o}^{p_R,\, p_I} + c \qquad 3.55$$

we obtain

$$H_I = \int_{o,o}^{p_R,o} \sum_{s=o}^{m} a_s (s+k)\, h e^{-p_R h(s+k)} \sin p_I h(s+k)\, dp_R -$$

$$- \int_{p_R,o}^{p_R,p_I} \sum_{s=o}^{m} a_s (s+k)\, h e^{-p_R h(s+k)} \cos p_I h(s+k)\, dp_I + c =$$

$$= - \sum_{s=o}^{m} a_s e^{-p_R h(s+k)} \sin p_I h(s+k) + c \qquad 3.56$$

Thus c must equal zero.

As we have already stated, H_R and H_I satisfy Laplace's equation. In order to uniquely determine a function which satisfies Laplace's equation it is necessary to know either the value of the function or its slope on a closed boundary.[5/] However, when we say that we prescribe the real part of the filter characteristics, we are in reality only prescribing values of the function H_R on the p_I or real frequency axis. Nevertheless, since we know the form of H_R,

$$H_R = \sum_{s=o}^{m} a_s e^{-p_R h(s+k)} \cos p_I h(s+k), \qquad 3.57$$

under certain conditions, i.e., ranges of k, knowledge of H_R along the p_I axis is sufficient to determine H_R over half of the complex plane. Suppose, for instance, that we know $\lim H_R = o$ as $p_R \to \infty$. Thus, there are no singularities in the right half plane. Now we know, from the theory of complex variables,[4/] that every harmonic function of p_R and p_I transforms into a harmonic function of u and v under the change of variables

$$u + jv = f(p_R + jp_I) \qquad 3.58$$

where f is an analytic function. Now the particular transformation which we wish to consider is

$$u + jv = \frac{p_R+jp_I-1}{p_R+jp_I+1} .$$

Since we only consider the right half of the p plane where $p_R \geqslant o$, then f is analytic in this region. This transformation maps the right

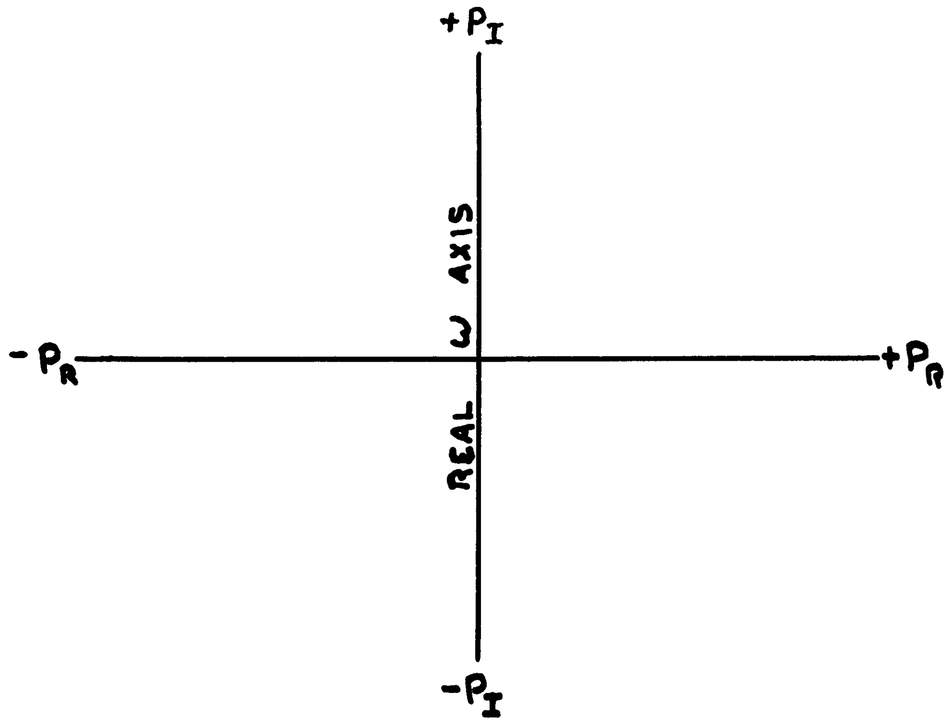


FIGURE 2

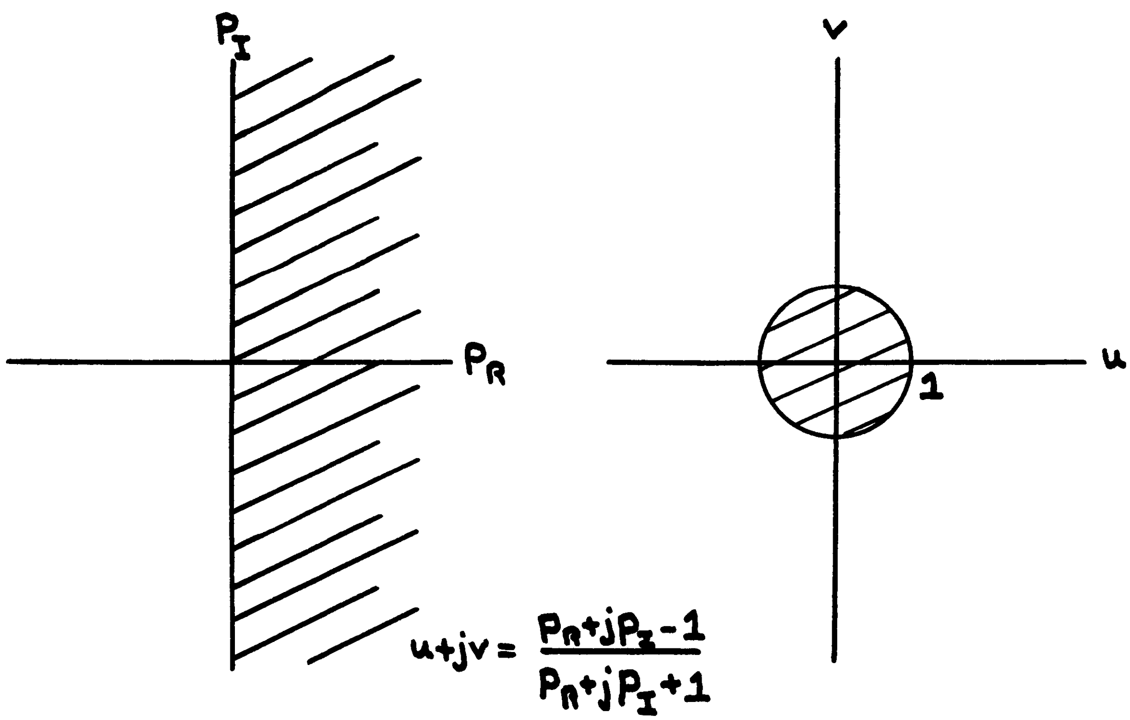


FIGURE 3

half of the p plane into the interior of the unit circle in the $u + jv$ plane and the p_I axis is mapped into the closed boundary of the unit circle (Fig. 3). Therefore, if H_R is a harmonic function in the entire right half of the p plane, it will transform into a harmonic function in the $u + jv$ plane and, thus, must satisfy Laplace's equation within the unit circle and on the boundary. But we have prescribed the value of H_R on the p_I axis. Therefore, we have prescribed the value of the transform of H_R on a closed boundary in the $u + jv$ plane and hence, we have prescribed the value of the transform of H_R throughout the unit circle, or throughout the right half of the p plane.

Now let us see under what conditions H_R is a harmonic function in the entire right half of the p plane. We see from equation (3.57) that this will be so if $(s+k) \geq \epsilon$ for all values of s where ϵ is a small positive quantity. Thus we must have $k \geq s + \epsilon$ for all s. Now s has a minimum value of 0. Therefore $k \geq \epsilon$. A similar treatment may be carried through in the left half of the p plane, and, in this case, we find that H_R is only completely specified by its Dirichlet conditions if $s+k \leq -\epsilon$ where ϵ is a small positive number. Since in this case the maximum value of s is $s=m$, $k \leq -(m+\epsilon)$.

We therefore see that H_R is completely specified by its boundary conditions on the p_I axis only if $k \geq \epsilon$ or $k \leq -(m+\epsilon)$. This treatment does not, however, determine the degree of specification when $-(m+\epsilon) < k < \epsilon$ or when $\epsilon = 0$. In conclusion, we can say that in view of equation (3.51), the imaginary component of the filter characteristics will be completely determined by the real frequency characteristics of the real component of $H(\omega)$ if $k \geq \epsilon$ or $k \leq -(m+\epsilon)$, and that if $-(m+\epsilon) < k < \epsilon$ then $Im[H(\omega)]$ is at least partly independent of $Re[H(\omega)]$. Only when $k = -\frac{m}{2}$ is $Im[H(\omega)]$ completely independent of $Re[H(\omega)]$. Nevertheless, in all the above situations, $Re[H(\omega)]$ must be an even function of ω and $Im[H(\omega)]$ must be an odd function.

3.6 Physical Realizability of the One Dimensional Linear Operator Filter

The requirements that a transfer function $H(\omega)$ be realizable in terms of a physical, stable network, whether active or passive, may be stated as follows.[6]

1. The zeros and poles of $H(p)$ in the p plane must be real, or they must occur in conjugate complex pairs.

2. The real and imaginary components of $H(\omega)$ must be, respectively, even and odd functions of p on the imaginary p axis (the real frequency axis).

3. Poles on the imaginary p axis must be simple.

4. None of the poles of $H(\omega)$ can be found in the right half of the p plane.

Now let us examine equation (3.49),

$$H(p) = \sum_{s=o}^{m} a_s e^{-p_R h(s+k)} e^{-jp_I h(s+k)}, \qquad 3.49$$

with respect to these requirements. Requirement (1) is fulfilled with respect to poles, for only infinite values of p_R can make $H(p) = \infty$. Zeros can only occur when $Re[H(p)] = 0$. Therefore if p_{I1} is a zero, then $-p_{I1}$ must also be a zero. Thus the zeros must be real or occur in conjugate complex pairs. We have already seen that $Re[H(\omega)]$ is even and $Im[H(\omega)]$ is odd, so that requirement (2) is also satisfied. Requirement (3) is certainly satisfied since no poles exist on the imaginary p axis. Requirement (4), however, will be violated unless $k \geq 0$. Nevertheless, this limitation on the realizability of $H(\omega)$ is only apparent since k must be finite. If, for example, $k < 0$, we can introduce a delay k such that $Y(\omega)$ is realizable,[7] where

$$Y(\omega) = H(\omega)\, e^{j\omega hk} = \sum_{s=o}^{m} a_s\, e^{-j\omega hs} \qquad 3.59$$

This situation may perhaps be visualized more clearly in terms of the discrete impulse response K_i, where K_i is given by the inverse Fourier transform of $H(\omega)$

$$= \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} H(\omega)\, e^{j\omega hi} d\omega \qquad 3.60$$

or

$$hK_i = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \sum_{s=o}^{m} a_s e^{-j\omega h(s+k)} e^{j\omega hi}\, d\omega = a_{i-k}. \qquad 3.61$$

K_i is interpreted as the output obtained from a filter in response to a unit impulse impressed upon its input at time $i = o$. To make physical sense, the output can only be non zero for $i \geq o$ and for stability, it must drop to zero again as i becomes large. Therefore, K_i must be zero for $i < 0$ and K_i must again become zero for $i > M$ where M is some large finite number. The second condition is automatically fulfilled, for k and m are both finite. The first condition, however, will not be fulfilled if $k<0$, for then $K_{-|k|} = a_0$, $K_{-|k|+1} = a_1$, etc. That is, the response to the unit impulse will occur before the unit impulse is impressed upon the filter. However, since k is finite, we can introduce a delay of $|k|$ into the filter such that the response is zero before the impulse is impressed.[7/]

One further problem exists in the comparison of the one dimensional linear operator filter with a real network, for as we have seen, the filter characteristics of the linear operator are periodic. This phenomenon is a result of the fact that the linear operator is discrete. However, we have seen that the computed spectra are also periodic of the same period, and therefore the whole spectrum of our discrete data lies between

$$-\frac{\pi}{h} \leq \omega \leq \frac{\pi}{h} \quad \text{or} \quad 0 \leq \omega \leq \frac{2\pi}{h}$$

where h is the spacing of the discrete data. The spectrum of the discrete data does not, in general, represent the spectrum of the continuous trace because the discrete computational procedure can not distinguish frequencies differing by $\frac{n}{h}$ or values of ω differing by $\frac{2\pi n}{h}$, where n is an integer. It is therefore necessary to choose h sufficiently small such that the power at frequencies $|f|>\frac{1}{h}$ is negligible. If this is the case, then the spectrum of interest in the continuous case lies between $0\leq\omega\leq\frac{2\pi}{h}$

and the filter characteristics of the continuous case correspond to the filter characteristics of the discrete case.

