

UNCLASSIFIED

AD 243719

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

AD No. ~~243~~ 719
ASTIA FILE COPY

COLUMBIA UNIVERSITY

3

A CLASS OF OPTIMUM NON-LINEAR FILTERS
FOR QUANTIZED INPUTS

by
Frank Tung

1243719



ASTIA
OCT 7 1968
JIPDR

DEPARTMENT OF ELECTRICAL ENGINEERING
SCHOOL OF ENGINEERING
COLUMBIA UNIVERSITY
NEW YORK 27, N. Y.

ERD - TN - 60 - 769

Technical Report T-2/N

CU-8-60-AF-4140-EE

July 1, 1960

A CLASS OF OPTIMUM NON-LINEAR FILTERS

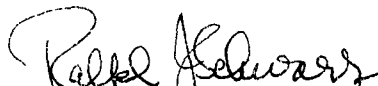
FOR QUANTIZED INPUTS

by

Frank Tung

Sponsored by and Prepared for
Electronics Research Directorate
Air Force Cambridge Research Labs
Air Force Research Division (AMDC)
United States Air Force
Bedford, Massachusetts
Contract No. AF 19(604)4140

Approved by:



Ralph J. Schwarz
Senior Investigator

The publication of this report does not necessarily constitute approval by the Air Force of the findings or conclusions contained herein.

Department of Electrical Engineering
COLUMBIA UNIVERSITY
New York 27, New York

Requests for additional copies by Agencies of the Department of Defense, their contractors, and other Government agencies should be directed to the:

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

Department of Defense contractors must be established for ASTIA services or have their 'need-to-know' certified by the cognizant military agency of their project or contract.

All other persons and organizations should apply to the:

U.S. DEPARTMENT OF COMMERCE
OFFICE OF TECHNICAL SERVICES
WASHINGTON 25, D.C.

ABSTRACT

Optimum non-linear filters belonging to Zadeh's class S_1 are considered. Attention is restricted to those systems whose present output is influenced only by a portion of the past input. The input signal consists of a message and noise both of which are stationary random processes. For this class of filters it is found that only the second order probability density functions of the message and the noise are necessary for obtaining the optimizing integral equation. It is assumed that the amplitude of the input time series is bounded and takes on discrete values at all times. This assumption is not too restrictive in practice since data supplied by computers and devices using digital read-out are quantized. By subjecting the joint probability density functions to a few mild restrictions, it is found that the optimizing integral equation reduces to a system of integral equations of the Wiener-Hopf type. By virtue of the assumptions made, the Fourier transforms of the kernels of these equations are rational functions. A method is developed for the solution of this set of simultaneous integral equations and three examples are given.

TABLE OF CONTENTS

	<u>Page</u>
Abstract	
Table of Contents	
I. Optimum Non-linear Filters	1
1. Introduction	1
2. Formulation of the Optimum Predictor	3
3. Quantized Inputs	7
4. Method of Solution	13
5. Examples	22
II. Multiple Predictions	32
References	36
Appendix A	
Appendix B	
Appendix C	

I. OPTIMUM NON-LINEAR FILTERS

1. Introduction

Broadly speaking, optimum filters and predictors are devices designed to produce, upon acting on the past of a time series, a value which estimates a desired value of some function of time in an optimum fashion according to some fixed error criterion. The most conventional optimum criterion has been the least square error criterion. Adoption of this criterion has often resulted in equations that can be handled analytically. When the input time series is stationary, the classic work of Wiener¹ has shown that the impulse response of the optimum stationary linear filter which acts on the infinite past of the input is the solution of the Wiener-Hopf equation. Subsequent to Wiener's publication, the subject of optimum linear filtering and prediction has been extended in many directions. Among them are the work of Zadeh and Ragazzini,² Boston,³ and Davis⁴ for continuous systems and the work of Blum,⁵ Chang,⁶ and others for sampled-data systems. By contrast, due to the lack of knowledge of characterizing a non-linear system as well as the inherent difficulty in any analytical treatment, relatively little work has been done with non-linear filters. Conditions under which optimum filters for the detection and prediction of signals are non-linear were studied by Laning⁷ and Drenick⁸ only for discrete data points. They have considered the case in which the message is a non-random linear combination of known time functions with unknown coefficients; in particular, polynomial functions were considered. On the other hand Zadeh⁹ has outlined an approach which is based on the consideration of a certain system of classes of non-linear filters. He has derived a sequence of linear integral equations for a class of optimum filters and has shown that

as the filter structure becomes more complicated, more and more information is necessary on the statistics of the input time series. A rather generalized result in optimum prediction and filtering using the least square error criterion was derived by Pugachev.¹⁰ The condition which he derived includes all of the optimizing equations as special cases.

Inasmuch as a linear filter is a degenerate case of a non-linear filter, improved results can usually be obtained by using a non-linear filter. However, we often find in practice that the amount of statistical data necessary for the design of non-linear filters far exceeds what is available. In addition, the complexity of the structure of the filters leads to problems that are unmanageable by analytical means. In order to circumvent this difficulty, it is common practice to make simplifying assumptions about the characteristics of the message and noise processes as well as to restrict attention to a certain class of filters.

The optimum filters to be considered in this report lie within a class of filters whose input-output relationship can be expressed as

$$y(t) = \int_0^t K[x(t - \tau), \tau] d\tau \quad 0 \leq t \leq T \quad (1.1)$$

Systems characterized by (1.1) have been designated by Zadeh⁹ as constituting the class $\mathcal{N}_1$. This class is a member of a sequence of classes of non-linear filters designated as $\mathcal{N}_1, \mathcal{N}_2, \dots$ classes. The class of linear filters is a subclass of $\mathcal{N}_1$. Consequently, every class in the hierarchy includes the class of linear filters. To cite an example in this class, consider

$$K[x(t - \tau), \tau] = f[x(t - \tau)]h(\tau) \quad (1.2)$$

where $f(x)$ is any function of its argument and $h(\tau)$ is the unit impulse response of a physically realizable linear filter. Then

$$y(t) = \int_0^t f(x(t-\tau))h(\tau)d\tau \quad 0 \leq t \leq T \quad (1.3)$$

represents the output of a system which consists of a cascade arrangement of an arbitrary zero memory non-linear device followed by a linear filter of memory size T .

As data supplied by computers and devices using digital read-out are quantized, it seems logical to incorporate this information in the design of filters. The following analysis shows that this a priori information can be used fruitfully when the form of the filter is given by (1.1).

2. Formulation of the Optimum Predictor *

Let it be supposed that the input signal $x(t)$ be the sum of two independent stationary random processes -- namely, the message $m(t)$ and the undesired noise $n(t)$,

$$x(t) = m(t) + n(t) \quad (1.4)$$

The problem is to find a filter belonging to the class $\mathcal{F}_T$ such that the difference between the actual output from the filter and the desired output is minimized in some sense. Let $q[m(t+T)]$ ($q \geq 0$), represent the desired output where q is any function of its argument whose ensemble mean is zero for all t , and let $e(t)$ be the error between the desired output and the actual output.

$$e(t) = y(t) - q[m(t+T)] \quad (1.5)$$

* Predictor is used here as a general term to indicate prediction in the presence of noise.

As in Wiener's theory, it will be postulated that the predictor is optimum when 1) the ensemble mean of $\epsilon(t)$ is equal to zero for all t , and 2) the ensemble variance of $\epsilon(t)$ is a minimum. Let us denote the ensemble average by $\langle \rangle_{AV}$. The filter is therefore optimum when

$$\langle \epsilon(t) \rangle_{AV} = 0 \quad (1.6)$$

and

$$\langle \epsilon^2(t) \rangle_{AV} = \text{minimum} \quad (1.7)$$

From (1.5), we have

$$\begin{aligned} \langle \epsilon^2(t) \rangle_{AV} &= \left\langle \left\{ \int_0^T K[x(t-\tau), \tau] d\tau \right\}^2 \right\rangle_{AV} \\ &\quad - 2 \left\langle \int_0^T K[x(t-\tau), \tau] d\tau \times q[m(t+\alpha)] \right\rangle_{AV} \\ &\quad + \left\langle \{q[m(t+\alpha)]\}^2 \right\rangle_{AV} \end{aligned} \quad (1.8)^*$$

Upon expansion, (1.8) becomes

$$\begin{aligned} \langle \epsilon^2(t) \rangle_{AV} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T \int_0^T K(x_1, \tau_1) \times K(x_2, \tau_2) \times p(x_1, x_2) d\tau_1 d\tau_2 dx_1 dx_2 \\ &\quad - 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T K(x_1, \tau_1) \times q(m_\alpha) \times p(x_1, m_\alpha) d\tau_1 dm_\alpha dx_1 \\ &\quad + \int_{-\infty}^{\infty} q^2(m_\alpha) p(m_\alpha) dm_\alpha \end{aligned} \quad (1.9)$$

* We assume here that enough time has elapsed since the signal started; consequently, the upper limit of the integral is replaced by T .

where in (1.9) we have circumvented the lengthy notation by letting

$$x_1 = x(t - \tau_1)$$

$$x_2 = x(t - \tau_2)$$

$$m_\alpha = m(t + \alpha)$$

$$\begin{aligned} p(x_1, x_2) &= p[x(t - \tau_1), x(t - \tau_2); \tau_1 - \tau_2] \\ &= \text{joint probability density function of } x(t - \tau_1) \text{ and } x(t - \tau_2) \end{aligned}$$

$$\begin{aligned} p(x_1, m_\alpha) &= p[x(t - \tau_1), m(t + \alpha); \tau_1 + \alpha] \\ &= \text{joint probability density function of } x(t - \tau_1) \text{ and } m(t + \alpha) \end{aligned}$$

$$\begin{aligned} p(m_\alpha) &= p[m(t + \alpha)] \\ &= \text{first order probability density function of } m(t + \alpha) \end{aligned}$$

Our aim is to find the kernel $K[x(t - \tau), \tau]$ which minimizes (1.9). The minimization of $\langle \epsilon^2(t) \rangle_{AV}$ is accomplished by the usual technique of variational calculus. The variation δI corresponding to an admissible variation δK in K is

$$\begin{aligned} \delta I &= \beta^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T \int_0^T \delta K(x_1, \tau_1) \delta K(x_2, \tau_2) \times p(x_1, x_2) d\tau_1 d\tau_2 dx_1 dx_2 \\ &+ 2\beta \int_{-\infty}^{\infty} \int_0^T \delta K(x_1, \tau_1) d\tau_1 dx_1 \times \left\{ \int_{-\infty}^{\infty} \int_0^T K(x_2, \tau_2) p(x_1, x_2) d\tau_2 dx_2 \right. \\ &\left. - \int_{-\infty}^{\infty} q(m_\alpha) p(m_\alpha, x_1) dm_\alpha \right\} \end{aligned} \quad (1.10)$$

