1961

A general class of discrete time-invariant filters

William F. Trench, *Trinity University*

J. Soc. Indust. Appl. Math.
Vol. 9, No. 3, September, 1961
Printed in U.S.A.

A GENERAL CLASS OF DISCRETE TIME-INVARIANT FILTERS*

WILLIAM F. TRENCH†

1. Introduction. In this paper we consider the problem of smoothing and prediction for sampled functions consisting of a polynomial plus stationary noise. Specifically, we consider real-valued sampled signals of the form

$$y_j = f_j + N_j \qquad (0 \leqq j < \infty),$$

where f_j is the value of a polynomial $f(t)$ at $t = jh$, and $\{N_j\}$ is a stationary (wide-sense) noise sequence of zero mean and known autocorrelation function, ϕ_j. In all that follows, $f(t)$ is assumed to be a fixed, but arbitrary polynomial of degree not exceeding P. We will investigate a certain class of time invariant filters for which the output $\{x_j\}$ is obtained from the input by means of the convolution

$$x_j = \sum_{r=0}^{\infty} w_r y_{j-r} .$$

A problem which has received much attention in the recent literature is that of choosing the weighting sequence $\{w_r\}$ so that

$$x_j = f^{(\nu)}(jh + \tau), \tag{1.1}$$

for some fixed integer ν and delay τ, provided N_j is identically zero. In addition, other conditions are placed on the weighting sequence to control the output standard deviation in the presence of noise. For instance, Blum [1], Johnson [5], and Lees [6] have considered the finite memory case, where $w_r = 0$ for $r \geqq N + 1$, N being an integer not less than P. In this case, the variance of the output noise is given by the quadratic form

$$\sigma^2 = \sum_{r,s=0}^{N} \phi_{r-s}\, w_r\, w_s ,$$

(provided $\phi_0 = 1$), and methods are given for choosing the vector $(w_0, w_1, \cdots, w_N)$ from the class of vectors such that (1.1) is satisfied, so that σ^2 is minimized. In addition, Blum [2] has considered the infinite memory case, where it is required that the weighting sequence $\{w_r\}$ be a solution of a linear, homogeneous difference equation with constant coefficients, the arbitrary constants being chosen to minimize the output variance subject to the constraints arising from (1.1). Filters of this type can

* Received by the editors June 6, 1960 and in revised form December 19, 1960.

The work reported in this paper was performed under Contract No. DA-36-ORD-2683 RD with the Army Rocket and Guided Missile Agency.

† Radio Corporation of America, Moorestown, New Jersey.

be implemented conveniently by means of recursion formulas, provided the difference equation satisfied by $\{w_r\}$ is stable.

In these approaches, the filter design is performed in the time domain. It is the purpose of this paper to show that these methods are special cases of a more general technique in which it is convenient to consider the filter as being defined by its transfer function. From this point of view, it becomes possible to impose additional conditions on the behavior of the filter in the frequency domain.

2. Preliminary Considerations. Let the polynomial $A(z)$ be defined by

$$A(z) = (1 - \gamma_1 z)(1 - \gamma_2 z) \cdots (1 - \gamma_k z) = \sum_{m=0}^{k} a_m z^m, \tag{2.1}$$

where the γ's are real or occur in complex conjugate pairs, and

$$\delta = \max_{1 \leqq i \leqq k} |\gamma_i| < 1. \tag{2.2}$$

Let $K_0, K_1, \cdots, K_N$ be arbitrary real numbers, and define

$$K(z) = \sum_{n=0}^{N} K_n z^n.$$

Then the operation which transforms the sequence $\{y_j\}$ into the sequence $\{x_j\}$ according to

$$\sum_{m=0}^{k} a_m x_{j-m} = \sum_{r=0}^{N} K_r y_{j-r} \qquad (j \geqq \max[k, N]), \tag{2.3}$$

with arbitrary initial conditions, defines a linear, time invariant filter. If $\gamma_1 = \gamma_2 \cdots = \gamma_k = 0$, then (2.3) is a finite memory filter, while if this is not true, it is an exponential filter similar to those studied by Blum [2]. In either case, if we define the generating function, or z-transform of a sequence $\{s_j\}$ to be the formal power series

$$S(z) = \sum_{j=0}^{\infty} s_j z^j,$$

then (2.3) can be solved in the z domain to obtain, by means of the convolution theorem [4],

$$X(z) = \frac{K(z)}{A(z)} Y(z) + \frac{I(z)}{A(z)}, \tag{2.4}$$

where $I(z)$ is a polynomial which depends on the initial conditions.

DEFINITION. *Two formal power series, $P(z)$ and $Q(z)$, are said to satisfy the relation*

$$P(z) \sim Q(z), \tag{2.5}$$

provided the difference $P(z) - Q(z)$ is analytic for $|z| < 1/\delta$, where δ is defined by (2.2).