To summarize the above results, we can say that a one dimensional linear operator filter may be represented by an equivalent active, stable, physical network. Further, it may be approximated to any degree of accuracy by a passive lumped element network, together with a single amplifier.[7/] On the other hand, can a finite linear operator possess filter characteristics equivalent to those of any active, stable, physical network? The answer to this question is probably no. We have seen that the transfer function of a linear operator which represents a physical network, can have only one pole at $p_R = -\infty$. The transfer function of a physical network, however, may have poles throughout the left half of the p plane and on the p_I axis. But poles on the p_I axis correspond to infinities on the real frequency axis, and since $H(\omega)$ is essentially a Fourier series representation of some transfer function $H'(\omega)$, we have a Fourier series approximation problem. Thus we can only approximate the infinities as maxima, and in general we must require an infinite number of operator coefficients for a good approximation.

We can, therefore, say that the one dimensional linear operator filter is a special case of all possible active, physical, stable filters. However, if we use a sufficient number of terms in the operator we can probably approximate any physical filter characteristics to a satisfactory degree of accuracy.

3.7 Filter Characteristics of the Two Dimensional Linear Operator

The filter characteristics of a two dimensional operator may be derived in a manner analagous to that used in the derivation of the one dimensional linear operator filter.[8/] Suppose we consider the input

$$x_{i_1,i_2} = e^{j(\omega_1 h_1 i_1 + \omega_2 h_2 i_2)} ,$$

then the output will be

$$e_{i_1,i_2} = \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1,s_2} \, e^{j\omega_1 h_1(i_1-k_1-s_1)} \, e^{j\omega_2 h_2(i_2-k_2-s_2)} \tag{3.62}$$

and the output over the input, $H(\omega_1,\omega_2)$, will be

$$H(\omega_1,\omega_2) = \frac{e_{i_1,i_2}}{x_{i_1,i_2}} = \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1,s_2} \, e^{-j\omega_1 h_1(s_1+k_1)} \, e^{-j\omega_2 h_2(s_2+k_2)} \tag{3.63}$$

Also, the two dimensional input spectrum may be defined as

$$X(\omega_1,\omega_2) = \sum_{i_1=A_1}^{B_1} \sum_{i_2=A_2}^{B_2} x_{i_1,i_2} \, e^{-j(\omega_1 h_1 i_1 + \omega_2 h_2 i_2)} \tag{3.64}$$

and the output spectrum as

$$E(\omega_1,\omega_2) = \sum_{i_1=A_1}^{B_1} \sum_{i_2=A_2}^{B_2} e_{i_1,i_2} \, e^{-j(\omega_1 h_1 i_1 + \omega_2 h_2 i_2)} . \tag{3.65}$$

Again, for the relation

$$E(\omega_1,\omega_2) = X(\omega_1,\omega_2)\ H(\omega_1,\omega_2) \qquad 3.66$$

to hold, the spectra must not vary too rapidly with respect to i_1 or i_2 in the intervals of length $m_1 + k_1$ and $m_2 + k_2$.

We can visualize the two dimensional spectrum as follows. If the subscript 1 again refers to the time axis of a seismogram and the subscript 2 refers to the ensemble axis, we can construct a three dimensional diagram where one axis is the i_1 axis, another is the i_2 axis and the third represents the amplitude of the trace. Thus the two dimensional amplitude is essentially a surface. If we now take the Fourier transform of the surface along the i_1 axis, we obtain a surface representing the spectrum of a single trace along one axial direction, say the ω_1 axis, while moving in the i_2 axial direction we pass from the spectrum of one trace to the spectrum of another. If we now take the Fourier transform of this surface along the i_2 axis, we obtain a new surface, with coordinates ω_1 and ω_2, which represents the two dimensional spectrum. This second transformation is essentially a spectrum analysis, in the i_2 direction, of the trace spectra at various values of ω_1.

To clarify the situation, let us consider an example. Suppose we have a seismogram with M traces and we wish to operate on it with the two dimensional linear operator

$$e_{i_1,i_2} = (a_{oo}x_{i_1,i_2} + a_{1o}x_{i_1-1,i_2}) + (a_{o1}x_{i_1,i_2-1} + a_{11}x_{i_1-1,i_2-1}) \qquad 3.67$$

That is, we wish to use an operator with four coefficients, two on each of two neighboring traces. We can compute the spectrum $X_{i_2}(\omega_1)$ of the i_2 trace and the spectrum $X_{i_2-1}(\omega_1)$ of the i_2-1 trace. We can also compute the filter characteristics, $H_{i_2-1}(\omega_1)$ for the i_2-1 trace. Further, we can write the spectrum of the output as $E_{i_2}(\omega_1)$ for this case of the operator applied to the i_2 and i_2-1 traces. Thus

$$E_{i_2}(\omega_1) = X_{i_2}(\omega_1)\ H_{i_2}(\omega_1) + X_{i_2-1}(\omega_1)\ H_{i_2-1}(\omega_1). \qquad 3.68$$

We see that $E_{i_2}(\omega_1)$ is the result of the filtering of $X_{i_2}(\omega_1)$ by $H_{i_2}(\omega_1)$ plus the vector addition of $X_{i_2-1}(\omega_1)$ filtered by $H_{i_2-1}(\omega_1)$. Equation (3.68) corresponds to taking the Fourier Transform of (3.67) in the i_1 direction only. But we wish to know the output spectrum of each set of traces to which the operator is applied. We could therefore compute E_{i_2} for all values of i_2 and compare this set of spectra. A corresponding form of this information can be obtained by taking the Fourier transform of (3.68) with respect to i_2. We then obtain

$$E(\omega_1,\omega_2) = X(\omega_1,\omega_2)\ H(\omega_1,\omega_2). \qquad 3.66$$

Since we know that a one dimensional linear operator corresponds to filtering, we can see from (3.67) and (3.68) that a two dimensional linear operator corresponds to filtering and mixing, or from (3.66), filtering in two dimensions. Since, however, the data is not continuous across the traces, the two dimensional filter is not realizable in terms of networks and we must be satisfied with a filtering and mixing device. To put it another way, electronic filtering in the ensemble direction must be carried out by using discrete data, just as is done in the linear operator.

IV Methods of Determining the Optimum Linear Operator or Filter

4.1 The Influence of the Statistical Properties of the Data on the Determination of the Linear Operator

In the two preceding sections, the characteristics of the linear operator have been developed, independent of assumptions as to the statistical nature of the input data. However, in order to evaluate the usefulness of the linear operator, (and therefore the linear filter), it is necessary to consider the characteristics of the signal and noise to which the linear operator will be applied. Also, these characteristics of the signal and noise will in large part determine the criterion by which the linear operator should be optimized.

We have seen that the operator coefficients are independent of time, and that therefore the filter characteristics must also be independent of time. Thus, if we design an optimum phase and amplitude filter for the signal and noise in one interval of time, and we wish this filter to be equally as satisfactory in another interval of time, then we must require the phase and amplitude spectra of the signal and noise to be essentially constant with respect to time. If we had only optimized the power characteristics of the filter, then we could reduce the restriction on the spectra to the requirement that the power spectra of the noise and signal be approximately constant with time. If, for instance, the noise power spectrum is approximately constant with time, but the signal power spectrum is not, then the criterion of optimization should be to design a filter with power characteristics which minimize the power in the noise spectrum.

Another criterion for designing the linear operator is the transient response or the impulse response. If, for example, the signal consisted of a series of similar transients, then we might wish to contract these transients in such a way that they would appear in the output as distinguishable pulses.

It is apparent from the above discussion that the criterion of optimization depends upon the nature of the input information and also upon the desired output. There are thus two alternative approaches to determining the best criterion. The first approach, and perhaps the more direct, is to thoroughly analyze the data, drawing on all available knowledge of its probable characteristics. The second approach is to apply the criteria of optimization in succession, beginning with the criterion which imposes the weakest requirements on the data. The optimum criterion in this case should be that which gives the best signal from noise separation.

As will be seen, the exact fitting of an $m+1$ term linear operator to $m+2$ points corresponds to optimization of the phase and amplitude filter characteristics in that interval. The least squares method of fitting an $m+1$ term linear operator to M points, where $M \gg m+1$, corresponds to the optimization of the power filter characteristics in the fitting interval.

In a previous section it was mentioned that seismic data might be considered as existing from $-\infty$ to ∞. However, it is obvious that the statistical properties of this data differ radically before the shot and just after it. Similarly, the statistics of the trace just after the shot must differ from the statistics as $t \to \infty$. Thus, the best we can hope for is that in the interval of the recorded trace, the characteristics of the signal or noise are approximately constant with time.

4.2 The Method of Least Squares Fitting of the Linear Operator to the Signal

This method requires that we know the signal in some interval of the data. We then write

$$g_i = s_i - e_i \qquad 4.1$$

where s_i is the value of the signal at i, e_i is the output of the linear operator at i, and g_i is the difference between s_i and e_i at i. We then sum the $(g_i)^2$ over some interval A to B and determine the coefficients of the linear operator such that

$$I = \sum_{i=A}^{B} g_i^2 \qquad 4.2$$

is a minimum.

Let us consider the case of the two dimensional linear operator.

$$e_{i_1+k_1, i_2+k_2} = \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1 s_2} x_{i_1-s_1, i_2-s_2} \qquad 4.3$$

If we know the signal $s_{i_1+k_1, i_2+k_2}$ over the interval A_1 to B_1 on each trace, we can sum the squared difference over A_1 to B_1 on traces A_2 to B_2.

First, however, let us assume that we wish to determine the operator only for trace $i_2+k_2 = k_2$. We can then simplify the notation of equation (4.3) to

$$e_{i_1+k_1} = \sum_{s_1=o}^{m_1} \sum_{s_2=o}^{m_2} a_{s_1 s_2} \; x_{i_1-s_1,\, o-s_2} \; . \qquad 4.4$$

Equation (4.4) can be written

$$e_{i_1+k_1} = a_{oo}x_{i_1,o} + a_{1o}x_{i_1-1,o}+\dots+ a_{o1}x_{i_1,-1}+a_{11}x_{i_1-1,-1}+\dots \qquad 4.5$$

The right hand side of this equation has the form of the product of a row matrix and a column matrix. Let us therefore write

$$e_{i_1+k_1} = \underline{x}_{i_1}\, \underline{a} \qquad 4.6$$

where

$$\underline{x}_{i_1} = [x_{i_1,o}\, x_{i_1-1,o} \dots\; x_{i_1,-1}\; x_{i_1-1,-1} \dots] \qquad 4.7$$

and $\underline{a} = \left\{ a_{oo} \;\; a_{1o} \dots\; a_{o1} \qquad a_{11} \dots \right\}$ 4.8

Therefore, $g_{i_1+k_1} = s_{i_1+k_1} - e_{i_1+k_1} = s_{i_1+k_1} - \underline{x}_{i_1}\, \underline{a}$ 4.9

For each value of i_2, we obtain a relationship of the form (4.9). Therefore, if we define a column matrix

$$\underline{g} = \left\{ g_A \; g_{A+1} \dots\; g_{i_1+k_1} \dots\; g_B \right\} \qquad 4.10$$

we can write

$$\underline{g} = \underline{s} - \underline{x}\, \underline{a} \qquad 4.11$$

where $\underline{s} = \left\{ s_A \; s_{A+1} \dots\; s_{i_1+k_1} \dots\; s_B \right\}$ 4.12

and $\underline{x} = \left\{ \underline{x}_{A-k_1} \;\; \underline{x}_{A-k_1+1} \dots\; \underline{x}_{i_1} \dots\; \underline{x}_{B-k_1} \right\}$ 4.13

or $\underline{x}$ is a rectangular matrix with rows $\underline{x}_{i_1}$ as defined in equation (4.7). From (4.10) we see that

$$I = \sum_{i_1+k_1=A}^{B} g^2_{i_1+k_1} = \underline{g}^T \underline{g}. \qquad 4.14$$

But $\underline{g}^T = [\underline{s} - \underline{x}\underline{a}]^T = \underline{s}^T - (\underline{x}\,\underline{a})^T$.