By evaluating $\frac{d\delta I}{d\beta}$ and setting it to zero at $\beta = 0$, we obtain, for all admissible $\delta K(x_1, \tau_1)$,

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_0^T \delta K(x_1, \tau_1) d\tau_1 dx_1 \\
& \left\{ \int_{-\infty}^{\infty} \int_0^T K(x_2, \tau_2) \times p(x_1, x_2) d\tau_2 dx_2 \right. \\
& \quad \left. - \int_{-\infty}^{\infty} q(m_\alpha) \times p(m_\alpha, x_1) dm_\alpha \right\} = 0 \quad (1.11)
\end{aligned}$$

The desired optimizing integral equation for the kernel K is therefore of the form

$$\begin{aligned}
\int_{-\infty}^{\infty} q(m_\alpha) \times p(m_\alpha, x_1) dm_\alpha &= \int_{-\infty}^{\infty} \int_0^T K(x_2, \tau_2) \times p(x_1, x_2) d\tau_2 dx_2 \\
& \quad 0 \leq \tau_1 \leq T \quad (1.12)
\end{aligned}$$

In order to show that (1.12) gives a true minimum rather than a mere stationary solution, we proceed as follows. Let us consider a filter $Q(x, \tau)$ which is different from $K(x, \tau)$. Inspection of (1.9) shows that the mean square error resulting from using the filter $Q(x, \tau)$ can be arranged to read

$$\begin{aligned}
\langle \epsilon^2(t) \rangle_{AV} &= \int_{-\infty}^{\infty} q^2(m_\alpha) p(m_\alpha) dm_\alpha \\
&+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T K(x_1, \tau_1) q(m_\alpha) \times p(m_\alpha, x_1) d\tau_1 dm_\alpha dx_1 \\
&+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T \int_0^T [Q(x_1, \tau_1) - K(x_1, \tau_1)] \cdot [Q(x_2, \tau_2) - K(x_2, \tau_2)] \\
&\quad \times p(x_1, x_2) d\tau_1 d\tau_2 dx_1 dx_2 \quad (1.13)
\end{aligned}$$

which is greater than or equal to the mean square error resulting from using the filter $K(x, \tau)$, namely

$$\int_{-\infty}^{\infty} q^2(m_\alpha) p(m_\alpha) dm_\alpha - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^T K(x_1, \tau_1) q(m_\alpha) \times p(m_\alpha, x_1) d\tau_1 dm_\alpha dx_1 \quad (1.14)$$

since the last term in (1.13) is non-negative. Consequently, we conclude that the stationary solution does, in fact, give a minimum mean square error.

3. Quantized Inputs

As mentioned in Section I.1, the amplitude of the input time series may be assumed to be bounded and discrete at all times in view of the nature of the measuring device. This is illustrated in Figure 1, where the observed input is denoted by $x^*(t)$. It is noted that no assumption is made regarding the amplitudes of the continuous processes $m(t)$ and $n(t)$. Our analysis includes the special case in which the amplitudes of both the message and the noise processes are bounded and discrete for all t .

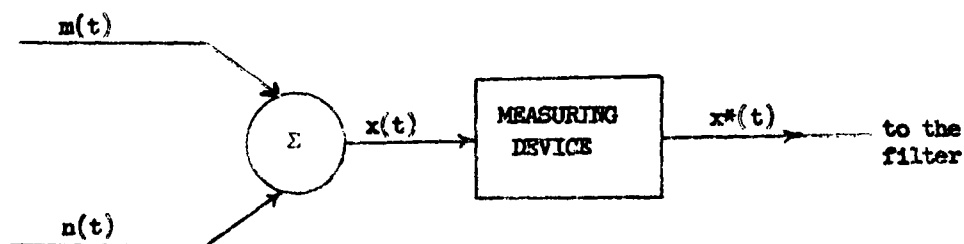


Fig. 1 Block Diagram Showing the Actual Inputs: $x^*(t)$
is a Quantized, Time-continuous Signal

If we let the number of the amplitudes of the input signal be $2N$, then the joint probability density functions $p(x_1^*, x_2^*)$ and $p(m_\alpha, x_1^*)$ can be expressed as

$$p(x_1^*, x_2^*) = \sum_{i=-N}^{+N} \sum_{\substack{j=-N \\ i,j \neq 0}}^{+N} A_{i,j}(|\tau_1 - \tau_2|) \delta(x_1^* - c_i) \delta(x_2^* - c_j) \quad (1.15)$$

$c_i = -c_{-i}$

and

$$p(m_\alpha, x_1^*) = \sum_{\substack{i=-N \\ i \neq 0}}^{+N} f_i(m_\alpha; \tau_1 + \alpha) \delta(x_1^* - c_i) \quad (1.16)$$

where $A_{i,j}(|\tau_1 - \tau_2|)$ is the probability that $x^*(t - \tau_1)$ equals c_i and $x^*(t - \tau_2)$ equals c_j , and $\int_{\mu}^{\mu+\Delta\mu} f_i(m_\alpha; \tau_1 + \alpha) dm_\alpha$ denotes the probability that $x^*(t - \tau_1)$ equals c_i and $m(t + \alpha)$ lies between μ and $\mu + \Delta\mu$. Substituting (1.15) and (1.16) into (1.11) and equating the corresponding coefficients of the delta functions associated with x_1 , we have

$$\int_{-\infty}^{\infty} f_i(m_\alpha; \tau_1 + \alpha) \delta(q(m_\alpha)) dm_\alpha = \sum_{\substack{j=-N \\ j \neq 0}}^{+N} \int_0^T A_{i,j}(|\tau_1 - \tau_2|) \delta(c_j, \tau_2) d\tau_2 \quad (1.17)$$

$0 \leq \tau_1 \leq T; \quad i = -N, \dots, -1, 1, \dots, N$

For arbitrary $A_{i,j}(|\tau_1 - \tau_2|)$ and $f_i(m_\alpha; \tau_1 + \alpha)$, the simultaneous set of integral equations in (1.17) is far too complex for any analytical solution. In fact, it is improbable that a direct solution can always be found. In order to obtain a reasonable solution, the following assumptions regarding the probability density functions are made:

A. The joint probability density function of the message and of the noise are symmetrical with respect to their arguments as well as symmetrical with respect to the origin, namely

$$p_m(m_1, m_2) = p_m(m_2, m_1) = p_m(-m_1, -m_2) \quad (1.18)$$

$$p_n(n_1, n_2) = p_n(n_2, n_1) = p_n(-n_1, -n_2) \quad (1.19)$$

It follows that $p(x_1, x_2)$ which is given by

$$p(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_m(m_1, m_2) \times p_n(x_1 - m_1, x_2 - m_2) dm_1 dm_2 \quad (1.20)$$

also has the same property. On the other hand $p(m_\alpha, x_1)$ which is given by

$$p(m_\alpha, x_1) = \int_{-\infty}^{\infty} p_m(m_\alpha, m_1) p_n(x_1 - m_1) dm_1 \quad (1.21)$$

is only symmetrical about the origin. In terms of the expressions in (1.16) and (1.17), this assumption can be written symbolically as

$$A_{i,j}(|\tau_1 - \tau_2|) = A_{j,i}(|\tau_1 - \tau_2|) = A_{-i,-j}(|\tau_1 - \tau_2|) \quad (1.22)$$

$$f_i(m_\alpha; \tau_1 + \alpha) = f_{-i}(-m_\alpha; \tau_1 + \alpha) \quad (1.23)$$

Let us define

$$\int_{-\infty}^{\infty} f_i(m_\alpha; \tau_1 + \alpha) q(m_\alpha) dm_\alpha = \mu_i(\tau_1 + \alpha) \quad (1.24)$$

Since $\langle q(m_\alpha) \rangle_{AV} = 0$, it follows from (1.23) that

$$z_1(\tau_1 + \alpha) = -z_{-1}(\tau_1 + \alpha) \quad (1.25)$$

B. For all i and j , the quantities $[A_{i,j}(|\tau_1 - \tau_2|)] = A_{i,-j}(|\tau_1 - \tau_2|)]$ are sums of a finite number of decaying exponentials.* It is the logical extension of the usual assumption that the autocorrelation function which is given by

$$R_x(|\tau_1 - \tau_2|) = \langle x_1 x_2 \rangle_{AV} = \sum_i \sum_j c_i c_j A_{i,j}(|\tau_1 - \tau_2|)$$

is the sum of exponential functions.

In view of assumption A, the optimum filter can be represented by the structure shown in Figure 2.

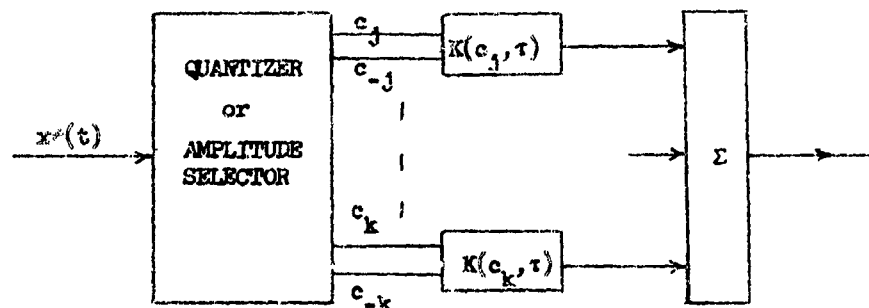


Fig. 2 Schematic Representation of a Non-linear Filter

* This is a weaker condition than that which requires all $A_{i,j}(|\tau_1 - \tau_2|)$ to be decaying exponentials.