An elementary consequence of this definition is

LEMMA 1. *If p_n and q_n are the nth coefficients of two power series satisfying*

(2.5), *then*

$$\lim_{n\to\infty} \frac{p_n - q_n}{\Delta^n} = 0 \tag{2.6}$$

for any $\Delta > \delta$. Thus, since $\delta < 1$, we can say that $p_n - q_n$ tends to zero exponentially. We will write (2.6) more compactly as

$$p_n = q_n + o(\delta^n).$$

(Often in the following presentation, it will be convenient to write the relation $P(z) \sim Q(z)$ even in cases where $P(z) - Q(z)$ is obviously a polynomial, so that $p_n = q_n$ for n sufficiently large).

The function

$$T(z) = \frac{K(z)}{A(z)} = \sum_{m=0}^{\infty} w_m z^m \tag{2.7}$$

is regular in $|z| < 1/\delta$, and we can write (2.4) as

$$X(z) \sim T(z)Y(z), \tag{2.8}$$

or

$$x_n = \sum_{r=0}^{n} w_r y_{n-r} + o(\delta^n), \tag{2.9}$$

which exhibits the fact that, for large n, the solution is essentially independent of the initial conditions.

$T(z)$ will be called the transfer function of the filter (2.3).

LEMMA 2. *Let* $\{N_j\}$, $(j \geqq 0)$, *be a real stationary random process with zero mean and autocorrelation function*

$$\phi_r = E[N_j N_{j+|r|}] \qquad (\phi_0 = 1),$$

and define

$$\Phi(z) = \sum_{r=-\infty}^{\infty} \phi_r z^r. \tag{2.10}$$

Let $\{M_j\}$ *be the output of the filter* (2.3) *when the input is* $\{N_j\}$. *Then*

$$E(M_j^2) = \sigma^2 + o(\delta^j),$$

where

$$\sigma^2 = \sum_{r,s=0}^{\infty} w_r w_s \phi_{r-s} = \frac{1}{2\pi i} \oint_{|z|=1} T(z)T(1/z)\Phi(z) \frac{dz}{z} < \infty. \tag{2.11}$$

The proof of this lemma can be obtained from (2.9) and the fact that $T(z)$ is regular in $|z| < 1/\delta$. For a related result, see Brown [4].

Thus, σ^2 is a "steady state variance." In the rest of this paper, it will simply be called the output variance.

The following lemma can be verified by straightforward manipulation, using (2.9).

LEMMA 3. *Let the input to* (2.3) *be of the form*

$$y_\nu = \sum_{\mu=0}^{M} \beta_\mu e^{i\nu\theta_\mu},$$

where θ_μ *is real* ($\mu = 0, 1, \cdots, M$). *Then the output is given by*

$$x_\nu = \sum_{\mu=0}^{M} \beta_\mu T(e^{-i\theta_\mu}) e^{i\nu\theta_\mu} + o(\delta^\nu).$$

Hence, if $T(z) = 0$ *for* $z = e^{-i\theta_\mu}(\mu = 0, 1, \cdots, M)$, x_ν *tends exponentially to zero as* ν *approaches infinity.*

3. The statement of the problems. Let $g(t)$ be obtained by applying an arbitrary constant-coefficient differential or difference operator to the polynomial $f(t)$, and $g(jh) = g_j$. Then, if deg $f(x) \leqq P$, we can find constants $c_0, c_1, \cdots, c_P$ such that

$$g_j = \sum_{r=0}^{P} c_r f_{j-r} \qquad (j \geqq P), \tag{3.1}$$

or, in terms of z-transforms

$$G(z) \sim C(z)F(z), \tag{3.2}$$

where $G(z) = \sum_{n=0}^{\infty} g_n z^n$, $F(z) = \sum_{n=0}^{\infty} f_n z^n$, and $C(z) = \sum_{n=0}^{P} c_n z^n$.

Problem 1. Given the polynomial $A(z)$, (2.1), characterize the class of polynomials $K(z)$, such that if the input to (2.3) is $y_j = F_j$, then the output is given by

$$x_j = g_j + o(\delta^j). \tag{3.3}$$

Problem 2. Given $R(z)$, a known real polynomial such that $R(1) \neq 0$, characterize the class of polynomials $S(z)$ such that, if $K(z) = R(z)S(z)$, then $K(z)$ still meets the requirements (3.3), when $y_j = f_j$.

Problem 3. From the class of polynomials of a given degree which satisfy the requirements of Problem 1 or 2, find the unique polynomial for which σ^2 is minimized.

Thus, Problem 1 is a requirement that (2.3) provide an asymptotically unbiased estimate of $\{g_j\}$, while Problem 3 requires that this estimate be the best of its kind in the least squares sense.