Therefore $\qquad I = \underline{g}^T\underline{g} = [\underline{s}^T - (\underline{x}\,\underline{a})^T][\underline{s} - \underline{x}\,\underline{a}] =$

$$= \underline{s}^T\underline{s} - \underline{s}^T\,\underline{x}\,\underline{a} - \underline{a}^T\underline{x}^T\,\underline{s} + \underline{a}^T\,\underline{x}^T\,\underline{x}\,\underline{a} \qquad 4.15$$

But $\underline{s}^T\,\underline{x}\,\underline{a} = (\underline{a}^T\underline{x}^T\underline{s})^T$ is a scalar and the transpose of a scalar is the scalar itself. Therefore, $\underline{s}^T\,\underline{x}\,\underline{a} = \underline{a}^T\,\underline{x}^T\underline{s}$ and equation (4.15) becomes

$$I = \underline{s}^T\underline{s} - 2\,\underline{a}^T\underline{x}^T\underline{s} + \underline{a}^T\underline{x}^T\underline{x}\,\underline{a} \qquad 4.16$$

Minimizing (4.16) we get

$$\delta I = -2\,\delta\,[\underline{a}^T\underline{x}^T\underline{s}] + \delta\,[\underline{a}^T\underline{x}^T\underline{x}\,\underline{a}] = 0 \qquad 4.17$$

But $\delta\,[\underline{a}^T\underline{x}^T\underline{x}\,\underline{a}] = \delta(\underline{a}^T)\,\underline{x}^T\underline{x}\,\underline{a} + \underline{a}^T\underline{x}^T\underline{x}\,\delta\,\underline{a} \qquad 4.18$

Again, since $\delta(\underline{a}^T)\,\underline{x}^T\underline{x}\,\underline{a} = [\underline{a}^T\,\underline{x}^T\,\underline{x}\;\;\delta\,\underline{a}]^T$ is a scalar equation, the transpose of the right hand side is the right hand side itself, and from the resulting equality we see that (4.17) may be written

$$\delta I = -2\,\delta\,[\underline{a}^T\underline{x}^T\underline{s}] + 2\;\;\delta(\underline{a}^T)\,\underline{x}^T\underline{x}\,\underline{a} = 0 \qquad 4.19$$

The resulting minimizing condition can then be written

$$\delta(\underline{a}^T)\ \underline{x}^T\underline{x}\ \underline{a} = \delta(\underline{a}^T)\ \underline{x}^T\underline{s} \qquad 4.20$$

If we premultiply both sides of equation (4.20) by $\delta(\underline{a})$ and denote $\delta(\underline{a})\ \delta(\underline{a}^T)$ by $\underline{G}$, (4.20) becomes

$$\underline{G}\ \underline{x}^T\underline{x}\ \underline{a} = \underline{G}\ \underline{x}^T\underline{s} \qquad 4.21$$

It is seen that the determinant of the square matrix $\underline{G}$ is the Gramian[9/] of $\delta(a_j)$. Therefore, if the $\delta(a_j)$'s are independent, $|\underline{G}| \neq o$ and thus $\underline{G}$ possesses an inverse. If we then premultiply both sides of (4.21) by $\underline{G}^{-1}$, we obtain the simplified minimizing condition

$$\underline{x}^T\underline{x}\ \underline{a} = \underline{x}^T\underline{s} \qquad 4.22$$

The matrix $\underline{x}^T\underline{x}$ is square, symetric, and in general $|\underline{x}^T\underline{x}| \neq o$. We may therefore compute the inverse of $\underline{x}^T\underline{x}$, and premultiply both sides of (4.22) by $[\underline{x}^T\underline{x}]^{-1}$. Therefore

$$\underline{a} = [\underline{x}^T\underline{x}]^{-1}\underline{x}^T\underline{s} \qquad 4.23$$

The requirement that the $\delta(a_j)$'s be independent is the requirement that $a_{s_1s_2}$'s be independent variables. If, however, some of the a_{s_1s}'s are restricted to be constants, then the corresponding $\delta(a_j)$'s will be zero and the matrix equation (4.20) may be partitioned[10/] such that a new $\underline{G}'$ is found which has a determinant $|\underline{G}'| \neq 0$. Then a similar simplification may be carried out with the result

$$\underline{a}' = [\underline{x}'^T\ \underline{x}']^{-1}\underline{x}'^T\underline{s}'. \qquad 4.24$$

in the special case when some of the $a_{s_1 s_2}$'s are dependent variables, the situation is more complex, but if the dependence is linear, then a set of linear simultaneous equations again results.

It will be remembered that the preceding development was carried through for the determination of the least squares linear operator from the signal contained in one trace only. Let us now determine the best least squares linear operator using the signal information contained in traces A_2 through B_2. We see that equation (4.14) will then be modified to

$$I = \sum_{i_2+k_2=A_2}^{B_2} \sum_{i_1+k_1=A_1}^{B_1} g^2_{i_1+k_1, i_2+k_2} = \sum_{i_2+k_2=A_2}^{B_2} \underline{g}^T_{i_2+k_2} \underline{g}_{i_2+k_2} \tag{4.25}$$

or

$$I = \sum_{i_2+k_2=A_2}^{B_2} [\underline{s}^T_{i_2+k_2} \underline{s}_{i_2+k_2} - 2\underline{a}^T \underline{x}^T_{i_2} \underline{s}_{i_2+k_2} + \underline{a}^T \underline{x}^T_{i_2} \underline{x}_{i_2} \underline{a}] \tag{4.26}$$

If we let

$$\underline{R}_{ss} = \sum_{i_2+k_2=A_2}^{B_2} \underline{s}^T_{i_2+k_2} \underline{s}_{i_2+k_2} , \tag{4.27}$$

$$\underline{R}_{xs} = \sum_{i_2+k_2=A_2}^{B_2} \underline{x}^T_{i_2} \underline{s}_{i_2+k_2} , \tag{4.28}$$

and

$$\underline{R}_{xx} = \sum_{i_2+k_2=A_2}^{B_2} \underline{x}^T_{i_2} \underline{x}_{i_2}$$

Then we may write (4.26) as

$$I = \underline{R}_{ss} - 2\underline{a}^T \underline{R}_{xs} + \underline{a}^T \underline{R}_{xx} \underline{a} \qquad 4.29$$

This equation has the same form as (4.16) and thus the resulting minimizing condition, or normal equation matrix is

$$\underline{a} = [\underline{R}_{xx}]^{-1} \underline{R}_{xs} \qquad 4.30$$

Thus, the coefficients as determined by equation (4.30) constitute the operator coefficients of the best least squares two dimensional linear operator. One could also determine a two dimensional linear operator by finding the least squares operator for each signal, as given by equation (4.23), and averaging the resulting coefficients. However, it can be shown that the resulting two dimensional linear operator is not the "best" in the least squares sense.

The application of this technique to seismograms is somewhat specialized since the signal s_1 is usually not considered as a continuous function of time. That is, we think of each reflection as a separate unit. However, if we know the signal to be zero in some interval, then we can use this interval of $s_1 = o$ to determine the operator.[3/] The resulting operator will then be the best least squares approximation to $s_1 = o$ in the fitting interval and, if the statistical properties of the traces are near constant with respect to time, the operator should continue to give a good approximation to zero as it moves along the time axis. If, however, the presence of a reflection modifies the properties of the trace in some region of time, we would expect the approximation e_i to differ from zero in the reflection interval. This does not imply, however, that the output e_i in this interval will resemble the signal s_i in form, for we included no information about the form of reflections when we designed the operator in the interval $s_i = o$. The best we can therefore expect is that the amplitude of e_i will be greater in a reflection interval than in a non-reflection interval, and since this is our only criterion by the least squares method, the output may best be displayed as e_i^2.

As an alternative procedure in applying the least squares method of determining the linear operator, we might fit e_i to a reflection interval. In this case, we would expect e_i to be small when in any reflection interval, if we assume the characteristics of all the reflections to be similar, and we would expect e_i to have a greater amplitude when in a non-reflection interval. The above problems will be considered again, later in this section.

It should be noted that if we assume $s_i = o$ in the fitting interval, then the right side of equation (4.22) will be zero, and the solution of this equation will be the trivial solution $a_{s_1 s_2} = o$. We must therefore require that one of the coefficients, say a_{oo}, be equal to a constant. As we have seen, we must then go back and partition the matrix in (4.20) and obtain a modified form of (4.22). This procedure corresponds to throwing away one of the homogeneous equations of the set

$$\underline{x}^T \underline{x}\, \underline{a} = \underline{o} \ , \qquad 4.31$$

and thus reducing the rank of the matrix $\underline{x}^T \underline{x}$ such that it is one less than the number of unknown coefficients $a_{s_1 s_1}$. The solutions we then obtain will be multiples of an arbitrary constant which we can specify for convenience. This case corresponds to the prediction operator where the first coefficient, multiplying the predicted trace, is required to be equal to 1.

4.3 The Least Squares Linear Operator Filter Characteristics

Equation (4.14) is the expression for the sum of squared differences between the signal and the operator output along one trace.

$$I = \sum_{i_1+k_1=A_1}^{B_1} g^2_{i_1+k_1} \qquad 4.14$$

But $\sum_{i_1+k_1=A_1}^{B_1} g^2_{i_1+k_1}$ is the computed autocorrelation of $g_{i_1+k_1}$ at zero lag. That is

$$\phi_{gg}(o) = \sum_{i_1+k_1=A_1}^{B_1} g^2_{i_1+k_1} \quad . \qquad 4.32$$

Parseval's Theorem for a periodic spectrum tells us that

$$\phi_{gg}(o) = 2\pi h \int_{-\pi/h}^{\pi/h} \overline{G}(\omega)\, G(\omega)\, d\omega \qquad 4.33$$

where $G(\omega)$ is the spectrum of $g_{i_1+k_1}$. Therefore

$$I = 2\pi h \int_{-\pi/h}^{\pi/h} \overline{G}(\omega)\, G(\omega)\, d\omega \qquad 4.34$$

Thus, since the minimizing condition of least squares is $\delta I = 0$, we see that in the frequency domain we require the filter characteristics which minimize the total power in the spectrum of $g_{i_1+k_1}$, in the data interval A_1 to B_1.

$\overline{G}(\omega)\, G(\omega)$ is an even function. We may therefore write

$$I = 4\pi h \int_0^{\pi/h} \overline{G}(\omega)\, G(\omega)\, d\omega \qquad 4.35$$

But $g_{i_1+k_1} = s_{i_1+k_1} - e_{i_1+k_1}$. Therefore, if the spectrum of $s_{i_1+k_1}$ is $S(\omega)$, then $G(\omega) = S(\omega) - E(\omega)$, and since $E(\omega) = X(\omega)H(\omega)$, we have

$$G(\omega) = S(\omega) - X(\omega)\, H(\omega)$$

and thus

$$I = 4\pi h \int_0^{\pi/h} [\overline{S}S - \overline{S}XH - S\overline{X}\overline{H} + \overline{X}X\overline{H}H]\, d\omega \qquad 4.36$$

We have seen previously that we can write $H = H_R + j\, H_I$ where H_R and H_I are independent. Therefore we may write

$$I = 4\pi h \int_0^{\pi/h} [\overline{S}S - \overline{S}XH_R - S\overline{X}H_R + \overline{X}XH_R^2 - j\overline{S}XH_I + jS\overline{X}H_I + \overline{X}XH_I^2\,]\, d\omega$$

and

$$\delta I = 4\pi h \int_0^{\pi/h} \left\{[-\overline{S}X - S\overline{X} + 2\overline{X}XH_R]\ \delta H_R + [-j\overline{S}X + jS\overline{X} + 2\overline{X}XH_I]\ \delta H_I\right\} d\omega = 0$$

4.37

Since H_R and H_I are independent, δH_R and δH_I are also independent and we must have

$$\int_0^{\pi/h} \left[-\bar{S}X - S\bar{X} + 2\bar{X}XH_R\right] \delta H_R \, d\omega = 0 \qquad 4.38$$

and

$$\int_0^{\pi/h} \left[-j\bar{S}X + jS\bar{X} + 2\bar{X}XH_I\right] \delta H_I \, d\omega = 0 \qquad 4.39$$

Since H_R is a finite Fourier cosine series and H_I is a finite Fourier sine series, in general δH_R and δH_I are not arbitrary functions in the interval $0 \leq \omega \leq \pi/h$ and we must be content with equations (4.38) and (4.39). The integration of equations (4.38) and (4.39) yields a pair of matrix equations equivalent to the matrix relation (4.19), and if the coefficients in the H_R and H_I series are independent, then the matrix equation (4.23) results.