In Fig. 1, $\frac{K(c_j, \tau)}{c_j}$ are the unit impulse responses of linear filters. The system becomes linear when

$$\frac{K(c_j, \tau)}{c_j} = \frac{K(c_k, \tau)}{c_k} \quad \text{for all } j \text{ and } k. \quad (1.26)$$

To see the validity of Fig. 1, it is necessary to show that

$$K(c_j, \tau) = -K(c_{-j}, \tau) \quad j = 1, 2, \dots, N \quad (1.27)$$

From (1.6) we conclude that

$$\langle y(t) \rangle_{AV} = 0 \quad (1.28)$$

Consequently

$$\langle K[x^*(t - \tau), \tau] \rangle_{AV} = 0 \quad (1.29)$$

Let f_j = the probability that $x^*(t - \tau)$ takes on value c_j .
Then

$$f_j = f_{-j} \quad j = 1, 2, \dots, N \quad (1.30)$$

which can be deduced from our previous assumptions on the process $x^*(t)$. Eq. (1.27) becomes

$$\sum_{j=1}^N 2f_j \left\{ K(c_j, \tau) + K(c_{-j}, \tau) \right\} = 0 \quad (1.31)$$

In order for (1.31) to hold for any set of f_j satisfying $\sum_j f_j = 1$, it is necessary that

$$K(c_j, \tau) + K(c_{-j}, \tau) = 0 \quad j = 1, 2, \dots, N \quad (1.32)$$

which agrees with (1.27). Eq. (1.18) can now be written as

$$z_i(\tau_1 + \alpha) = \sum_{\substack{j=-N \\ j \neq 0}}^{+N} \int_0^T [A_{i,j}(|\tau_1 - \tau_2|) - A_{i,-j}(|\tau_1 - \tau_2|)] \times K(c_j, \tau_2) d\tau_2$$

$$0 \leq \tau_1 \leq T; \quad i = -N, \dots, -1, 1, \dots, N \quad (1.33)$$

By using (1.22) and (1.25), it is seen that N of the $2N$ equations in (1.33) are redundant. Let us replace τ_1 and τ_2 by t and τ respectively, and let

$$w_{ij}(|t - \tau|) = A_{i,j}(|t - \tau|) - A_{i,-j}(|t - \tau|) \quad (1.34)$$

$$K_j(\tau) = K(c_j, \tau) \quad (1.35)$$

Eq. (1.33) finally reduces to

$$z_i(t + \alpha) = \sum_{j=1}^N \int_0^T w_{ij}(|t - \tau|) K_j(\tau) d\tau$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N \quad (1.36)$$

From (1.9) we see that the minimum mean square error is given by

$$\text{Min} \langle \epsilon^2 \rangle_{AV} = \langle q^2[m(t + \alpha)] \rangle_{AV} - 2 \sum_{i=1}^N \int_0^T z_i(t + \alpha) K_i(t) dt \quad (1.37)$$

4. Method of Solution

It was shown in the previous section that $w_{ij}(|\tau|)$ is the sum of a number of decaying exponential functions; hence its Fourier transform, defined as

$$w_{ij}(\lambda^2) = \int_{-\infty}^{\infty} w_{ij}(|\tau|) e^{-j\omega\tau} d\tau \quad \lambda = 2\pi f \quad (1.38)$$

is a rational function of λ^2 . Suppose*

$$w_{ij}(\lambda^2) = \frac{D_{ij} P_{ij}(\lambda^2)}{Q(\lambda^2)} \quad (1.39)$$

where D_{ij} are constants, $Q(\lambda^2)$ is of order d and $P_{ij}(\lambda^2)$ is of order n_{ij} ($n_{ij} < d$). We shall derive in this section a necessary and sufficient condition under which a unique, absolutely integrable solution of (1.36) exists. $K_j(t)$ is absolutely integrable if

$$\int_0^{\infty} |K_j(t)| dt < \infty \quad \text{for } j = 1, 2, \dots, N \quad (1.40)$$

which is the usual stability condition for linear systems. A system is, for our purposes, defined to be stable if all bounded inputs result in bounded outputs. Our result will also indicate that a formal solution can always be obtained if the restriction imposed by (1.40) is removed. The approach here is first to transform (1.36) to a simpler system of integral equations. It is then shown that the solution of the modified system of integral equations does, in fact, satisfy (1.36). When convenient, the following notation will be used:

* The common denominator of all $w_{ij}(\lambda^2)$ has been used.

- $[D]$ = an $N \times N$ square matrix whose elements are D_{ij} (this matrix is assumed to be non-singular)
 $[D]^{-1}$ = inverse of $[D]$
 $\underline{K}(t)$ = a column matrix whose elements are time functions $K_j(t)$, $j = 1, 2, \dots, N$.
 $\underline{z}(t|\alpha)$ = a column matrix whose elements are time functions. $z_j(t|\alpha)$, $j = 1, 2, \dots, N$.
 $\underline{y}(t|\alpha)$ = a column matrix whose elements are time functions, $y_j(t|\alpha)$, $j = 1, 2, \dots, N$.
 $\hat{w}(|\tau|)$ = inverse Fourier transform of $1/Q(\lambda^2)$
 $\left[\int_0^T w(|t-\tau|) \underline{K}(\tau) d\tau \right]$ = a column matrix whose elements are $\int_0^T \hat{w}(|t-\tau|) K_j(\tau) d\tau$, $j = 1, 2, \dots, N$.
 $[DP(\frac{t^2}{2})]$ = an $N \times N$ square matrix whose elements are linear operators, $D_{ij} P_{ij}(-\frac{d^2}{dt^2})$; $i, j = 1, 2, \dots, N$.

Let $\underline{y}(t|\alpha)$ be any solution satisfying the following system of differential equations:

$$\underline{z}(t|\alpha) = [D P(-\frac{d^2}{dt^2})][D]^{-1} \underline{y}(t|\alpha) \quad 0 \leq t \leq T \quad (1.41)$$

and let the modified set of integral equations be given by

$$\underline{y}(t|\alpha) = [D] \left[\int_0^T \hat{w}(|t-\tau|) \underline{K}(\tau) d\tau \right] \quad 0 \leq t \leq T \quad (1.42)$$

We will now show that the solution of (1.42) also satisfies (1.36)

To see this we pre-multiply both sides of (1.42) by $[D]^{-1}$, which becomes

$$\left[\int_0^T \hat{w}(|t-\tau|) \underline{K}(\tau) d\tau \right] = [D]^{-1} \underline{y}(t|\alpha) \quad 0 \leq t \leq T \quad (1.43)$$

Operating on both sides of (1.43) by $[D P (-\frac{d^2}{dt^2})]$ yields

$$[DP(-\frac{d^2}{dt^2})][\int_0^T \hat{w}(|t-\tau|) \underline{K}(\tau) d\tau] = [DP(-\frac{d^2}{dt^2})][D]^{-1} \underline{y}(t+\alpha) \quad 0 \leq t \leq T \quad (1.44)$$

The left side of (1.44) is a column matrix whose elements are

$$\sum_{j=1}^N D_{ij} P_{ij} (-\frac{d^2}{dt^2}) \int_0^T \hat{w}(|t-\tau|) K_j(\tau) d\tau \quad i = 1, 2, \dots, N$$

which, as shown in Appendix A, becomes

$$\sum_{j=1}^N \int_0^T w_{ij}(|t-\tau|) K_j(\tau) d\tau; \quad i = 1, 2, \dots, N$$

Equation (1.44) therefore reduces to

$$\sum_{j=1}^N \int_0^T w_{ij}(|t-\tau|) K_j(\tau) d\tau = z_i(t+\alpha); \quad 0 \leq t \leq T \quad i = 1, 2, \dots, N \quad (1.45)$$

which is, in fact, (1.36).

In section 1.2 we have shown that any solution (if it is not unique) of (1.36) will give the same mean square error. It is, therefore, immaterial whether other solutions of (1.36) exist which do not satisfy (1.40). Nevertheless, as we will show in Appendix C, all solutions of (1.36) are necessarily the solutions of (1.40).

We shall now investigate the solution of the modified system of integral equations (1.42). It is shown in Appendix A that if a solu-

tion exists, it satisfies a set of simultaneous differential equations

$$Q\left(-\frac{d^2}{dt^2}\right) \underline{y(t+\alpha)} = [D]\underline{K(t)} \quad 0 \leq t \leq T \quad (1.46)$$

The solution of (1.46) can be written as

$$\underline{K(t)} = Q\left(-\frac{d^2}{dt^2}\right)[D]^{-1} \underline{y(t+\alpha)} \quad 0 \leq t \leq T \quad (1.47)$$

Since the information pertaining to the derivatives of $\underline{K(t)}$ at $t = 0$ and $t = T$ is not included in the derivation of (1.46), we find that certain conditions on $\underline{y(t+\alpha)}$ are necessary in order that the $\underline{K(t)}$ so obtained from (1.47) do satisfy (1.42). Those conditions are obtained by substituting (1.46) into (1.42) and solving the resultant equations as an identity. The development parallels that appearing in Appendix 2 of Davenport and Root.¹¹

Let us first establish a useful result. From the definition of $\hat{w}(t)$, that is,

$$\hat{w}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\lambda t}}{Q(\lambda^2)} d\lambda \quad (1.48)$$

we obtain the corresponding differential equation

$$Q\left(\frac{d^2}{dt^2}\right)\hat{w}(t) = \delta(t) \quad (1.49)$$

In particular, let

$$Q(\lambda^2) = \sum_{k=0}^d q_{2k}(\lambda)^{2k} \quad (1.50)$$

Eq. (1.49) can be expressed as*

$$\sum_{k=0}^d (-1)^k q_{2k} \hat{w}^{(2k)}(|t|) = \delta(t) \quad (1.51)$$

This relation will be used later on.

Using (1.50), Eq. (1.46) can be rewritten as

$$[D] \underline{K}(\tau) = \left(\sum_{k=0}^d (-1)^k q_{2k} \left(\frac{d}{d\tau} \right)^{2k} \right) \underline{Y}(\tau+\alpha) \quad 0 \leq \tau \leq T \quad (1.52)$$

Multiplying both sides of (1.52) by $\hat{w}(|t-\tau|)$ and integrating the resultant expression from 0 to T, we obtain

$$\begin{aligned} [D] \left[\int_0^T \hat{w}(|t-\tau|) \underline{K}(\tau) d\tau \right] &= \int_0^t \sum_{k=0}^d (-1)^k q_{2k} \underline{Y}^{(2k)}(\tau+\alpha) \hat{w}(t-\tau) d\tau \\ &+ \int_t^T \sum_{k=0}^d (-1)^k q_{2k} \underline{Y}^{(2k)}(\tau+\alpha) \hat{w}(\tau-t) d\tau \end{aligned} \quad (1.53)$$

$$0 \leq t \leq T$$

Here we have separated the range of integration into two regions; for $0 \leq \tau < t$, the kernel is $\hat{w}(t-\tau)$, while for $t \leq \tau \leq T$, the kernel is $\hat{w}(\tau-t)$. After integrating the right-hand side of (1.53) by parts $2d$ times and making use of the property that**

$$\hat{w}^{(j)}(0^+) = (-1)^j \hat{w}^{(j)}(0^-) \quad (1.54)$$

we are left with integrals

$$(-1)^k q_{2k} \int_0^T \underline{Y}(\tau+\alpha) \hat{w}^{(2k)}(|t-\tau|) d\tau$$

* $\hat{w}^{(2k)}(|t|)$ denotes the $2k^{\text{th}}$ derivative of $\hat{w}(|t|)$.