The motivation for Problem 2 requires more explanation. First, in cases where the solution of Problem 3 is not feasible from a practical standpoint, it may still be possible to decrease σ^2 by choosing $R(z)$ suitably. Furthermore, in a situation where the detector itself introduces an extraneous periodic component to the signal, one may exploit Lemma 3 to remove it.

4. The solutions of problems 1 and 2. In several of the references cited, the constraints which arise from the requirements of Problem 1 are treated in the time domain, giving rise to a set of $P + 1$ relationships of the form

$$\sum_{r=0}^{N} K_r r^s = u_s \qquad (s = 0, 1, \cdots, P). \tag{4.1}$$

If $N < P$, (4.1) generally has no solution. If $N = P$, the solution is unique, while if $N > P$ there are $N + 1 - P$ linearly independent solutions. The procedure developed below has the advantage of exhibiting the relationship between the form of $K(z)$ and the constraints in a more explicit fashion.

Equation (3.3) requires that

$$G(z) \sim T(z)F(z),$$

and from (2.7), (2.8) and (3.2) it follows that

$$[K(z) - A(z)C(z)] \frac{F(z)}{A(z)} \sim 0. \tag{4.2}$$

Now, this relationship must hold whenever $F(z)$ is the generating function of a polynomial sequence of maximal degree P. But

$$F(z) = \frac{1}{(z-1)^{P+1}}$$

is such a generating function. This, coupled with (4.2) leads to the conclusion that the function

$$\frac{K(z) - A(z)C(z)}{(z-1)^{P+1}}$$

is regular in $|z| < 1/\delta(> 1)$. Hence $K(z)$ is of the form

$$K(z) = A(z)C(z) + (z-1)^{P+1}D(z), \tag{4.3}$$

where $D(z)$ is a polynomial. Conversely, and $K(z)$ of the form (4.3) is a solution of Problem 1, because then

$$T(z)F(z) = C(z)F(z) + \frac{D(z)(z-1)^{P+1}F(z)}{A(z)} \sim C(z)F(z) \sim G(z),$$

by virtue of the fact that $(z-1)^{P+1}F(z)$ is a polynomial, since

$$\sum_{\nu=0}^{P+1} (-1)^{\nu} \binom{P+1}{\nu} f_{j-\nu} = 0 \qquad (j \geqq P+1).$$

This completes the solution of Problem 1.

For Problem 2, it follows from (4.3) that if $K(z) = R(z)S(z)$, then

$$\frac{d^{\nu}}{dz^{\nu}} [R(z)S(z) - A(z)C(z)]|_{z=1} = 0 \qquad (\nu = 0, 1, \cdots, P).$$

This may be considered as a $(P+1)$-square system in the unknowns $S(1), S^{(1)}(1), \cdots, S^{(P)}(1)$, which, if $R(1) \neq 0$, has the unique solution

$$S^{(\nu)}(1) = b_{\nu} \qquad (\nu = 0, 1, \cdots, P),$$

where $b_0, b_1, \cdots, b_P$ are defined by

$$\frac{A(z)C(z)}{R(z)} = \sum_{n=0}^{\infty} \frac{b_n}{n!} (z-1)^n. \tag{4.4}$$

Since no conditions are placed on $S^{(\nu)}(1)$ for $\nu > P$, we conclude that a solution of Problem 2 is of the form

$$K(z) = R(z) \left[\sum_{n=0}^{P} \frac{b_n}{n!} (z-1)^n + (z-1)^{P+1} E(z) \right],$$

where $E(z)$ is an arbitrary real polynomial. Conversely, any $K(z)$ of this form is a solution of Problem 2.

5. The minimization of σ^2. Let $K(z)$ be a solution of Problem 1 or 2. That is

$$K(z) = R(z)[B(z) + (z-1)^{P+1}E(z)], \tag{5.1}$$

where

$$B(z) = \sum_{n=0}^{P} \frac{b_n}{n!} (z-1)^n,$$

and $b_0, b_1, \cdots, b_P$ are defined by (4.4). Assume that $\deg R(z) = M \geqq 0$. The solution of Problem 3 consists of finding the polynomial $K(z)$ of degree N, which is of the form (5.1), and minimizes σ^2. Clearly, no solution exists unless $N \geqq M + P$, and if equality holds, the solution $E(z) \equiv 0$ is the only one. If $N - (M + P + 1) = V \geqq 0$, then $\deg E(z) = V$, and the choice of $K(z)$ has $V + 1$ degrees of freedom $E_0, E_1, \cdots, E_V$, where