If, however, as the number of terms in the Fourier series H_R and H_I is increased, it happens that the functions δH_R and δH_I become sufficiently arbitrary such that we can choose two functions δH_R and δH_I which make the integrands in (4.38) and (4.39) positive whenever they are not zero, then we must require that

$$-\bar{S}X - S\bar{X} + 2\bar{X}XH_R = 0 \qquad 4.40$$

and

$$-j\bar{S}X + jS\bar{X} + 2\bar{X}XH_I = 0 \qquad 4.41$$

Therefore, the minimizing condition may be written in terms of filter characteristics and spectra as

$$H_R = \frac{\bar{S}X + S\bar{X}}{2\bar{X}X} \qquad 4.42$$

$$H_I = j\,\frac{\bar{S}X - S\bar{X}}{2\bar{X}X} \qquad 4.43$$

In Section I, the input information x_i was defined

$$x_i = s_i + n_i \qquad 4.44$$

where s_i represents the signal and n_i represents the background interference. It was stated that, in general, n_i consists of a predictable component p_i and a random component r_i. Therefore, if the corresponding spectra of x_i, s_i, n_i, p_i, and r_i are $X(\omega)$, $S(\omega)$, $N(\omega)$, $P(\omega)$, and $R(\omega)$ respectively, where $R(\omega)$ is undefined, then

$$X(\omega) = S(\omega) + P(\omega) + R(\omega) \qquad 4.45$$

$$\bar{X}X = \bar{S}S + \bar{P}P + \bar{R}R + \bar{S}P + \bar{P}S + \bar{S}R + \bar{R}S + \bar{P}R + \bar{R}P, \qquad 4.46$$

$$\bar{S}X = \bar{S}S + \bar{S}P + \bar{S}R, \qquad 4.47$$

and

$$S\bar{X} = S\bar{S} + S\bar{P} + S\bar{R} \qquad 4.48$$

But since r_i is a random component, the cross power spectra of r_i and any other component will be zero if the mean of r_i is zero.[11] Only the power spectrum of r_i will be non zero. Therefore (4.46) reduces to

$$\bar{X}X = \bar{S}S + \bar{P}P + \bar{R}R + \bar{S}P + \bar{P}S \qquad 4.49$$

and (4.47) and (4.48) become

$$\bar{S}X = \bar{S}S + \bar{S}P \qquad 4.50$$

and

$$S\bar{X} = S\bar{S} + S\bar{P} \qquad 4.51$$

Thus equations (4.42) and (4.43) may be written as

$$H_R = \frac{\overline{S}S + \mathrm{Re}[\overline{S}P]}{\overline{S}S + \overline{P}P + \overline{R}R + 2\mathrm{Re}[\overline{S}P]} \qquad 4.52$$

and

$$H_I = \frac{-\mathrm{Im}[\overline{S}P]}{\overline{S}S + \overline{P}P + \overline{R}R + 2\mathrm{Re}[\overline{S}P]} \qquad 4.53$$

and the best least squares linear operator transfer function is therefore

$$H(\omega) = \frac{\overline{S}S + S\overline{P}}{\overline{S}S + \overline{P}P + \overline{R}R + 2\mathrm{Re}[\overline{S}P]} \qquad 4.54$$

Let us consider some special cases of equation (4.54). If the random component of the noise r_i is equal to zero, then $\overline{R}R = 0$. Equation (4.54) then reduces to

$$H(\omega) = \frac{\overline{S}S + S\overline{P}}{\overline{S}S + 2\mathrm{Re}[\overline{S}P] + \overline{P}P} \qquad 4.55$$

But $2\mathrm{Re}[\overline{S}P] = \overline{S}P + S\overline{P}$. Therefore

$$H(\omega) = \frac{S(\overline{S} + \overline{P})}{(S+P)\ (\overline{S}+\overline{P})} = \frac{S}{S+P} \qquad 4.56$$

Equation (4.56) tells us that in the absence of a random component, the best least squares linear operator is the one which transforms the signal plus noise exactly into the signal. This corresponds to perfect fitting in the fitting interval. For a finite interval, it is possible to describe the noise as a function of time. Therefore, for this finite interval there is no random component.

As another special case, let us suppose that the predictable component of the noise is zero, but the random component is not. Then by equations (4.52) and (4.53)

$$H_R = \frac{\overline{S}S}{\overline{S}S + \overline{R}R} \qquad 4.56$$

and $$H_I = 0 \qquad 4.57$$

Therefore $$H(\omega) = \frac{\overline{S}S}{\overline{S}S + \overline{R}R} \qquad 4.58$$

Since $\overline{S}S$ and $\overline{R}R$ must be positive, then not only is $H(\omega)$ in (4.58) real, but it must also be positive. Therefore, the phase angle of $H(\omega)$ is zero or some multiple of 2π. In this case it is seen that the best least squares transfer function is determined solely from the power spectra of the signal and noise.

Another special case, which has been treated in Part 4.2 of this section, is the case in which we wish to determine the operator in an interval of zero signal. It is evident from equation (4.54) that the trivial solution $H(\omega) = 0$ will be the optimum solution in this interval. In order to prevent this solution from occurring, it is necessary to restrict $H(\omega)$ in such a manner that it can not be identically zero for all values of ω. A restriction we have already considered is that one of the operator coefficients be equal to a constant c. This restriction may be written as

$$\frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} H(\omega)\, e^{j\omega h(\ell + k)}\, d\omega = a_\ell = c \quad 0 \leq \ell \leq m \qquad 4.59$$

or

$$\frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} H_R \cos \omega h\,(\ell+k)\,d\omega + j\,\frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} H_R \sin \omega h(\ell+k)\,d\omega +$$

$$+\ j\,\frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} H_I \cos \omega h\,(\ell+k)\,d\omega - \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} H_I \sin \omega h(\ell+k)\,d\omega = c. \qquad 4.60$$

However, since H_R is even and H_I is odd, the second and third integrals in (4.60) must be zero. Further, since the integrands in the first and fourth integrals are even functions, the constrain equation (4.59) may be written as

$$4\pi\,h\int_{0}^{\pi/h} [H_R \cos \omega h(\ell+k) - H_I \sin \omega h(\ell+k)]\,d\omega = 4\pi^2 D \quad 0 \leq \ell \leq m \qquad 4.61$$

We therefore wish to minimize

$$I = 4\pi h\int_{0}^{\pi/h} \overline{X}X\,\overline{H}H\,d\omega \qquad 4.62$$

subject to the constraint (4.61). If we write $\overline{H}H = H_R^2 + H_I^2$, we can set

$$F = \overline{X}X\,[H_R^2 + H_I^2] + \lambda H_R \cos \omega h(\ell+k) - \lambda H_I \sin \omega h(\ell+k) \qquad 4.63$$

and minimize

$$4\pi h\int_{0}^{\pi/h} F\,d\omega$$

subject to no constraints.

The resulting equation is

$$\int_{0}^{\pi/h} \left\{ [2\overline{X}XH_R + \lambda \cos \omega h\ (\ell + k)]\ \delta H_R + [2\overline{X}XH_I - \lambda \sin \omega h\ (\ell + k)] \delta H_I \right\} d\omega = 0 \qquad 4.64$$

Again, as the number of terms in the operator goes to infinity, we may obtain the equations

$$H_R = \frac{-\lambda \cos \omega h\ (k+\ell)}{2\overline{X}X} \qquad 4.65$$

and

$$H_I = \frac{\lambda \sin \omega h\ (k+\ell)}{2\overline{X}X} \qquad 4.66$$

Then

$$H(\omega) = -\lambda \frac{e^{-j\omega h(k+\ell)}}{2\overline{X}X} \qquad 4.67$$

where is to be determined such that equation (4.61) is satisfied. Equation (4.67) may be written in the alternative form

$$H(\omega) = -\lambda \frac{e^{-j\omega h(k+\ell)}}{2\ \overline{P}P + 2\ \overline{R}R} \qquad 4.68$$

since $S(\omega) = 0$.

A formula somewhat similar to (4.68) may be obtained using equation (4.54). If, instead of assuming the signal to be zero in the fitting interval and non zero in some other interval, we assume that the signal is a random function with a constant power spectrum λ, we can obtain from equation (4.54)

$$H(\omega) = \frac{\lambda}{\lambda + \overline{P}P + \overline{R}R} \ . \qquad 4.69$$

An extension of this development to the case of the two dimensional linear operator will not be considered. Only a brief mention will be made of the case involving the prediction operator using more than one trace. Let us consider the two trace prediction operator of equation (2.13), where the spectrum of the error series was given by equation (3.22)

$$E(\omega) = X(\omega)\ [1-A(\omega)] - Y(\omega)\ B(\omega) \qquad 3.22$$

Therefore the sum of squared errors is given by

$$I = 4\pi h \int_{-\pi/h}^{\pi/h} \overline{E}E\ d\omega = 4\pi h \int_{-\pi/h}^{\pi/h} [\overline{X}-\overline{XA}-\overline{YB}][X-XA-YB]\ d\omega =$$

$$= 4\pi h \int_{-\pi/h}^{\pi/h} [\overline{X}X(1-A)(1-\overline{A}) - \overline{X}Y(1-\overline{A})\ B-X\overline{Y}\ (1-A)\ \overline{B} + \overline{Y}Y\overline{B}B]\ d\omega \qquad 4.70$$

The minimizing condition is then, that δI in (4.70) equal zero. The first and last terms in the integrand in (4.70) are positive. The middle two terms may be negative for some ω. If the cross spectrum $\overline{X}Y$ is approximately zero, as it would be if the x and y traces were uncorrelated, then the minimization of (4.70) consists essentially of minimizing the integrals

$$I_{xx} = \int_{-\pi/h}^{\pi/h} \overline{X}X(1-A)(1-\overline{A})\ d\omega \qquad 4.71$$

and

$$I_{yy} = \int_{-\pi/h}^{\pi/h} \overline{Y}Y\ \overline{B}B\ d\omega \qquad 4.72$$

Equation (4.72) will be minimized by $B = 0$. Therefore we would expect the b' coefficients in the linear operator to be zero.

If $\overline{X}Y$ is not approximately zero, then we must essentially minimize (4.71) and (4.72) while maximizing

$$I_{xy} = 2 \int_{-\pi/h}^{\pi/h} \mathrm{Re}\left[\overline{X}Y(1-\overline{A})B\right] \, d\omega \qquad 4.73$$

The various expressions for the optimum transfer functions in this section have been derived with no upper limit on the number of terms in the linear operator. Therefore, if the linear operator is restricted to $m+1$ terms, it will usually be necessary to approximate the optimum transfer function. But the optimum transfer function must be a complex Fourier series with M terms where M may approach ∞. Therefore the best approximation to the optimum transfer function $H'(\omega)$, in the least squares sense, will be a complex Fourier series with $m+1$ terms where the $m+1$ coefficients are the same as the corresponding set of $m+1$ coefficients of the series with M terms.

Unfortunately this simple approximation procedure does not, in general, conform to the requirements that the sum of squared errors, I, be a minimum. This may be seen from the following discussion. Let $H'(\omega)$ represent the optimum least squares transfer function, where $H'(\omega)$ may contain an infinite number of terms. The most general representation of $H'(\omega)$ is given by equations (4.42) and (4.43) or in a slightly different form

$$H'(\omega) = \frac{\overline{S}X}{\overline{X}X} \qquad 4.74$$

The equation for the sum of squared error is given by (4.36)

$$I = 4\pi h \int_{o}^{\pi/h} [\bar{S}S - \bar{S}XH - S\overline{XH} + \overline{XXH}H]\, d\omega \qquad 4.36$$

substituting (4.74) into (4.36) we obtain

$$I = 4\pi h \int_{o}^{\pi/h} \left\{ \bar{S}S + \bar{X}X\, [-H'H - \bar{H}'\bar{H} + \bar{H}H] \right\} d\omega =$$

$$= 4\pi h \int_{o}^{\pi/h} \left\{ \bar{S}S - \bar{X}X\bar{H}'H' + \bar{X}X[H' - H)(\bar{H}' - \bar{H})] \right\} d\omega \qquad 4.75$$

Since $\bar{S}S$ and $\bar{X}X\bar{H}'H'$ are constants, the minimization of I results in the equation

$$\delta \int_{o}^{\pi/h} \bar{X}X[(H'-H)(\bar{H}'-\bar{H})]\, d\omega = 0 \qquad 4.76$$

Thus H is a weighted least squares approximation to H' where the weighting factor is the input power spectrum.