** See Appendix A.

unintegrated as they occur at every other step. In addition, terms involving the derivatives of $y(t+\alpha)$ at $t=0$ and $t=T$ are carried over from each of the $2d$ integrations.

It is shown in Appendix B that the right-hand side of (1.53) can be expressed as

$$\begin{aligned} \sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(t) \cdot \underline{Y}_{\ell} + \sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(T-t) \underline{Z}_{\ell} \\ + \int_0^T \sum_{k=0}^d (-1)^k q_{2k} \hat{w}^{(2k)}(|t-\tau|) \underline{y}(\tau+\alpha) d\tau \end{aligned}$$

where $\underline{Y}_{\ell}$ and $\underline{Z}_{\ell}$ are column matrices whose elements are $Y_{i\ell}$ and $Z_{i\ell}$ ($i = 1, 2, \dots, N$). $Y_{i\ell}$ and $Z_{i\ell}$ are linear combinations of the derivatives of $y_i(t+\alpha)$ ($i = 1, 2, \dots, N$) at $t=0$ and $t=T$ defined by (B.8) and (B.9) respectively. By (1.51), the summation in the last integral of the above expression is $\delta(t-\tau)$. Hence, using the property of the delta function,

$$\int_0^T \delta(t-\tau) \underline{y}(\tau+\alpha) d\tau = \underline{y}(t+\alpha) \quad (1.55)$$

Eq. (1.53) can be reduced to

$$\begin{aligned} [D] \left[\int_0^t \hat{w}(|t-\tau|) \underline{K}(\tau) d\tau \right] &= \underline{y}(t+\alpha) + \sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(t) \underline{Y}_{\ell} \\ &+ \sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(T-t) \underline{Z}_{\ell} \\ &0 \leq t \leq T \end{aligned} \quad (1.56)$$

We observe from (1.56) that in order for the solution obtained from the differential equation (1.46) to satisfy the system of integral equations (1.42), means should be provided to take care of the additional terms

$$\sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(t) \underline{y}_{\ell} + \sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(T-t) \underline{z}_{\ell}$$

Let us add to the solution of the differential equation two terms of the form

$$\underline{b}_1 \delta(t) + \underline{c}_1 \delta(t - T)^*$$

where $\underline{b}_1$ and $\underline{c}_1$ are column matrices consisting of elements b_{j1} , c_{j1} , for $j = 1, 2, \dots, N$. Then it can be easily shown that (1.56) is identically satisfied if

$$\sum_{\ell=1}^{2d} w^{(\ell-1)}(t) \underline{y}_{\ell} = \hat{w}(t) \underline{G} \quad (1.57)$$

and

$$\sum_{\ell=1}^{2d} \hat{w}^{(\ell-1)}(T-t) \underline{z}_{\ell} = \hat{w}(T-t) \underline{H} \quad (1.58)$$

where $\underline{G}$ and $\underline{H}$ are column matrices defined by

$$\underline{G} = [D] \underline{b}_1 \quad (1.59)$$

and

$$\underline{H} = [D] \underline{c}_1 \quad (1.60)$$

It will now be shown that a necessary and sufficient condition that the solution of (1.36) be unique and absolutely integrable is that $\underline{y}(t; \alpha)$ obtained from (1.41) have $N(2d - 2)$ undetermined constants. This means that the determinant $|DF(\lambda^2)|$ is a polynomial of degree $N(2d - 2)$.

* This does not violate Eq. (1.40) since $\int_{-\infty}^{\infty} |\delta(t)| dt = 1$

We note from (B.8) and (B.9) that $\underline{Y}_\ell$ and $\underline{Z}_\ell$ are linear functions of $y(t+\alpha)$ and its successive derivatives evaluated at $t = 0$ and $t = T$ respectively. Let us assume that $y(t+\alpha)$ has $N(2d - 2)$ undetermined coefficients. Then both $\underline{Y}_\ell$ and $\underline{Z}_\ell$ can be expressed as linear combinations of the same number of coefficients. From (1.48), we see that $\hat{W}(t)$ can be expressed as

$$\hat{W}(t) = \sum_{i=1}^d A_i e^{-\alpha_i t} \quad t > 0 \quad (1.61)$$

It follows that

$$\hat{W}^{(l-1)}(t) = \sum_{i=1}^d (\alpha_i)^{l-1} \cdot A_i e^{-\alpha_i t} \quad t > 0 \quad (1.62)$$

Upon equating the corresponding coefficients of $e^{-\alpha_i t}$ ($i = 1, 2, \dots, d$) in Eqs. (1.57) and (1.58), we obtain $2Nd$ algebraic equations. Since the number of unknowns is also $2Nd$ -- $N(2d - 2)$ of which belong to $\underline{Y}_\ell$ and $\underline{Z}_\ell$ and the remaining $2N$ are contributed by $\underline{G}$ and $\underline{H}$ -- a unique solution can always be obtained.

So far, we have only considered the finite memory filter. The result, however, can be applied directly to the infinite memory filter, in which T is infinite. For the infinite memory filter, the stability condition becomes

$$\int_0^\infty |K_j(t)| dt < \infty \quad j = 1, 2, \dots, N \quad (1.63)$$

Eq. (1.63) implies that

A. The roots of the equation

$$|DP(\lambda^2)| = 0 \quad (1.64)$$

cannot be purely real, and

B. The terms in $K(t)$ which belong to the roots of (1.64) that lie in the lower half of the λ plane should be discarded.

When condition (A) is satisfied, a unique solution again can be obtained. In the infinite memory case, it is necessary to use only (1.57). The number of equations as well as the number of unknowns are reduced by a factor of 2.

As a final remark, we see from our result that it is not necessary for $y(t_0)$ to have $N(2d - 2)$ undetermined constants if the stability conditions -- (1.40) for the finite memory filter and (1.63) for the infinite memory filter -- are removed. In this case, we can always obtain a formal solution by adding delta functions and higher order derivatives of delta functions to the solution obtained from the differential equations. The successive derivatives of delta functions are defined by

$$\int_{-\infty}^{\infty} \delta^{(l)}(t - t_0) f(t) dt = (-1)^l f^{(l)}(t_0) \quad (1.65)$$

As an example, consider the solution of (1.36) when all $P_{ij}(\lambda^2) = 1$. It can be shown that (1.56) will be identically satisfied if we add to the solution of $K(t)$ obtained from (1.47) the additional terms

$$\sum_{l=1}^{2d} \delta^{(l-1)}(t) \cdot \underline{b}_l + \sum_{l=1}^{2d} \delta^{(l-1)}(t - T) \underline{c}_l$$

and choose $\underline{b}_l$ and $\underline{c}_l$ so that the following equations are satisfied:

$$[D] \underline{b}_l = \underline{Y}_l \quad l = 1, 2, \dots, 2d \quad (1.66)$$

and

$$[D] \underline{c}_l = \underline{Z}_l \quad l = 1, 2, \dots, 2d \quad (1.67)$$

The physical significance of the higher order impulse functions is that the filters are required to perform differentiation operations.

At least two procedures can be adopted to obtain the solution of (1.36). For the numerical examples in the next section, the following procedures have been adopted.

1. Use (1.41) to obtain the homogeneous and the particular solution of $y(t+\alpha)$.
2. From (1.47), we find $K(t)$.
3. Add to $K(t)$ two terms of the form $b_1 \delta(t) + c_1 \delta(t-T)$.
4. Substitute the solutions obtained in steps (2) and (3) above into (1.36) and perform the integration. The unspecified constants are determined by solving the resultant set of algebraic equations.

It should be noted that $K(t)$ obtained from steps (1) and (2) above is essentially the solution of the set of differential equations

$$[D(-\frac{d^2}{dt^2})]K(t) = Q(-\frac{d^2}{dt^2})z(t+\alpha) \quad (1.68)$$

5. Examples

1. In this example we consider a prediction problem in the presence of noise. Let $m(t+\alpha)$ be the desired output. The message $m(t)$ is a random square wave taking on the values ± 1 with equal probability. The probability density of the time duration between transitions, τ , is $\beta_1 e^{-\beta_1 \tau}$ (a Poisson process). The second order probability density function of the message is given by

$$\begin{aligned} p_m(m_1, m_2) = & \frac{1}{4}(1 - e^{-\beta_1 |\tau_1 - \tau_2|})[(1, -1) + (-1, 1)] \\ & + \frac{1}{4}(1 + e^{-\beta_1 |\tau_1 - \tau_2|})[(1, 1) + (-1, -1)] \end{aligned} \quad (1.69)$$

where (i,j) denotes $\delta(m_1 - i)\delta(m_2 - j)$. In other words it is the point where $m(t - \tau_1)$ takes on value i and $m(t - \tau_2)$ takes on value j . The probability density of the noise, $n(t)$, is of the same form with β_1 replaced by β_2 . We shall choose $\beta_1 = 2$, $\beta_2 = 4$, and $e^{-2\alpha} = \frac{1}{2}$.

By using (1.20) and (1.21), we first obtain $p(x_1, x_2)$ and $p(m_0, x_1)$ which upon substituting into (1.12) yields two simultaneous integral equations. Inspection of the resultant integral equations shows that $K_0(t) = 0$. The integral equation for $K_2(t)$ becomes

$$e^{-2t} = \int_0^T [e^{-2|t-\tau|} + e^{-4|t-\tau|}] K_2(\tau) d\tau \quad 0 \leq t \leq T \quad (1.70)$$

The Fourier transform of the kernel is

$$W(\lambda^2) = \frac{DP(\lambda^2)}{Q(\lambda^2)} = \frac{12(\lambda^2 + 8)}{\lambda^4 + 20\lambda^2 + 64} \quad (1.71)$$

From step 1 of the procedure, we have

$$\left(-\frac{d^2}{dt^2} + 8\right)y(t) = e^{-2t} \quad 0 \leq t \leq T \quad (1.72)$$

Therefore,

$$y(t) = q_{11}e^{-\sqrt{8}t} + q_{12}e^{\sqrt{8}t} + \frac{1}{4}e^{-2t} \quad 0 \leq t \leq T \quad (1.73)$$

Step 2 yields

$$K_2(t) = -\left(\frac{32}{12}\right) [q_{11}e^{-\sqrt{8}t} + q_{12}e^{\sqrt{8}t}] \quad 0 \leq t \leq T \quad (1.74)$$

We now add to $K_2(t)$ the terms $b_{11}\delta(t) + c_{11}\delta(t-T)$ and substitute the complete expression for $K_2(t)$ into (1.70). Assuming $T = 100$ milliseconds, we find that the integral equation can be satisfied

when the unknowns are

$$\begin{aligned} q_{11} &= -183 \times 10^{-3} ; \quad b_{11} = 420 \times 10^{-3} \\ q_{12} &= -3 \times 10^{-3} ; \quad c_{21} = 63 \times 10^{-3} \end{aligned} \quad (1.75)$$

Hence

$$\begin{aligned} K_0(t) &= 0 \\ K_2(t) &= 0.428e^{-\sqrt{8}t} + 0.008e^{\sqrt{8}t} + 0.425(t) + 0.0638(t-0.1) \\ &\quad 0 \leq t \leq 0.1 \quad (1.76) \\ K_2(t) &= 0 \quad \text{for } t > 0.1 \end{aligned}$$

Since $K_0(t) = 0$, the optimum filter takes the form of a linear filter whose unit impulse response is $\frac{1}{2} K_2(t)$ for $0 \leq t \leq 0.1$ and is zero for $t > 0.1$. This is in fact the optimum linear filter since (1.70) is actually the modified Wiener-Hopf equation where the upper limit on the integral has been changed from ∞ to T .

2. Here we consider a pure prediction problem, namely $x(t) = m(t)$. The desired output is again $m(t+\alpha)$. We assume that the second order probability density of the message is

$$\begin{aligned} p(x_1, x_2) &= \\ &\frac{1}{16} [1 - e^{-\beta_1|\tau_2-\tau_1|} - e^{-\beta_2|\tau_2-\tau_1|} - e^{-(\beta_1+\beta_2)|\tau_2-\tau_1|}] [(2,-2)+(1,-1)+(-2,2)+(-1,1)] \\ &+ \frac{1}{16} [1 + e^{-\beta_1|\tau_2-\tau_1|} - e^{-\beta_2|\tau_2-\tau_1|} - e^{-(\beta_1+\beta_2)|\tau_2-\tau_1|}] [(2,-1)+(1,-2)+(-2,1)+(-1,2)] \\ &+ \frac{1}{16} [1 - e^{-\beta_1|\tau_2-\tau_1|} + e^{-\beta_2|\tau_2-\tau_1|} - e^{-(\beta_1+\beta_2)|\tau_2-\tau_1|}] [(2,1)+(1,2)+(-2,-1)+(-1,-2)] \\ &+ \frac{1}{16} [1 + e^{-\beta_1|\tau_2-\tau_1|} + e^{-\beta_2|\tau_2-\tau_1|} + e^{-(\beta_1+\beta_2)|\tau_2-\tau_1|}] [(2,2)+(1,1)+(-2,-2)+(-1,-1)] \end{aligned} \quad (1.77)$$

This situation arises, for instance, when the message is the sum of two independent Poisson processes whose amplitudes take on the values $\pm \frac{1}{2}$ and $\pm \frac{3}{2}$ respectively. The second order probability density function of a Poisson process is of the form given by (1.65).

In this example, let us allow the filter to have an infinite memory. To simplify the calculation, we choose $\beta_1 = 2$, $\beta_2 = 1$. The prediction time α is arbitrary. Let us denote $e^{-\alpha}$ by k ($0 < k \leq 1$). Upon substituting this information into (1.12) we obtain the following two simultaneous integral equations.

$$\begin{aligned}
 -k^2 e^{-2t} + 3k e^{-t} &= \int_0^{\infty} [e^{-2|t-\tau|} + e^{-|t-\tau|}] K_1(\tau) d\tau \\
 &+ \int_0^{\infty} [-e^{-2|t-\tau|} + e^{-|t-\tau|}] K_2(\tau) d\tau \\
 \\
 +k^2 e^{-2t} + 3k e^{-t} &= \int_0^{\infty} [-e^{-2|t-\tau|} + e^{-|t-\tau|}] K_1(\tau) d\tau \\
 &+ \int_0^{\infty} [e^{-2|t-\tau|} + e^{-|t-\tau|}] K_2(\tau) d\tau \quad t \geq 0 \quad (1.78)
 \end{aligned}$$

It is easily verified that the solution is

$$\begin{aligned}
 K_1(\tau) &= \frac{1}{2} (3k - k^2) \delta(\tau) \\
 K_2(\tau) &= \frac{1}{2} (3k + k^2) \delta(\tau) \quad (1.79)
 \end{aligned}$$

It is observed that the only solution for which $K_2(\tau) = 2K_1(\tau)$ [a linear filter] is when $k = 1$, which corresponds to the trivial case of zero prediction time. For any finite prediction time α , a non-linear zero memory filter results.

By using (1.37), the normalized mean square error can be evaluated as

$$\begin{aligned} \left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{Linear}} &= 1 - \frac{2}{2.5} \times \left[\int_0^{\infty} \left[\frac{1}{8} k^2 e^{-2t} + \frac{3}{8} k e^{-t} \right] K_1(t) dt \right. \\ &\quad \left. + \int_0^{\infty} \left[\frac{1}{8} k^2 e^{-2t} + \frac{3}{8} k e^{-t} \right] K_2(t) dt \right] \\ &= 1 - \left\{ 0.9k^2 + 0.1k^4 \right\} \end{aligned} \quad (1.80)$$

For the sake of comparison, the mean square error of the linear filter is also obtained. The autocorrelation function of the message is

$$R_X(\tau) = \frac{36}{16} e^{-2|\tau|} + \frac{4}{16} e^{-|\tau|} \quad (1.81)$$

The Wiener-Hopf equation becomes

$$36k e^{-2t} + 4k^2 e^{-t} = \int_0^{\infty} [36e^{-2|t-\tau|} + 4e^{-|t-\tau|}] h(\tau) d\tau \quad t \geq 0 \quad (1.82)$$

The solution of this equation is found to be

$$h(\tau) = 0.12(k - k^2)e^{-1.86\tau} + (0.139k^2 + 0.861k)\delta(\tau) \quad (1.83)$$

The corresponding normalized mean square error is given by

$$\begin{aligned} \left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{Linear}} &= 1 - \frac{1}{2.5} \int_0^{\infty} h(t) \times \left[\frac{36}{16} k e^{-2t} + \frac{4}{16} k^2 e^{-t} \right] dt \\ &= 1 - \left\{ 0.012k^4 + 0.176k^3 + 0.812k^2 \right\} \end{aligned} \quad (1.84)$$

Inspection of (1.84) and (1.80) shows that (1.84) is always greater than or equal to (1.80); that is

$$\left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{Linear}} - \left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{NL}} = 0.088k^4 - 0.176k^3 + 0.088k^2 \\ = 0.088k^2(1-k)^2 \geq 0 \quad (1.85)$$

The non-linear zero memory filter, therefore, always gives a lower mean square error than that of the linear filter. As a numerical example, let $k = 0.5$ ($\alpha = 0.694$ sec). We find

$$\left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{Linear}} = 0.819; \quad \left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{NL}} = 0.781 \quad (1.86)$$

or, approximately a 5% improvement.

3. For purposes of illustrating the procedures outlined in section 4, we consider here another example of pure prediction. Let the desired output be $m(t+\alpha)$. The amplitude of the input process at any time can take on any one of the four values ± 1 and ± 2 with equal probability. The second order probability density function is

$$P(x_1, x_2) = \frac{1}{16}(1 - e^{-\beta_1|\tau_1 - \tau_2|})[(2, -2) + (1, -1) + (-2, 2) + (-1, 1)] \\ + \frac{1}{16}(1 - e^{-\beta_2|\tau_1 - \tau_2|})[(2, -1) + (1, -2) + (-2, 1) + (-1, 2)] \\ + \frac{1}{16}(1 - e^{-\beta_3|\tau_1 - \tau_2|})[(2, 1) + (1, 2) + (-2, -1) + (-1, -2)] \\ + \frac{1}{16}(1 + e^{-\beta_1|\tau_1 - \tau_2|} + e^{-\beta_2|\tau_1 - \tau_2|} + e^{-\beta_3|\tau_1 - \tau_2|}) \\ \times [(2, 2) + (1, 1) + (-2, -2) + (-1, -1)] \quad (1.87)$$

Here, we let $\beta_1 = \beta_2 = 1$, $\beta_3 = 2$ and we denote $e^{-\alpha}$ by k . The optimizing integral equations are

$$\begin{aligned}
5ke^{-t} - k^2e^{-2t} &= \int_0^T [3e^{-|t-\tau|} + e^{-2|t-\tau|}]K_1(\tau)d\tau \\
&+ \int_0^T [e^{-|t-\tau|} - e^{-2|t-\tau|}]K_2(\tau)d\tau \\
7ke^{-t} + k^2e^{-2t} &= \int_0^T [e^{-|t-\tau|} - e^{-2|t-\tau|}]K_1(\tau)d\tau \\
&+ \int_0^T [3e^{-|t-\tau|} + e^{-2|t-\tau|}]K_2(\tau)d\tau \quad 0 \leq t \leq T \quad (1.88)
\end{aligned}$$

The Fourier transforms of the respective kernels are

$$W_{11}(\lambda^2) = W_{22}(\lambda^2) = \frac{10(\lambda^2 + 2.8)}{\lambda^4 + 5\lambda^2 + 4} \quad (1.89)$$

and

$$W_{12}(\lambda^2) = W_{21}(\lambda^2) = \frac{2(-\lambda^2 + 2)}{\lambda^4 + 5\lambda^2 + 4} \quad (1.90)$$