4.4 The Ratio of Squares Method

Let us again consider the problem of detecting the reflections on a single seismogram trace. In this case we will assume that we know a particular interval, A to B, in which the signal is zero, and that we know another interval, C to D, in which the signal is non zero. We can therefore introduce the following criterion for the determination of the linear operator. We wish to minimize the output e_i in the interval

A to B, and we wish to maximize the output e_i in the interval C to D. Therefore, if we write

$$I_1 = \sum_{i=A}^{B} e_i^2 \qquad 4.77$$

and

$$I_2 = \sum_{i=C}^{D} e_i^2 \qquad 4.78$$

then we wish to minimize λ, where

$$\lambda = \frac{I_1}{I_2} \qquad 4.79$$

$$\text{But } \delta\lambda = \frac{I_2\,\delta I_1 - I_1\,\delta I_2}{I_2^2} = \frac{1}{I_2}\left(\delta I_1 - \lambda\delta I_2\right) = 0 \qquad 4.80$$

We therefore want

$$\delta I_1 - \lambda\,\delta I_2 = 0 \qquad 4.81$$

Using the notation of the first part of Section IV, we then have

$$\delta I_1 = \delta\,[\underline{a}^T\,\underline{x}_A^T\,\underline{x}_A\,\underline{a}\,] \qquad 4.82$$

and

$$\delta I_2 = \delta\,[\underline{a}^T\underline{x}_C^T\,\underline{x}_C\,\underline{a}\,] \qquad 4.83$$

Where $\underline{x}_A$ represents the matrix of the data from the interval A to B and $\underline{x}_C$ represents the matrix of the data from the interval C to D. Therefore,

$$\delta I_1 - \lambda\delta I_2 = \delta(\underline{a}^T)\;[\underline{x}_A^T\,\underline{x}_A - \lambda\,\underline{x}_C^T\,\underline{x}_C]\,\underline{a} = 0, \qquad 4.84$$

and since the a_s's are independent, we get

$$[\underline{x}_A^T \underline{x}_A - \lambda \underline{x}_C^T \underline{x}_C] \underline{a} = \underline{0} \qquad 4.85$$

The solution of this matrix equation will be the trivial solution unless the determinant

$$|[\underline{x}_A^T \underline{x}_A - \lambda \underline{x}_C^T \underline{x}_C]| = 0 \qquad 4.86$$

Thus we have a characteristic value problem. We must therefore determine the minimum characteristic number, since we wish to minimize λ, and inserting this value of λ_{min} in equation (4.85) we solve for $\underline{a}$.

4.5 Other Methods of Determining the Optimum Linear Operator or Filter

As was mentioned at the beginning of this section, it may be desirable that the linear operator have a certain transient response or impulse response. The impulse response is the easiest criterion from which to design a linear operator since, as we have seen, the operator coefficient series constitutes the impulse response. The transient response criterion may be more difficult to handle, depending upon the situation. As an example, suppose the signal occurs as a wavelet f(t), and we wish to design a linear operator which will contract this wavelet into an impulse. We can consider f(t) as the impulse response of some linear operator or filter with a transfer function $H(\omega)$. If this filter represents a minimum phase network,[7] then it has a physically realizable inverse with transfer function $H^{-1}(\omega)$. Therefore the linear operator corresponding to $H^{-1}(\omega)$ should compress f(t) approximately into an impulse.

Other criterion may be developed for the two dimensional linear operator or for the prediction operator using more than one trace. For instance, if the signal is in-phase on all the traces, it might be desirable to add the effects of the signal in-phase. This would correspond to in-phase mixing. Many other possible criteria can be developed.

V. Applications in a Controlled Experiment

5.1 Description of the Controlled Experiment

As a means of illustrating certain of the results of previous sections, a controlled set of data has been chosen consisting of a known signal pulse superimposed on a "noise" seismogram. See figure 4. This artificial seismic record appeared in a recent paper by Frank and Doty[11], where it was referred to as case No. 1. Frank and Doty illustrate the relative effects of filtering this record by three different filters. One filter is peaked at maximum signal, and two are peaked at maximum signal-to-noise ratio. See figure 5. The authors also show the effect of several types of mixing procedures. Frank and Doty conclude that the optimum filter for this record is the one peaked at maximum signal-to-noise ratio, with the sharper high pass cut off characteristics. They also find that graded or multiple mixing after filtering by the optimum filter tends to improve the "pick".

It will be of interest, in view of our previous results, to determine the optimum amplitude filter response according to several different criteria, and to compare these results with the Frank and Doty optimum filter. Unfortunately, their paper gives only the amplitude spectra of the signal and noise, and the amplitude filter responses. We are therefore necessarily limited in our comparison. Further, we will determine three linear operators by the least squares method of fitting in a noise interval, and compare these with the optimum filter determined from the noise spectrum. Finally, we will determine a linear operator which closely approximates the optimum Frank and Doty amplitude filter response.

5.2 Determination of the Optimum Least Squares Amplitude Filter Response from the Amplitude Spectra of the Signal and Noise

It has been shown in Section IV that for a small interval, if we are given $S(\omega)$ and $N(\omega)$ where $S(\omega)$ is the signal amplitude and phase spectrum and $N(\omega)$ is the noise amplitude and phase spectrum, we can determine an

$$H(\omega) = \frac{S(\omega)}{S(\omega) + N(\omega)}$$

which gives the signal exactly. This situation, however, is of little interest in the present case.

Equation (4.54) tells us that $H(\omega)$ should be given by

$$H(\omega) = \frac{\overline{S}S + \overline{S}P}{\overline{S}S + \overline{P}P + \overline{R}R + 2Re\,[\overline{S}P]} \qquad 4.54$$

However, we know only $|S|$ and $|N| = \sqrt{\overline{P}P + \overline{R}R}$. Therefore, if we assume that $\overline{S}P = S\overline{P} = 0$, then (4.54) may be written as

$$H(\omega) = \frac{|S|^2}{|S|^2 + |N|^2} \qquad 5.1$$

If $\rho = \frac{|S|}{|N|}$, where $\frac{|S|}{|N|}$ is the signal to noise ratio, then (5.1) may be written as

$$H(\omega) = \frac{\rho^2}{\rho^2 + 1} \qquad 5.2$$

It is apparent that this expression will be a maximum when ρ is a maximum. Using the amplitude spectra for the signal and noise as given by Frank and Doty, we obtain the optimum filter characteristics for the assumption $\overline{SP} = \overline{SP} = 0$ as shown in figure 6.

Using equations (4.52) and (4.53), and letting $S = |S| e^{j\theta_s}$ and $P = |P| e^{j\theta_p}$, we obtain the expression for the power response of the filter as

$$\overline{H}H = \frac{|S|^2[|S|^2 + 2|S||P| \cos(\theta_s - \theta_p) + |P|^2]}{[|S|^2 + 2|S||P| \cos(\theta_s - \theta_p) + |P|^2 + |R|^2]^2} \qquad 5.3$$

If we assume that all values of $\theta_s - \theta_p$ are equally likely, then we can write the expression for the expected value of $\overline{H}H$ as

$$Av(\overline{H}H) = \frac{1}{\pi} \int_0^\pi \frac{|S|^2[|S|^2 + 2|S||P| \cos x + |P|^2]\, dx}{[|S|^2 + 2|S||P| \cos x + |P|^2 + |R|^2]^2} =$$

$$= |S|^2 \left\{ \frac{(|S|^2+|P|^2+|R|^2)(|S|^2+|P|^2) - 4(|S|^2|P|^2)}{[(|S|^2+|P|^2+|R|^2)^2 - 4(|S|^2|P|^2)]^{3/2}} \right\} =$$

$$= |S|^2 \left\{ \frac{(|S|-|P|)^2(|S|+|P|)^2 + |R|^2(|S|^2+|P|^2)}{[(|S|-|P|)^2+|R|^2]^{3/2}[(|S|+|P|)^2+|R|^2]^{3/2}} \right\}$$

Reference 12/ 5.4

If $|R| = 0$, then

$$AV[\overline{H}H] = \frac{|S|^2}{|S|^2 - |P|^2} \quad \text{for } |S|^2 > |P|^2$$

$$= \infty \quad \text{for } |S|^2 = |P|^2$$

$$= \frac{|S|^2}{|P|^2 - |S|^2} \quad \text{for } |S|^2 < |P|^2$$

Using the Frank and Doty values of $|S|$ and $|N| = |P|$, we obtain a curve $Av[\overline{H}H]$ for $|R| = 0$. The square root of this curve is plotted in figure 6. If $\frac{|P|}{|R|} = 1$, then $|N|^2 = |P|^2 + |R|^2 = 2|P|^2$, and using equation (5.4) we obtain

$$\text{Av}[\bar{H}H] = |S|^2 \left\{ \frac{(|S|^2 + |N|^2)(|S|^2 + \frac{1}{2}|N|^2) - 2(|S|^2|N|^2)}{[(|S|^2+|N|^2)^2 - 2(|S|^2|N|^2)]^{3/2}} \right\} =$$

$$= \rho^2 \left\{ \frac{(\rho^2+1)(\rho^2 + \frac{1}{2}) - 2\rho^2}{[(\rho^2+1)^2 - 2\rho^2]^{3/2}} \right\} =$$

$$= \rho^2 \left\{ \frac{\rho^4 - \frac{1}{2}\rho^2 + \frac{1}{2}}{[\rho^4 + 1]^{3/2}} \right\} \qquad 5.6$$

Using the Frank and Doty values of $|S|$ and $|N|$, we obtain a curve Av $[\bar{H}H]$ for $|P| = |R|$. The square root of this curve is also plotted in figure 6.

5.3 Determination of the Optimum Least Square Amplitude Filter Response from the Amplitude Spectrum of the Noise

Equation (4.68) gives the optimum least squares transfer function for minimizing the power in the noise spectrum subject to the restriction that one of the operator coefficients equal a constant. The amplitude response of this transfer function is given by

$$|H(\omega)| = \frac{|\lambda|}{2\bar{N}N} \qquad 5.7$$

Using the Frank and Doty values of $|N|$, we obtain the curve plotted in figure 7. The value of $|\lambda|$ has not been determined from equation (4.59). A value of $|\lambda| = 2\cdot 10^{-3}$ has been assumed.

5.4 An Example of One Dimensional Linear Operators Determined by the Least Squares Method from the Input Noise

Three one dimensional linear operators have been determined from the input noise by the least squares method described in part 4.2. The fitting interval was from .3 to .6 seconds, as shown on the record in figure 4. Operators No. 1, No. 2, and No. 3 were determined in this time interval from traces T1, T2, and T3 respectively. All three operators we determined subject to the restrictions that $a_0 = 1.000$ and $a_1 = .000$. The coefficients for these operators are given in table 1. The amplitude filter responses of these operators are shown in figure 7.

5.5 A One Dimensional Linear Operator with Amplitude Response Approximating a 30 cps. Step.

We have seen in Section III that the real part of $H(\omega)$ is independent of the imaginary part if $k = \frac{m}{2}$ and m is an even integer. We can therefore design a linear operator in which $Re[H(\omega)] \doteq |H'(\omega)|$ and $Im[H(\omega)] = 0$, where $|H'(\omega)|$ is the amplitude response of some desired filter. For $|H'(\omega)|$ let us choose the 30 cps. step as shown in figure 8. We have seen that the equation which determines the operator coefficients is, in this case,

$$(a_s + a_{m-s}) = \frac{h}{\pi} \int_{-\pi/h}^{\pi/h} |H'(\omega)| \cos \omega h \left(s - \frac{m}{2}\right) d\omega. \qquad 5.8$$

But

$$|H'(\omega)| = \begin{cases} 0 & -\omega_o \leq \omega \leq \omega_o \\ 1 & \omega_o < |\omega| \end{cases} \qquad \text{where } \omega_o = 2\pi \times 30 \text{ cps.}$$

Therefore (5.8) becomes

$$(a_s + a_{m-s}) = \frac{h}{\pi} \int_{-\pi/h}^{-\omega_o} \cos \omega h(s - \tfrac{m}{2})\, d\omega + \frac{h}{\pi} \int_{\omega_o}^{\pi/h} \cos \omega h(s - \tfrac{m}{2})\, d\omega \qquad 5.9$$

or

$$(a_s + a_{m-s}) = \frac{2}{\pi} \frac{\left[-\sin \omega_o h(s - \frac{m}{2})\right]}{(s - \frac{m}{2})} \quad \text{for } s \neq \frac{m}{2} \qquad 5.10$$

and

$$(a_{\frac{m}{2}} + a_{\frac{m}{2}}) = \frac{2}{\pi} [\pi - \omega_o h] \quad \text{for } s = \frac{m}{2} \qquad 5.11$$

Since we also want $Im[H(\omega)] = 0$, we must require that the coefficients a_s be symmetrical about $a_{\frac{m}{2}}$. Therefore, equations (5.10) and (5.11), and the requirement of symmetry, determine the operator coefficients. Operator coefficients are given in the right hand column of table 1 for $m = 20$, $k = 10$, and $\omega_o = 2\pi \times 30$ cps. The amplitude response of this linear operator filter is shown in figure 8.

5.6 Discussion of Results

The amplitude responses of the optimum least squares filters, shown in figure 6, correspond in a general way with the response of the Frank and Doty optimum filter. Apparently, the assumption $\overline{SP} = S\overline{P} = 0$ gives the response most similar to that of Frank and Doty.

Operators No. 1, No. 2, and No. 3 shown in figure 7 are quite similar in the frequency range of the noise spectrum. This is to be expected since they were determined from the noise alone. The variation observed between the responses of these operators could easily be accounted for by variations in the noise spectra of the traces from which these operators were determined. All three operators have responses similar in general character to that of the optimum filter determined from N . As we have seen, H(ω), as determined by the least squares criterion, is a weighted least squares approximation to H'(ω) where the weighting factor is the spectrum. This is exactly the result shown in figure 7. The three operator responses best approximate the optimum response in the region of maximum noise spectrum. From the standpoint of signal amplification, the three operators would be better than the optimum filter. This is, of course, because the criterion of optimization is minimization of the noise spectrum. It is interesting to note that the responses of these three linear operators are quite similar to the response of the Frank and Doty optimum filter in the region of significant noise and signal spectra.

Figure 8 shows a 21 term linear operator approximation to a 30 cps step response. It is interesting that the amplitude response of this linear operator is almost identical with the response of the Frank and Doty optimum filter between 20 and 40 cps. There is no question but that a linear operator of approximately 20 terms could be determined which would have an amplitude response identical to that of the Frank and Doty optimum filter. More important, perhaps, is the fact that the 21 term linear operator in figure 8 is an approximation to a step response.

Thus, as the number of terms is increased, the response of the linear operator will approach a step. Further, the transfer function of this linear operator is real. Thus, the phase angle must be a multiple of π. In fact, for frequencies above 20 cps the transfer function is positive and the phase angle must be a multiple of 2π. This is a desirable filter property from the standpoint of not distorting the signal.

Frank and Doty determined that a combination of optimum filtering and mixing gave the best "pick". This result is perhaps trivial since the same signal was superimposed on the noise traces with no step out. Nevertheless, a similar filtering and mixing may be achieved by use of the two dimensional linear operator. The operator coefficients for each trace, in this case, would be proportional to those of the optimum one dimensional linear operator. The proportionality factor would determine the type of mixing.

TABLE 1

Operator	No. 1	No. 2	No. 3	Approximation to 30 cps Step
k	0	0	0	10
m	17	17	17	20
a_o	1.000	1.000	1.000	.032
a_1	.000	.000	.000	.032
a_2	-1.640	-2.170	-2.245	.023
a_3	.390	.690	1.267	.007
a_4	.870	.999	.600	-.016
a_5	-.133	.065	-.062	-.045
a_6	-.480	-.231	-.277	-.076
a_7	.280	-.163	-.110	-.105
a_8	.266	.000	.202	-.127
a_9	.048	.076	.197	-.150
a_{10}	.457	.102	.092	.850
a_{11}	-.386	-.217	-.070	-.150
a_{12}	-.522	.109	-.253	-.127
a_{13}	.350	.123	-.037	-.105
a_{14}	.533	.317	.237	-.076
a_{15}	-.332	-.213	.012	-.045
a_{16}	-.515	-.540	-.055	-.016
a_{17}	.627	.443	.143	.007
a_{18}				.023
a_{19}				.032
a_{20}				.032

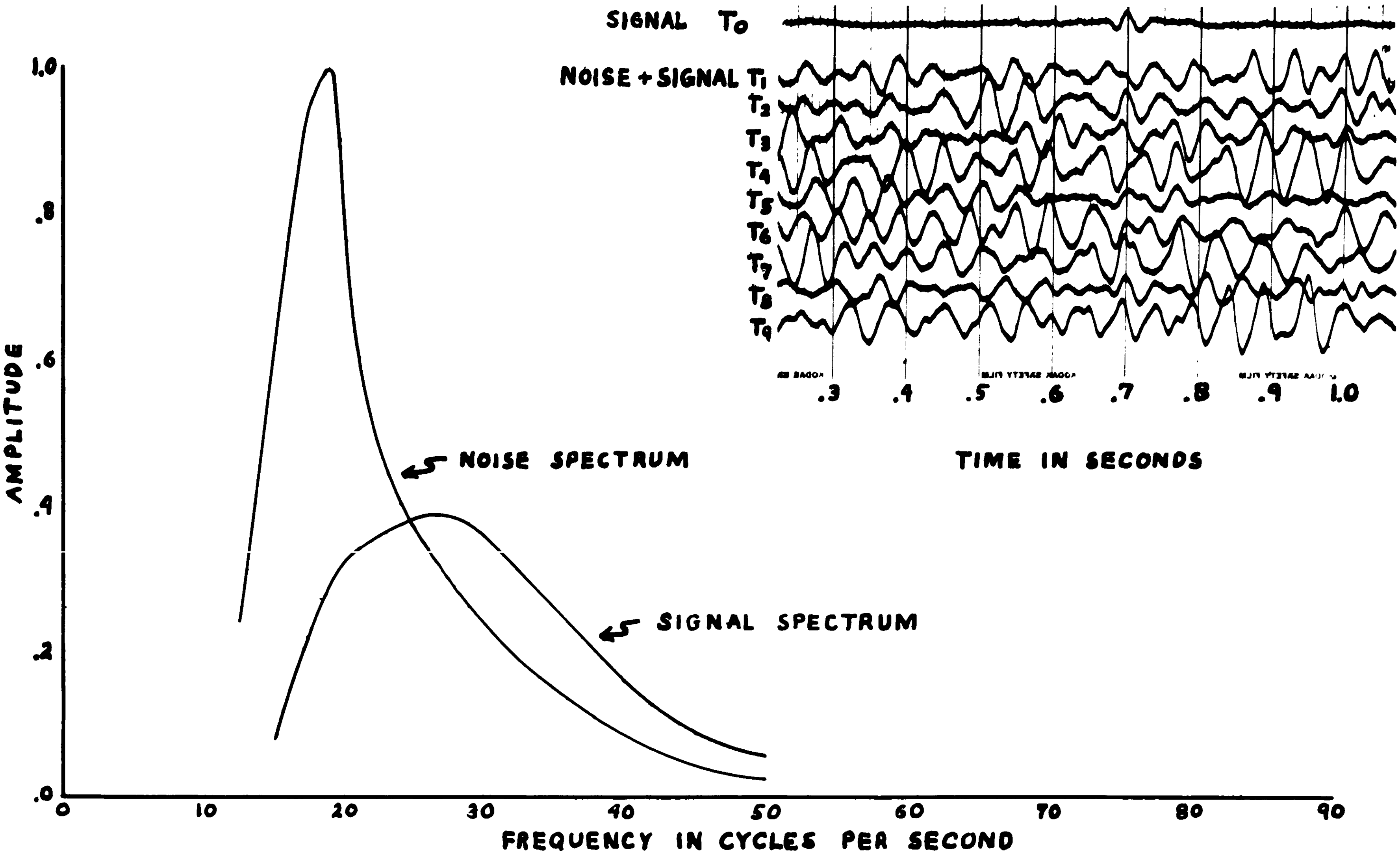

SIGNAL T0
NOISE + SIGNAL T1
T2
T3
T4
T5
T6
T7
T8
T9
.3
.4
.5
.6
.7
.8
.9
1.0
TIME IN SECONDS
1.0
.8
.6
.4
.2
.0
AMPLITUDE
NOISE SPECTRUM
SIGNAL SPECTRUM
0
10
20
30
40
50
60
70
80
90
FREQUENCY IN CYCLES PER SECOND

FIGURE 4

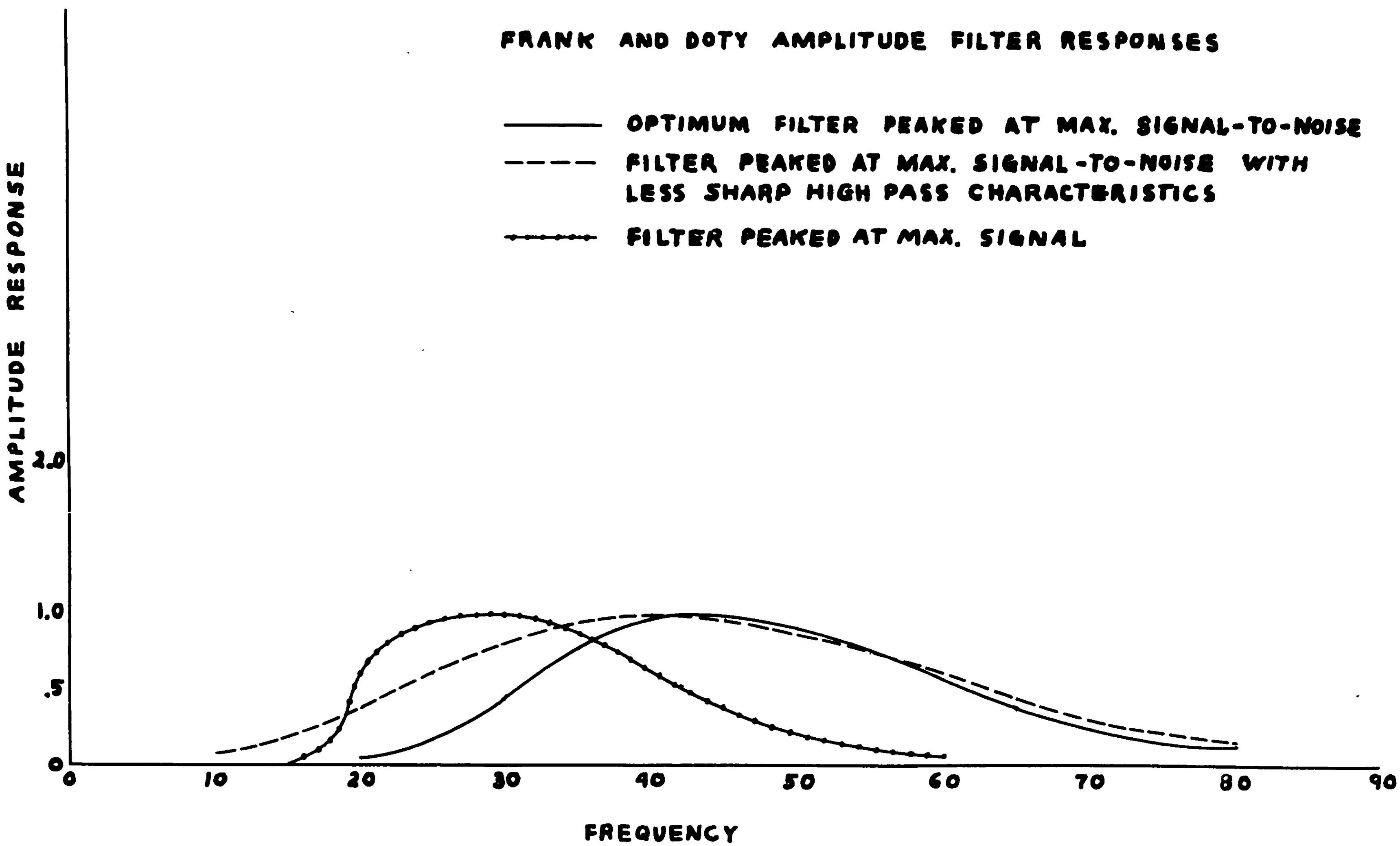

FRANK AND DOTY AMPLITUDE FILTER RESPONSES
OPTIMUM FILTER PEAKED AT MAX. SIGNAL-TO-NOISE
FILTER PEAKED AT MAX. SIGNAL-TO-NOISE WITH LESS SHARP HIGH PASS CHARACTERISTICS
FILTER PEAKED AT MAX. SIGNAL
AMPLITUDE RESPONSE
2.0
1.0
.5
0
0
10
20
30
40
50
60
70
80
90
FREQUENCY