It follows that

$$[D] = \begin{bmatrix} 10 & 2 \\ 2 & 10 \end{bmatrix} \quad \text{and} \quad [D]^{-1} = \frac{1}{96} \begin{bmatrix} 10 & -2 \\ -2 & 10 \end{bmatrix} \quad (1.91)$$

Following the above mentioned procedures, we find that (1.41) in our example is

$$\begin{bmatrix} 5ke^{-t} - k^2e^{-2t} \\ 7ke^{-t} + k^2e^{-2t} \end{bmatrix} = \begin{bmatrix} 10(-\frac{d^2}{dt^2} + 2.8) & 2(\frac{d^2}{dt^2} + 2) \\ 2(\frac{d^2}{dt^2} + 2) & 10(-\frac{d^2}{dt^2} + 2.8) \end{bmatrix} \begin{bmatrix} \frac{10}{96} & \frac{-2}{96} \\ \frac{-2}{96} & \frac{10}{96} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

(1.92)

which after simplification becomes

$$\begin{bmatrix} 5ke^{-t} - k^2e^{-2t} \\ 7ke^{-t} + k^2e^{-2t} \end{bmatrix} = \frac{1}{24} \begin{bmatrix} -26 \frac{d^2y_1}{dt^2} + 68y_1 + 10 \frac{d^2y_2}{dt^2} - 4y_2 \\ 10 \frac{d^2y_1}{dt^2} - 4y_1 - 26 \frac{d^2y_2}{dt^2} + 68y_2 \end{bmatrix} \quad (1.93)$$

$0 \leq t \leq T$

The solution for $y(t)$ is found to be

$$\begin{aligned} y_1(t) &= Q_{11}e^{\sqrt{2}t} + Q_{12}e^{-\sqrt{2}t} + Q_{13}e^{2t} + Q_{14}e^{-2t} + \frac{7}{3}ke^{-t} \\ y_2(t) &= -Q_{11}e^{\sqrt{2}t} - Q_{12}e^{-\sqrt{2}t} + Q_{13}e^{2t} + (Q_{14} - \frac{2}{3}k^2)e^{-2t} + \frac{11}{3}ke^{-t} \end{aligned} \quad (1.94)$$

Proceeding to the next step, we obtain $K(t)$ from (1.47):

$$\begin{bmatrix} K_1(t) \\ K_2(t) \end{bmatrix} = \begin{bmatrix} \frac{d^4}{dt^4} - 5 \frac{d^2}{dt^2} + 4 \\ \frac{d^4}{dt^4} - 5 \frac{d^2}{dt^2} + 4 \end{bmatrix} \times \begin{bmatrix} \frac{10}{96} & \frac{-2}{96} \\ \frac{-2}{96} & \frac{10}{96} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \quad (1.95)$$

$0 \leq t \leq T$

which after simplification becomes

$$\begin{bmatrix} K_1(t) \\ K_2(t) \end{bmatrix} = \begin{bmatrix} B_{11}e^{2t} + B_{12}e^{-2t} \\ -B_{11}e^{2t} - B_{12}e^{-2t} \end{bmatrix} \quad \begin{matrix} \text{where } B_{11} \text{ and } B_{12} \\ \text{are constants} \end{matrix} \quad (1.96)$$

Now let us add $b_{11}\delta(t) + c_{11}\delta(t-T)$ and $b_{21}\delta(t) + c_{21}\delta(t-T)$ to $K_1(t)$ and $K_2(t)$ respectively and substitute the complete expression into (1.83). We obtain, upon letting $T = 100$ milliseconds,

$$\begin{aligned}
B_{11} &= -2.76 \times 10^{-3} (k - k^2) \\
B_{12} &= -121.31 \times 10^{-3} (k - k^2) \\
b_{11} &= (1293k - 293k^2) \times 10^{-3} \\
b_{21} &= (1707k + 293k^2) \times 10^{-3} \\
c_{11} &= -37.21(k - k^2) \times 10^{-3} \\
c_{21} &= -c_{11}
\end{aligned} \tag{1.97}$$

It is clear that $K_2(t) \neq 2K_1(t)$ (excluding the case where $k = 1$). Consequently, the optimum predictor for this problem is always non-linear. It is interesting to point out that in this particular example, one can also obtain the same solution by solving just a single integral equation. By adding the two equations in (1.88), one can immediately obtain the relation that

$$K_1(t) + K_2(t) = 3000k \times 10^{-3} \delta(t) \tag{1.98}$$

Upon substituting (1.98) into (1.88), the two simultaneous equations can be reduced to a single one.

The normalized mean square error for this non-linear filter is

$$\left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{non-linear}} = 1 - \frac{1}{40} [37k^2 + 2k^3 + k^4] \tag{1.99}$$

Correspondingly, we find the normalized mean square error for the linear filter to be

$$\left\{ \langle \epsilon^2 \rangle_{AV} \right\}_{\text{linear}} = 1 - \frac{1}{40} [35.1k^2 + 3.8k^3 + 0.1k^4] \tag{1.100}$$

We observe that the difference between (1.100) and (1.99) is always greater than or equal to zero; that is

$$\frac{1}{40} [0.9k^2 - 1.8k^3 + 0.9k^4] = \frac{1}{40} \times 0.9k^2 \times (1 - k^2) \geq 0 \quad (1.101)$$

Consequently, we have shown that improved results can be obtained by using a non-linear filter of the $\mathcal{H}_1$ class in place of a linear filter. The amount of improvement depends on the problem at hand. For the particular example we have chosen here as well as the numerical values we have assumed, the improvement is negligible. At $k = 0.5$, the improvement is about 0.15%. Inasmuch as there are five parameters involved in this example ($\beta_1, \beta_2, \beta_3, \alpha$ and T), it is unlikely that any general statement can be made with regard to their effects on the mean square error without first obtaining an explicit expression in terms of these parameters.

II. MULTIPLE PREDICTIONS

The formulation of problems into a set of simultaneous integral equations occurs in many fields; for example, in the study of radiation and wave propagation involving quite general boundary conditions.¹² One problem which is of particular interest to us is the synthesis of linear predictors and filters for multiple time series. Briefly, the problem is as follows:

Given a number of stationary time series,

$$\begin{aligned} r_1(t) &= m_1(t) + n_1(t) \\ r_2(t) &= m_2(t) + n_2(t) \\ &\vdots \\ r_k(t) &= m_k(t) + n_k(t) \end{aligned} \quad (2.1)$$

we wish to find a set of linear filters $h_{1j}(t)$ so that the sum over all inputs j ($j = 1, 2, \dots, k$) of the outputs of the $h_{1j}(t)$ operating on the past of $m_j(t) + n_j(t)$, respectively, is the best approximation in the least square sense to $m_1(t+\alpha)$, ($i = 1, 2, \dots, k$). Specifically, let $i = 1$. Our aim is to find $h_{1j}(t)$, ($j = 1, 2, \dots, k$) so as to minimize

$$\left\langle \left[m_1(t+\alpha) - \sum_{\ell=1}^k \int_0^T r_\ell(t-\tau) h_{1\ell}(\tau) d\tau \right]^2 \right\rangle_{AV} \quad (2.2)$$

A block diagram of such a k -input, 1-output network is shown in Fig. 3, where $\hat{m}_1(t)$ denotes the best estimate of $m_1(t+\alpha)$.

Wiener¹ has shown that the minimum of (2.2) occurs when the $h_{1\ell}(\tau)$ satisfy the following set of integral equations:

$$\begin{aligned} x_j(\tau+\alpha) &= \sum_{\ell=1}^k \int_0^T \phi_{\ell j}(\tau-\sigma) h_{1\ell}(\sigma) d\sigma \\ 0 \leq \tau \leq T; \quad j &= 1, 2, \dots, k \end{aligned} \quad (2.3)$$

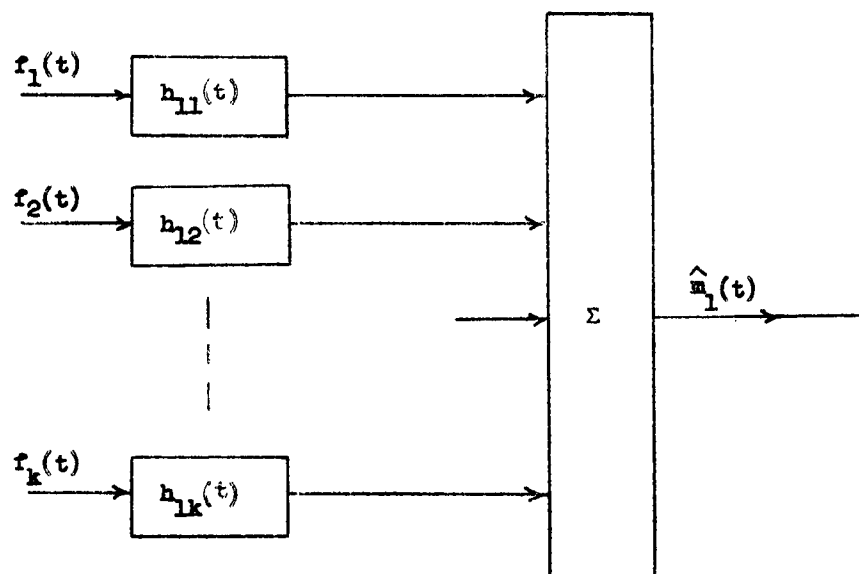


Fig. 3 A K-1 Pole Network

where $X_j(\tau+\alpha)$ is the correlation function between $m_1(t+\alpha)$ and $f_j(t-\tau)$, and $\phi_{lj}(t-\alpha)$ is the correlation function between $f_j(t-\tau)$ and $f_l(t-\sigma)$.

If we suppose that the Fourier transforms of the respective $\phi_{lj}(x)$ are bounded at ∞ and each Fourier transform can be represented by a rational function in $j\lambda$, namely,

$$\Phi_{\ell j}(j\lambda) = \frac{D_{\ell j} P_{\ell j}(j\lambda)}{Q(\lambda^2)} \quad (2.4)$$

then the technique developed in the previous section is also applicable for obtaining the solution of (2.3).

By paralleling the steps in Appendix A, it is seen that the solution of (2.3) satisfies a system of differential equations

$$[DP(\frac{d}{dt})]\underline{h}(t) = Q(-\frac{d^2}{dt^2})\underline{X}(t+\alpha) \quad 0 \leq t \leq T \quad (2.5)$$

where

$\underline{h}(t)$ = a column matrix whose elements are $h_{1\ell}(t)$, ($\ell = 1, 2, \dots, k$)

$\underline{X}(t+\alpha)$ = a column matrix whose elements are $X_j(t+\alpha)$
($j = 1, 2, \dots, k$), and

$[DP(\frac{d}{dt})]$ = a $k \times k$ square matrix whose elements are linear operators $D_{1j}P_{1j}(\frac{d}{dt})$.

The essential difference in the derivation of (2.5) is that here it is unnecessary to break the off-diagonal integrals into two separate integrals when each differentiation is performed. This is because only the kernels on the diagonal axis are even functions of their arguments. In other words

$$\phi_{\ell j}(\tau-\sigma) = \phi_{\ell j}(|\tau-\sigma|) \quad (2.6)$$

only for $\ell = j$.

If a unique, absolutely integrable solution exists, that is, if the impulse response of all the linear filters shown in Fig. 3 satisfies the stability condition

$$\int_0^T |h_{1\ell}(t)| dt < \infty; \quad \ell = 1, 2, \dots, k \quad (2.7)$$

the solution can be obtained by adding delta functions located at $t = 0$ and $t = T$ to the solutions of (2.5). The unspecified constants can be determined by substituting the complete expression for $h(t)$ into (2.3) and solving it as an identity.

REFERENCES

1. Wiener, N., "Extrapolation, Interpolation and Smoothing of Stationary Time Series," John Wiley and Sons, Inc., New York; 1950.
2. Zadeh, L.A., and Ragazzini, J.R., "An Extension of Wiener's Theory of Prediction," J. Appl. Phys., vol. 21, pp. 645-55; July, 1950.
3. Booton, R.C., Jr., "An Optimization Theory for Time-Varying Linear Systems with Non-Stationary Statistical Inputs," Proc. IRE, vol. 40, pp. 977-81; August, 1952.
4. Davis, R.C., "On the Theory of Prediction of Non-stationary Stochastic Processes," J. Appl. Phys., vol. 23, pp. 1047-1053; September, 1952.
5. Blum, M., "An Extension of the Minimum Mean Square Prediction Error Theory for Sampled Data," IRE Trans. on Information Theory, vol. IT-2, pp. 176-84; September, 1956.
6. Chang, S.S.L., "Statistical Design Theory for Strictly Digital Sampled Data Systems," Trans. AIEE I [Communication and Electronics], no. 54, pp. 702-09; January, 1958.
7. Laning, J.H., Jr., "Prediction and Filtering in the Presence of Gaussian Interference," M.I.T. Instrumentation Laboratory, Cambridge, Mass., Tech. Rept. R-27; October, 1951.
8. Drenick, R., "A Nonlinear Prediction Theory," IRE Trans. on Information Theory, vol. IT-4, pp. 146-62; September, 1954.
9. Zadeh, L.A., "Optimum Non-linear Filters," J. Appl. Phys., vol. 24, pp. 396-404; April, 1953.
10. Pugachev, V.S., "A General Condition for Minimum Mean Square Error in a Dynamic System," Automation and Remote Control, vol. 17, pp. 289-295; 1956. (In Russian, translated by Consultant Bureau, Inc., N.Y.)
11. Davenport, W.Jr., and Root, W., "An Introduction to the Theory of Random Signals and Noise," McGraw Hill Book Co., New York; 1958.
12. Heins, A., "System of Wiener-Hopf Equations," Proc. Symp. in Appl. Math. II, McGraw-Hill Book Co., New York, pp. 76-81; 1950.

APPENDIX A

We wish to show that the solution of

$$z_i(t+\alpha) = \sum_{j=1}^N \int_0^T w_{ij}(|t-\tau|) K_j(\tau) d\tau \quad \begin{array}{l} 0 \leq t \leq T \\ i = 1, 2, \dots, N \end{array} \quad A.1$$

with the condition that

$$w_{ij}(\lambda^2) = \int_{-\infty}^{\infty} w_{ij}(|\tau|) e^{-j\lambda\tau} d\tau = \frac{D_{ij} P_{ij}(\lambda^2)}{Q(\lambda^2)} \quad A.2$$

satisfies the following system of differential equations.

$$Q\left(-\frac{d^2}{dt^2}\right) z_i(t+\alpha) = \sum_{j=1}^N D_{ij} P_{ij}\left(-\frac{d^2}{dt^2}\right) K_j(t) \quad A.3$$

$$i = 1, 2, \dots, N \quad 0 \leq t \leq T$$

On account of our assumption that $w_{ij}(|t|)$ is of the form

$$w_{ij}(|t|) = \sum_k B_k e^{-\beta_k |t|} \quad A.4$$

it is readily verified, by taking the limits of the derivatives of any order from both sides of $t = 0$, that

$$w_{ij}^{(k)}(0^+) = (-1)^k w_{ij}^{(k)}(0^-) \quad A.5$$

Equation (A.1) can be written as

$$z_i(t+\alpha) = \sum_{j=1}^N \left\{ \int_0^t w_{ij}(t-\tau) K_j(\tau) d\tau + \int_t^T w_{ij}(\tau-t) K_j(\tau) d\tau \right\} \quad A.6$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N$$

Taking derivatives on both sides with respect to time yields*

* $w_{ij}^{(\beta)}(X)$ refers to the derivatives with respect to the argument X .

$$\begin{aligned}
z_i^{(1)}(t+\alpha) = & \sum_{j=1}^N \left\{ K_j(t) [w_{ij}(0^+) - w_{ij}(0^-)] \right\} \\
& + \sum_{j=1}^N \left\{ \int_0^t w_{ij}^{(1)}(t-\tau) K_j(\tau) d\tau \right. \\
& \left. - \int_t^T w_{ij}^{(1)}(\tau-t) K_j(\tau) d\tau \right\} \quad \begin{matrix} 0 \leq t \leq T \\ i = 1, 2, \dots, N \end{matrix} \quad A.7
\end{aligned}$$

On account of (A.5), the first summation vanishes and we are left with

$$z_i^{(1)}(t+\alpha) = \sum_{j=1}^N \left\{ \int_0^t w_{ij}^{(1)}(t-\tau) K_j(\tau) d\tau - \int_t^T w_{ij}^{(1)}(\tau-t) K_j(\tau) d\tau \right\} \quad A.8$$

$0 \leq t \leq T; \quad i = 1, 2, \dots, N$

Similarly

$$\begin{aligned}
z_i^{(2)}(t+\alpha) = & \sum_{j=1}^N \left\{ K_j(t) [w_{ij}^{(1)}(0^+) + w_{ij}^{(1)}(0^-)] \right\} \\
& + \sum_{j=1}^N \left\{ \int_0^t w_{ij}^{(2)}(t-\tau) K_j(\tau) d\tau + \int_t^T w_{ij}^{(2)}(\tau-t) K_j(\tau) d\tau \right\} \quad A.9
\end{aligned}$$

$0 \leq t \leq T; \quad i = 1, 2, \dots, N$

which in view of (A.5) reduces to

$$z_i^{(2)}(t+\alpha) = \sum_{j=1}^N \left\{ \int_0^t w_{ij}^{(2)}(t-\tau) K_j(\tau) d\tau + \int_t^T w_{ij}^{(2)}(\tau-t) K_j(\tau) d\tau \right\} \quad A.10$$

$0 \leq t \leq T; \quad i = 1, 2, \dots, N$

Following the same reasoning, one can show that

$$\begin{aligned}
z_i^{(k)}(t+\alpha) = & \sum_{j=1}^N \left\{ \int_0^t w_{ij}^{(k)}(t-\tau) K_j(\tau) d\tau \right. \\
& \left. + (-1)^k \int_t^T w_{ij}^{(k)}(\tau-t) K_j(\tau) d\tau \right\} \quad \begin{matrix} 0 \leq t \leq T \\ i = 1, 2, \dots, N \end{matrix} \quad A.11
\end{aligned}$$

Since the even derivative of an even function is again an even function, we can write (A.11) as

$$z_1^{(k)}(t+\alpha) = \sum_{j=1}^N \int_0^T w_{1j}^{(k)}(|t-\tau|) K_j(\tau) d\tau \quad \text{for } k = \text{even} \quad \text{A.12}$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N$$

From the inverse Fourier transform relationship of $w_{1j}(X)$, we obtain

$$w_{1j}^{(k)}(|X|) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\lambda)^k \frac{D_{1j} P_{1j}(\lambda^2)}{Q(\lambda^2)} e^{j\lambda X} d\lambda \quad \text{for } k = \text{even} \quad \text{A.13}$$

Consequently if we operate on both sides of (A.6) by $Q(-\frac{d^2}{dt^2})$, the denominator $Q(\lambda^2)$ vanishes and the resultant equation becomes

$$Q(-\frac{d^2}{dt^2}) z_1(t+\alpha) = \sum_{j=1}^N D_{1j} \int_0^T K_j(\tau) d\tau \times \int_{-\infty}^{\infty} \frac{1}{2\pi} P_{1j}(\lambda^2) e^{j\lambda|t-\tau|} d\lambda$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N \quad \text{A.14}$$

The numerator polynomial $P_{1j}(\lambda^2)$ can be built up by differentiating the integral. Then (A.14) can be written

$$Q(-\frac{d^2}{dt^2}) z_1(t+\alpha) = \sum_{j=1}^N D_{1j} P_{1j}(-\frac{d^2}{dt^2}) \int_0^T K_j(\tau) d\tau \times \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\lambda|t-\tau|} d\lambda$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N$$

or

$$Q(-\frac{d^2}{dt^2}) z_1(t+\alpha) = \sum_{j=1}^N D_{1j} P_{1j}(-\frac{d^2}{dt^2}) K_j(t) \quad \text{A.15}$$

$$0 \leq t \leq T; \quad i = 1, 2, \dots, N$$

where use has been made of the relation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\lambda|t-\tau|} d\lambda = \delta(|t-\tau|)$$