FIGURE 5

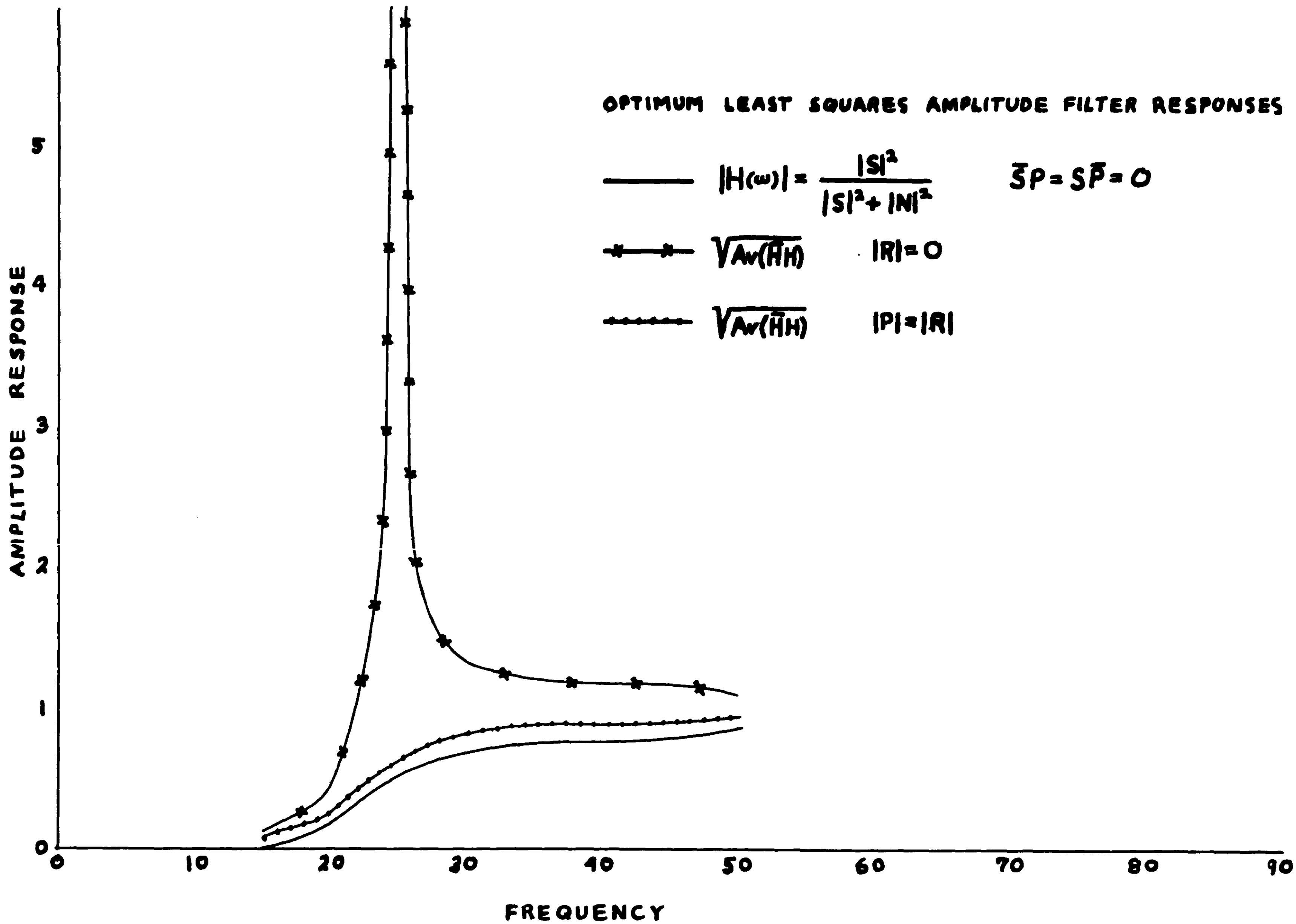

OPTIMUM LEAST SQUARES AMPLITUDE FILTER RESPONSES
$|H(\omega)| = \frac{|S|^2}{|S|^2 + |N|^2}$ $\bar{S}P = S\bar{P} = 0$
$\sqrt{Av(\bar{H}H)}$ $|R| = 0$
$\sqrt{Av(\bar{H}H)}$ $|P| = |R|$
AMPLITUDE RESPONSE
0
1
2
3
4
5
0
10
20
30
40
50
60
70
80
90
FREQUENCY

FIGURE 6

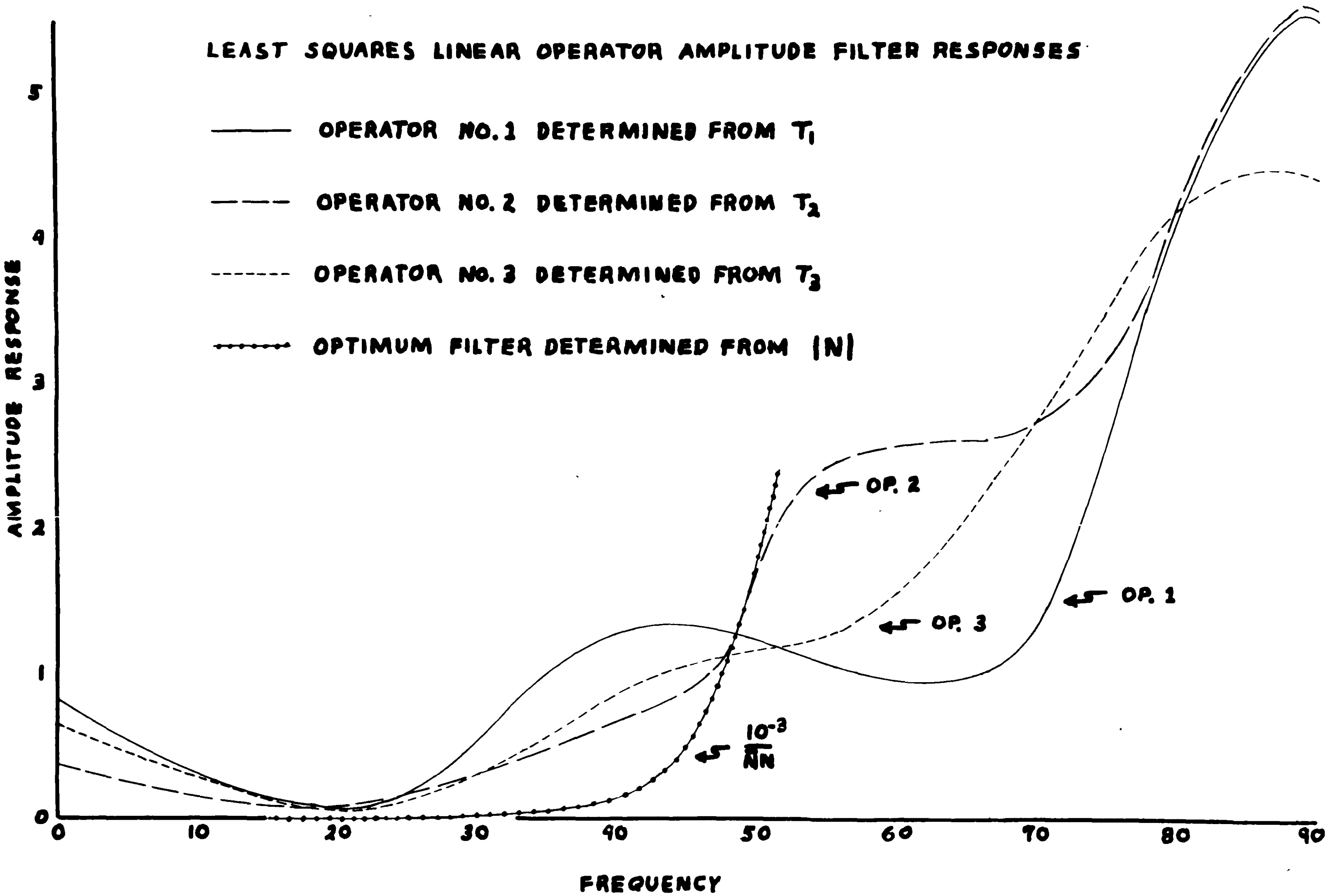
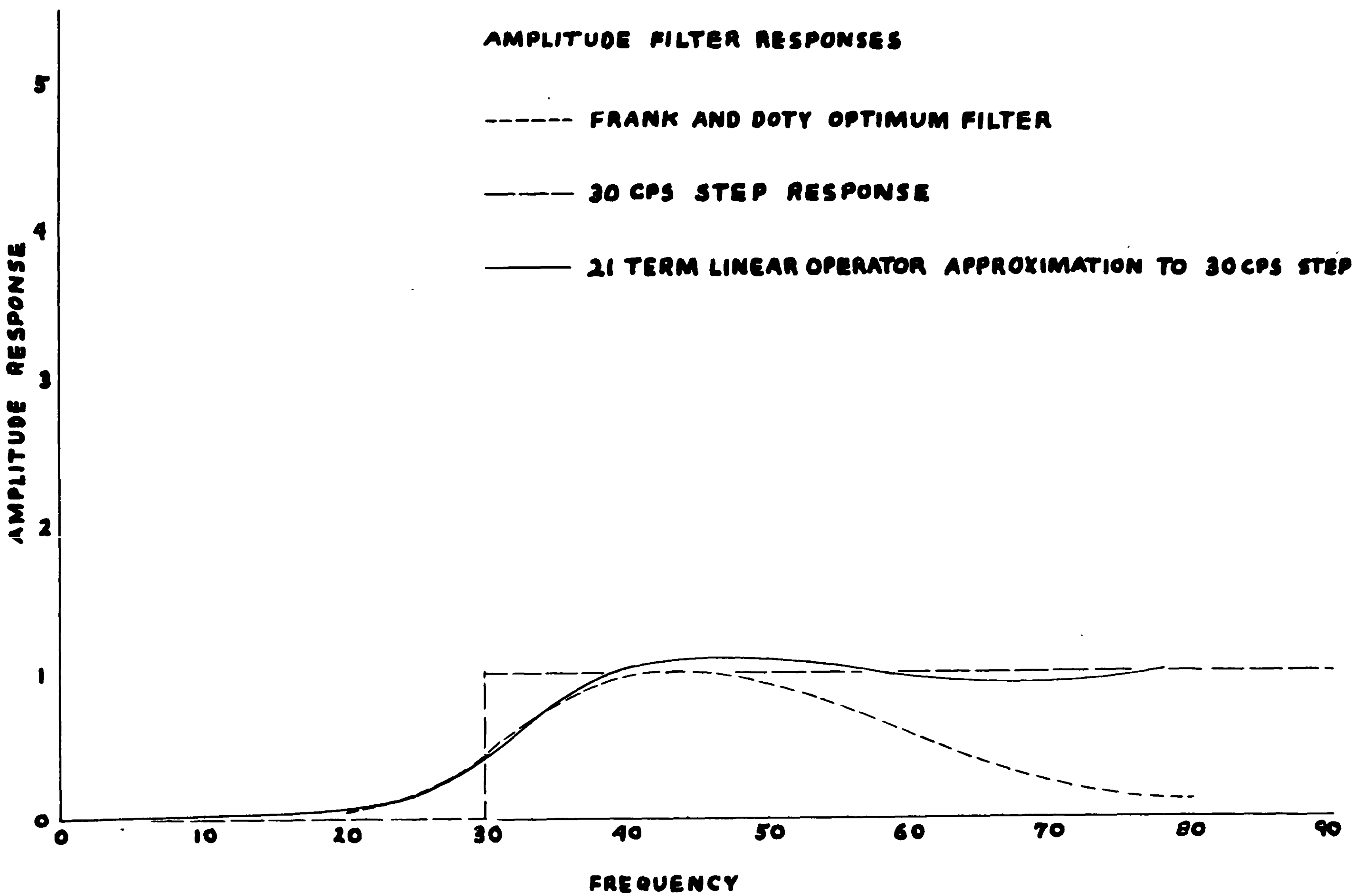

LEAST SQUARES LINEAR OPERATOR AMPLITUDE FILTER RESPONSES
OPERATOR NO. 1 DETERMINED FROM T_1
OPERATOR NO. 2 DETERMINED FROM T_2
OPERATOR NO. 3 DETERMINED FROM T_3
OPTIMUM FILTER DETERMINED FROM |N|
OP. 2
OP. 1
OP. 3
$\frac{10^{-3}}{\bar{N}N}$
AMPLITUDE RESPONSE
0
1
2
3
4
5
0
10
20
30
40
50
60
70
80
90
FREQUENCY

FIGURE 7



AMPLITUDE FILTER RESPONSES
FRANK AND DOTY OPTIMUM FILTER
30 CPS STEP RESPONSE
21 TERM LINEAR OPERATOR APPROXIMATION TO 30 CPS STEP
AMPLITUDE RESPONSE
0
1
2
3
4
5
0
10
20
30
40
50
60
70
80
90
FREQUENCY

FIGURE 8

VI. Conclusions and Recommendations

In Section III, it has been shown that the linear operator and the linear electrical filter are fundamentally equivalent. Thus they may both be used to perform linear operations on data. Each method has its computational advantages depending upon the form of the input data and the available facilities, such as a digital computer. However, the linear operator method possesses a basic advantage which outweighs computational considerations. This advantage may be stated as follows.

Based on knowledge of or assumptions as to the statistical nature of the data, criteria may be developed (as shown in Section IV) which give the best filter characteristics as a function of the data. Thus, given a criterion for optimizing the filter characteristics, we may determine the best filter for each set of input data. However, since the linear operator method is a mathematical method, we may use the same criterion to design automatically the best linear operator for each set of data, and having designed the linear operator, proceed to filter the data in the same operation.

In Section II, the concept of the Generalized Linear Operator indicated how the linear operator might be extended to operate simultaneously on a set of seismic records plus a set of geologic information. With the determination of the proper criterion for optimizing the linear operator, this method might provide an automatic technique for correlating reflections with subsurface structure. Thus, a greater share of the work of separating desired information from background interference might be assigned to the linear operator.