APPENDIX B

In this appendix we carry out the first few integrations of (1.53) and show how it leads to (1.56). Let us rewrite (1.53),

$$\begin{aligned} \sum_{j=1}^N \int_0^T \hat{w}(|t-\tau|) K_j(\tau) d\tau &= \int_0^t \sum_{k=0}^d (-1)^k q_{2k} y_1^{(2k)}(\tau+\alpha) \hat{w}(t-\tau) d\tau \\ &+ \int_t^T \sum_{k=0}^d (-1)^k q_{2k} y_1^{(2k)}(\tau+\alpha) \hat{w}(\tau-t) d\tau \\ 0 \leq t \leq T; \quad i = 1, 2, \dots, N \end{aligned} \quad \text{B.1}$$

Integrating the right hand side of (B.1) once by parts yields

$$\begin{aligned} q_0 \int_0^T y_1(\tau+\alpha) \hat{w}(|t-\tau|) d\tau &+ \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-1)}(\tau+\alpha) \times \hat{w}(t-\tau)] \Big|_0^t \\ &+ \sum_{k=1}^d (-1)^k q_{2k} \times \int_0^t y_1^{(2k-1)}(\tau+\alpha) \times \hat{w}^{(1)}(t-\tau) d\tau \\ &+ \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-1)}(\tau+\alpha) \times \hat{w}(\tau-t)] \Big|_t^T \\ &- \sum_{k=1}^d (-1)^k q_{2k} \times \int_t^T y_1^{(2k-1)}(\tau+\alpha) \hat{w}^{(1)}(\tau-t) d\tau \\ 0 \leq t \leq T; \quad i = 1, 2, \dots, N \end{aligned} \quad \text{B.2}$$

Since

$$\hat{w}^{(k)}(0^+) = (-1)^k \hat{w}^{(k)}(0^-) \quad \text{B.3}$$

the terms evaluated at $\tau = t$ cancel. The remaining terms evaluated at $t = 0$ and $t = T$ are

$$\begin{aligned}
 & - \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-1)}(\alpha)] \times \hat{w}(t) \\
 & + \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-1)}(T+\alpha)] \times \hat{w}(T-t) \\
 & i = 1, 2, \dots, N
 \end{aligned}
 \tag{B.4}$$

Performing the same integration by parts on the remaining integrals gives

$$\begin{aligned}
 & \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-2)}(\tau+\alpha) \times \hat{w}^{(1)}(t-\tau)] \Big|_0^t \\
 & + \sum_{k=1}^d (-1)^k q_{2k} \times \int_0^t y_1^{(2k-2)}(\tau+\alpha) \times \hat{w}^{(2)}(t-\tau) d\tau \\
 & - \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-2)}(\tau+\alpha) \times \hat{w}^{(1)}(\tau-t)] \Big|_t^T \\
 & + \sum_{k=1}^d (-1)^k q_{2k} \times \int_t^T y_1^{(2k-2)}(\tau+\alpha) \times \hat{w}^{(2)}(\tau-t) d\tau \\
 & 0 \leq t \leq T; \quad i = 1, 2, \dots, N
 \end{aligned}
 \tag{B.5}$$

The terms evaluated at $\tau = t$ again cancel. The two other terms evaluated at $t = 0$ and $t = T$ are

$$\begin{aligned}
 & - \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-2)}(\alpha)] \hat{w}^{(1)}(t) \\
 & - \sum_{k=1}^d (-1)^k q_{2k} \times [y_1^{(2k-2)}(T+\alpha)] \hat{w}^{(1)}(T-t) \quad i = 1, 2, \dots, N
 \end{aligned}
 \tag{B.6}$$

The remaining integrals are again integrated by parts and the process is continued $2d$ times, leaving integrals

$$(-1)^k q_{2k} \times \int_0^T y_1(\tau+\alpha) \hat{w}^{(2k)}(|t-\tau|) d\tau; \quad i = 1, 2, \dots, N \quad B.7$$

unintegrated as they occur at every other step. In addition, we have terms in the form of (B.4) and (B.6) left over at each of the $2d$ integrations. Inspection of (B.4) and (B.6) shows that the terms left over can be put in a compact form. Let us define

$$Y_{1l} = - \sum_{k=\frac{l+1}{2}}^d (-1)^k q_{2k} \times \begin{cases} y_1^{(2k-l)}(\alpha) \\ y_1^{(2k-l)}(\tau+\alpha) \end{cases} \quad \begin{matrix} i = 1, 2, \dots, N \\ B.8 \end{matrix}$$

$= \frac{l}{2} \text{ for } l = \text{even}$

and

$$Z_{1l} = \sum_{k=\frac{l+1}{2}}^d (-1)^{l+k+1} q_{2k} \times \begin{cases} y_1^{(2k-l)}(\tau+\alpha) \\ y_1^{(2k-l)}(\tau+\alpha) \end{cases} \quad \begin{matrix} B.9 \\ i = 1, 2, \dots, N. \end{matrix}$$

$= \frac{l}{2} \text{ for } l = \text{even}$

Then the total number of terms which are evaluated at $t = 0$ and $t = T$ can be written as

$$\sum_{l=1}^{2d} Y_{1l} \hat{w}^{(l-1)}(t) + \sum_{l=1}^{2d} Z_{1l} \hat{w}^{(l-1)}(T-t); \quad i = 1, 2, \dots, N \quad B.10$$

It is noted that Y_{1l} and Z_{1l} are linear combinations of the derivatives of $y_1(t+\alpha)$ at $t = 0$ and $t = T$. By making use of (B.7)

and (B.10), the right hand side of (B.1) becomes

$$\begin{aligned} & \sum_{\ell=1}^{2d} Y_{i\ell} \hat{w}^{(\ell-1)}(t) + \sum_{\ell=1}^{2d} Z_{i\ell} \hat{w}^{(\ell-1)}(T-t) \\ & + \int_0^T y_1(\tau+\alpha) \times \sum_{k=0}^d (-1)^k q_{2k} \hat{w}^{(2k)}(|t-\tau|) d\tau \end{aligned} \quad \text{B.11}$$

which can be identified as the right hand side of (1.56).

APPENDIX C

We wish to show that if $\underline{K}(t)$ is the solution of (1.36), it is also a solution of (1.42) where $\underline{y}(t+\alpha)$ is any solution satisfying the set of differential equations (1.41). Let the solution of (1.36) be $\underline{K}(t)$. Then

$$\begin{aligned} & [DP(-\frac{d^2}{dt^2})][\int_0^T \hat{w}(t-\tau)\underline{K}(\tau)d\tau] \\ & = \underline{z}(t+\alpha) = [DP(-\frac{d^2}{dt^2})][D]^{-1}\underline{y}_1(t+\alpha) \end{aligned} \quad \text{C.1}$$

$0 \leq t \leq T$

where $\underline{y}_1(t+\alpha)$ could be any solution of (1.41). Let us introduce $\underline{G}(t)$ where

$$\begin{aligned} \underline{G}(t) &= [\int_0^T \hat{w}(t-\tau)\underline{K}(\tau)d\tau] - [D]^{-1}\underline{y}_1(t+\alpha) \\ & \quad 0 \leq t \leq T \end{aligned} \quad \text{C.2}$$

It is readily seen from (C.1) that

$$[DP(-\frac{d^2}{dt^2})]\underline{G}(t) = 0 \quad \text{C.3}$$

If we let

$$\hat{\underline{y}}(t+\alpha) = \underline{y}_1(t+\alpha) + [D]\underline{G}(t) \quad \text{C.4}$$

then we see that $\hat{\underline{y}}(t+\alpha)$ is a solution of (1.44) if we operate on both sides of (C.4) by $[DP(-\frac{d^2}{dt^2})][D]^{-1}$. Inspection of (1.43) and (C.1) shows that $\underline{K}(t)$ is the solution of both equations when $\hat{\underline{y}}(t+\alpha) = \underline{y}(t+\alpha)$, which establishes the desired result.

DISTRIBUTION LIST

AF Missile Test Cen
Patrick AFB, Fla
Attn: AFMTC Tech Library, MU-135

ADL
Maxwell AFB, Ala

WADD (WCLJA-2)
Aeronautical Research Lab
Research Division
Wright-Patterson AFB, Ohio

RIHC
Griffies AFB, N.Y.

U.S. Army Signal Engineering Labs
Technical Documents Center
Evans Signal Lab
Belmar, New Jersey

Department of the Army
Office of Chief Signal Officer
Washington 25, D.C.
Attn: Research & Dev. Div., OCSigO

ACTIA
Arlington Hall Station
Arlington 12, Virginia
E

DET #2 AFRL (CRPS-2)
L.G. Hanscom Field,
Bedford, Massachusetts
10

Director, Avionics Division (AV)
Bureau of Aeronautics
Department of the Navy
Washington 25, D.C.
2

Director (Code 2027)
U.S. Naval Research Lab
Washington 25, D.C.
2

Director, USAF Project RAND
via: AF Liaison Office
Rand Corporation
1700 Main Street
Santa Monica, California

U.S. Advisory Group
Attn: Mr. E.B. Staples
SHAPE Air Defense Technical Center
P.O.B. 174
The Hague, Netherlands

Boston Sub Office
Patent Prosecution Branch (Hq. AMC)
Murphy General Hospital, Building 133
424 Trapelo Road
Waltham 54, Mass.

Office of Technical Services
Department of Commerce
Washington 25, D.C.
Attn: Tech. Reports Section
2

Library
Boulder Labs
National Bureau of Standards
Boulder, Colorado
2

MIT
Research Lab of Electronics
Room 20B-221
Cambridge 39, Mass.
Attn: John H. Hewitt

ANDC (RDRR)
Andrews AFB,
Washington 25, D.C.

AFCOR (Harold Wooster)
Director, Research Information Office
Washington 25, D.C.

WADC (WCLJR)
Attn: R.M. Lemmen
Wright-Patterson AFB, Ohio

RADC (RCWID)
Attn: Capt. B.J. Long
Griffies AFB, N.Y.

Atomic Energy Commission (J.R. Pasta)
Division of Research
Washington 25, D.C.

Roger Moulton
National Security Agency
Ft. George G. Meade, Maryland

169-LSFO (OSIS, Step-2)
(Space Track)
121 Tropic Road
Waltham 54, Mass.

ASISOL (Robert E. Acker)
Chief, Data Techniques Communication Dept.
Fort Monmouth, N.J.

National Science Foundation
Engineering Sciences Program
1051 Constitution Avenue
Washington 25, D.C.

Office of Naval Research
Code 477, Information Systems Branch
Attn: Marshall C. Levits
Washington 25, D.C.