In conclusion, it should be re-emphasized that the linear operator and the linear electrical filter are computational techniques corresponding to linear mathematical operations, and that, therefore, the theory of linear operations applies to both. The fundamental problem is, thus, to determine the linear operation which best performs the desired function. More generally, the problem is to determine the best operation, whether linear or non linear, and, if need be, design a non linear computational technique. For cases in which the spectra of the signal and noise vary with time, it may be desirable to design a mathematical operation which also varies with time. Thus, the best computational form for use in the analysis of seismic records might conceivably be a discrete non linear operator with coefficients which are functions of time. It is therefore recommended, particularly in the case of seismograms, that a thorough study be made of the amplitude and phase spectra (and their variation with time) of the signal and noise in an attempt to determine the operation which will best separate the desired information from the background interference.

References

1. Wiener, Norbert, Extrapolation, Interpolation and Smoothing of Stationary Time Series, The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York (1950).

2. Kolmogoroff, A. "Interpolation und Extrapolation von Stationaren Zufalligen Folgen", Bull. Acad. Sci. (URSS) Ser. Math. 5, pp 3-14; 1941.

3. Wadsworth, G. P., Robinson, E. A., Bryan, J. G., and Hurley, P. M. "Detection of Reflections on Seismic Records by Linear Operators", Geophysics XVIII, 3, pp. 539-586; July 1953.

4. Churchill, R. V., Introduction to Complex Variables and Applications, McGraw-Hill Book Co., Inc., New York (1948).

5. Morse, P. M. and Feshbach, H., Methods of Theoretical Physics, Part 1, McGraw-Hill Book Co., Inc., New York (1953).

6. Bode, H. W., Network Analysis and Feedback Amplifier Design, D. Van Nostrand Co., Inc., New York (1952).

7. Bode, H. W. and Shannon, C. E., "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory", Institute of Radio Engineers Proceedings, 38, pp. 417-425; April 1950.

8. Courant, R. and Hilbert, D., Methods of Mathematical Physics, Vol. I, Interscience Publishers, Inc., New York (1953).

9. Hildebrand, F. B., Methods of Applied Mathematics, Prentice-Hall, Inc., New York, (1952).

10. Frazer, R. A., Duncan, W. J., and Collar, A. R., Elementary Matrices, Cambridge University Press, Cambridge, (1952).

11. Frank, H. R. and Doty, W. E. N., "Signal-to-Noise Ratio Improvements by Filtering and Mixing", Geophysics XVIII, 3, pp. 587-604; July 1953.

12. DeHaan, D. B., Nouvelles Tables d'Integrales Definies, G. E. Stechert and Co., New York (1939).

Unpublished References

13. Lee, Y. W., Class Notes, M. I. T. Course in "Statistical Theory of Communication", Fall Term, 1953.

14. Wadsworth, G. P. and Robinson, E. A., Directors, M. I. T. Geophysical Analysis Group, "A Prospectus on the Applications of Linear Operators to Seismology," Nov. 5, 1952.

15. M. I. T. Geophysical Analysis Group, Reports No. 3, 4, and 5, August 1953.

16. Stumpers, F. L., A Bibliography of Information Theory, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., Feb. 2, 1953.

Other References

17. Guillemin, E. A., The Mathematics of Circuit Analysis, The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York (1951).

18. Guillemin, E. A., Communication Networks, Vol. II, John Wiley and Sons., Inc., New York (1951).

Appendix I

General Notation Used

t	=	the independent variable; usually time.
$f(t)$	=	some function of t.
$x(t)$	=	trace, input, data, perturbed signal, or information.
$s(t)$	=	signal, or desired information.
$n(t)$	=	noise or background disturbance.
$p(t)$	=	interference or predictable element of the noise.
$r(t)$	=	random noise or random element of the noise.
$e(t)$	=	approximation to the signal $s(t)$ or $s(t + \quad)$.
ω	=	the angular frequency.
$F(\omega)$	=	some function of angular frequency ω.
$X(\omega)$	=	amplitude and phase spectrum of $x(t)$.
$S(\omega)$	=	amplitude and phase spectrum of $s(t)$.
$N(\omega)$	=	amplitude and phase spectrum of $n(t)$.
$P(\omega)$	=	amplitude and phase spectrum of $p(t)$.
$R(\omega)$	=	amplitude and phase spectrum of $r(t)$.
$E(\omega)$	=	amplitude and phase spectrum of $e(t)$.
a_s	=	operator coefficients of the one dimensional linear operator.
$a_{s_1 s_2}$	=	operator coefficients of the two dimensional linear operator.

General Notation Used (cont.)

a'_s, b'_s	=	operator coefficients of the prediction operator
h	=	spacing between data points
$i = t/h$	=	index corresponding to the independent variable.
$K(t)$	=	impulse response of a linear system.
$H(\omega)$	=	transfer function of a linear system.
$H'(\omega)$	=	optimum transfer function of a linear system.

Complex Variable Notation Used

j	=	$\sqrt{-1}$
$\overline{F}(\omega)$	=	the complex conjugate of $F(\omega)$.
$F(\omega)$	=	$F_R + j\,F_I$
F_R	=	$Re[F(\omega)]$ = the real part of $F(\omega)$.
F_I	=	$Im[F(\omega)]$ = the imaginary part of $F(\omega)$.
p	=	$p_R + j\,p_I = j\omega = j\omega_R - \omega_I$

Matrix Notation Used

c_{ij}	=	the i, j element of a matrix
$\underline{c}$	=	the matrix of elements c_{ij}
$\underline{c}^T$	=	the transpose of the matrix $\underline{c}$.
$\underline{c}^{-1}$	=	the inverse of the matrix $\underline{c}$
$\underline{c}$	=	c_{ij}.......... = a column matrix.
$\underline{c}$	=	c_{ij}.......... = a row matrix.
$\lvert\underline{c}\rvert$	=	the determinant of the matrix $\underline{c}$.

Calculus of Variations Notation Used

δF = the variation of F. Solutions of $\delta F = 0$ give stationary values of F. The minimum value of F is a stationary value.

Generalized Harmonic Analysis Notation[13]

Periodic Functions

$$\phi_{f_1,f_2}(\tau) = \frac{1}{T}\int_0^T f_1(t)\, f_2(t+\tau)\, dt$$ = cross correlation of f_1 and f_2.

$$\phi_{f_1,f_1}(\tau) = \frac{1}{T}\int_0^T f_1(t)\, f_1(t+\tau)\, dt$$ = auto correlation of f_1.

Aperiodic Functions

$$\phi_{f_1,f_2} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^T f_1(t)\, f_2(t+\tau)\, dt$$ = cross correlation of f_1 and f_2.

$$\phi_{f_1,f_1} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^T f_1(t)\, f_1(t+\tau)\, dt$$ = auto correlation of f_1.

Computational Formulae

$$\phi_{x,y}(\tau) \doteq \frac{1}{N+1}\sum_{n=o}^{N-\tau} x_n\, y_{n+\tau}$$ = cross correlation of x and y.

$$\phi_{x,x}(\tau) \doteq \frac{1}{N+1}\sum_{n=o}^{N-\tau} x_n x_{n+\tau}$$ = auto correlation of x.

Spectra and Correlation Relations

$$\Phi_{f_1,f_2}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_{f_1,f_2}(\tau)\, e^{-j\omega\tau}\, d\tau$$ = cross spectrum of f_1 and f_2.

$$\Phi_{f_1,f_1}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_{f_1,f_1}(\tau)\, e^{-j\omega\tau}\, d\tau$$ = auto power spectrum of f_1.

$$\phi_{f_1,f_2}(\tau) = \int_{-\infty}^{\infty} \Phi_{f_1,f_2}(\omega)\, e^{j\omega\tau}\, d\omega$$

$$\phi_{f_1,f_1}(\tau) = \int_{-\infty}^{\infty} \Phi_{f_1,f_1}(\omega)\, e^{j\omega\tau}\, d\omega$$

$$\phi_{f_1,f_2}(\tau) = \phi_{f_2,f_1}(-\tau)$$

$$\phi_{f_1,f_1}(\tau) = \phi_{f_1,f_1}(-\tau)$$

$$\Phi_{f_1,f_2}(\omega) = \overline{\Phi}_{f_2,f_1}(\omega)$$

$$\Phi_{f_1,f_1}(\omega) = \overline{\Phi}_{f_1,f_1}(\omega)$$

Appendix II

Given

$$H(\omega) = \sum_{s=0}^{m} a_s e^{-j\omega h(s+k)} \qquad \text{A2.1}$$

and

$$\overline{G}(\omega) = \sum_{\ell=0}^{m} b_\ell \, e^{j\omega h(\ell+k)} \qquad \text{A2.2}$$

Then the product $\overline{G}(\omega)\, H(\omega)$ is given by the double sum

$$\overline{G}(\omega)\, H(\omega) = \sum_{s=0}^{m} \sum_{\ell=0}^{m} a_s b_\ell \, e^{-j\omega h(s-\ell)} \qquad \text{A2.3}$$

However, it would be more convenient to express $\overline{G}(\omega)\, H(\omega)$ as a single sum

$$\overline{G}(\omega)\, H(\omega) = \sum_{\tau=-M}^{M} \beta_\tau \, e^{-j\omega h\tau} \qquad \text{A2.4}$$

where the coefficients β_τ are real.

Let us therefore determine the coefficients β_τ and the limits M.

Since the single summation expression for $\overline{G}(\omega)\, H(\omega)$ is a complex Fourier series, we may determine the coefficients β_τ by taking the Fourier Transform of $\overline{G}(\omega)\, H(\omega)$. That is,

$$\beta_\tau = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{G}(\omega)\, H(\omega)\, e^{j\omega h\tau}\, d\omega \qquad \text{A2.5}$$

But $\overline{G}(\omega)\, H(\omega)$ is given by equation (A2.3). Therefore

$$\beta_\tau = \frac{1}{2\pi} \sum_{s=0}^{m} \sum_{l=0}^{m} a_s b_l \int_{-\pi}^{\pi} e^{-j\omega h(s-l-\tau)} \, d\omega$$

Therefore, the only terms in the double sum which will contribute to β_τ are those for which $s-l-\tau = 0$. Therefore

$$\beta_o = \sum_{s=0}^{m} a_s b_s \qquad \text{A2.7a}$$

$$\beta_\tau = \sum_{s=-\tau}^{m} a_{s+\tau} b_s \qquad \tau < o \qquad \text{A2.7b}$$

$$\beta_\tau = \sum_{s=0}^{m-\tau} a_{s+\tau} b_s \qquad \tau > o \qquad \text{A2.7c}$$

If we consider the coefficients a_s and b_s as being equal to zero for $s > m$ and $s < o$, then we may write

$$\beta_\tau = \sum_{s=0}^{m} a_{s+\tau} b_s \qquad \text{for all } \tau \qquad \text{A2.8}$$

This equation is nothing but the expression for the cross-correlation function of the transient series a_s and b_s and we may therefore write

$$\beta_\tau = \phi_{ba}(\tau) \qquad \text{A2.9}$$

Also, the largest value of τ for which β_τ exists is $\tau = m$, and similarly the smallest is $\tau = -m$. Therefore the limits M on the summation in (A2.4) are m and -m for the upper and lower respectively. We can therefore write

$$\bar{G}(\omega)\ H(\omega) = \sum_{\tau=-m}^{m} \phi_{ba}(\tau)\ e^{-j\omega h\tau} \qquad \text{A2.10}$$

For the special case of $G(\omega) = H(\omega)$ we see that

$$\bar{H}(\omega)\ H(\omega) = \sum_{\tau=-m}^{m} \phi_{aa}(\tau)\ e^{-j\omega h\tau} \qquad \text{A2.11}$$

Appendix III

An Alternative Method of Electrical Filtering

The previous discussion of the discrete linear operator implies an alternative method of electrical filtering. If the input data were recorded on magnetic tape, an analog discrete linear operator could be constructed, consisting of magnetic pick-ups spaced along the tape at intervals of h. Each pick-up would then be connected through a variable resistor to the output, and the settings of the variable resistors would correspond to the values of the operator coefficients. The effective value of the spacing h could be changed by varying the speed of the tape during the initial recording of data. With the exception of the tape recording equipment, the components of this filter would be cheap variable resistors as compared with the more expensive components of an equivalent conventional filter.

BIOGRAPHICAL NOTE: Markwick Kern Smith, Jr.

Attended the Massachusetts Institute of Technology from June 1945 to February 1946. Served in U. S. Navy, February 1946 to February 1948, attending the Naval Electronics Service School for one year. Enrolled for one semester at the University of California at Los Angeles. Returned to the Massachusetts Institute of Technology, September 1948. Received the degree of S. B. in Geophysics in June 1951. Entered the Graduate School at the Massachusetts Institute of Technology in the Department of Geology and Geophysics in June 1951.

Presently employed as a Research Assistant in the Department of Geology and Geophysics on the Geophysical Analysis Group Research Project.