$$E(z) = \textstyle\sum_{\nu=0}^{V} E_\nu z^\nu.$$

Substituting (5.1) into the contour integral (2.11), and defining

$$H(z) = \Phi(z) \frac{R(z)R(1/z)}{A(z)A(1/z)},$$

one obtains

$$\begin{aligned} \sigma^2 &= \frac{1}{2\pi i} \int_{|z|=1} H(z)B(z)B(1/z) \frac{dz}{z} \\ &\quad + \frac{2}{2\pi i} \int_{|z|=1} H(z)B(z) \left(\frac{1}{z} - 1\right)^{P+1} E(1/z) \frac{dz}{z} \\ &\quad + \frac{1}{2\pi i} \int_{|z|=1} H(z)(z-1)^{P+1} \left(\frac{1}{z} - 1\right)^{P+1} E(z)E(1/z) \frac{dz}{z}, \end{aligned} \tag{5.2}$$

where we have made use of the fact that $H(z) = H(1/z)$ in combining two

integrals to obtain the second term. The condition for a minimum is that

$$\frac{\partial \sigma^2}{\partial E_\nu} = 0 \qquad (\nu = 0, 1, \cdots, V),$$

which leads to

$$(5.3)\quad \frac{1}{2\pi i}\int_{|z|=1} H(z)\left(\frac{1}{z} - 1\right)^{P+1}[B(z) + (z-1)^{P+1}E(z)]\,\frac{dz}{z^{s+1}} = 0 \qquad s = 0, 1, \cdots, V).$$

There are two methods which suggest themselves for solving (5.3) for the optimum transfer function. The first leads to a result which is similar in form to those obtained in some of the previous works cited above. One defines

$$(5.4)\qquad L(z) = B(z) + (z-1)^{P+1}E(z) = \textstyle\sum_{r=0}^{N-M} L_r z^r,$$

and expands (5.3) to obtain

$$(5.5)\quad \sum_{\nu=0}^{P+1}(-1)^\nu\binom{P+1}{\nu}\frac{1}{2\pi i}\oint_{|z|=1} H(z)L(z)\,\frac{dz}{z^{\nu+s+1}} = 0 \qquad (s = 0, 1, \cdots, V).$$

This implies that

$$(5.6)\quad \frac{1}{2\pi i}\oint_{|z|=1} H(z)L(z)\,\frac{dz}{z^{r+1}} = \lambda_0 + \lambda_1 r + \cdots + \lambda_P r^P \qquad (r = 0, 1, \cdots, N-M),$$

since (5.5) is the homogeneous difference equation satisfied by polynomials of degree at most P. The quantities $(\lambda_0, \cdots, \lambda_P)$ play a role analogous to that of the Lagrange multipliers which are introduced in the conventional approach.

Expanding (5.6), we obtain

$$(5.7)\quad \textstyle\sum_{s=0}^{N-M} \alpha_{r-s}L_s = \lambda_0 + \lambda_1 r + \cdots + \lambda_P r^P \quad (r = 0, 1, \cdots, N-M),$$

where

$$\alpha_n = \frac{1}{2\pi i}\oint_{|z|=1} H(z)\,\frac{dz}{z^{n+1}}.$$

By using the transformation $z = 1/\zeta$, it is easy to show that $\alpha_n = \alpha_{-n}$.

It can be shown that

$$\sigma^2 = \textstyle\sum_{r,s=0}^{N-M} \alpha_{r-s}L_rL_s,$$

which is positive for every nonvanishing choice of $L_0, \cdots, L_{N-M}$. Hence

the matrix

$$\Gamma = (\alpha_{r-s}) \quad (r, s = 0, 1, \cdots, N - M),$$

is nonsingular. Defining

$$\boldsymbol{L} = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_{N-M} \end{bmatrix}, \qquad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_P \end{bmatrix},$$

$$\Delta_{N-M} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^P \\ \vdots & & & \\ 1 & (N-M) & \cdots & (N-M)^P \end{bmatrix},$$

the solution of (5.7) becomes

$$\boldsymbol{L} = \Gamma^{-1}\Delta_{N-M}\boldsymbol{\lambda}. \tag{5.8}$$

The unknown vector $\boldsymbol{\lambda}$ can be obtained by noting that (5.4) implies:

$$L^{(\nu)}(1) = B^{(\nu)}(1) = b_\nu \qquad (\nu = 0, 1, \cdots, P).$$

Hence,

$$\textstyle\sum_{n=0}^{N-M} L_n n^\nu = \sum_{m=0}^{P} B_m m^\nu \quad (\nu = 0, 1, \cdots, P), \tag{5.9}$$

(where $B(z) = \sum_{m=0}^{P} B_m z^m$), since for any power series

$$\left(z\frac{d}{dz}\right)^\nu \left(\sum_{m=0}^{\infty} \beta_n z^n\right)\bigg|_{z=1} = \sum_{m=0}^{\infty} \beta_n n^\nu,$$

and the differential operator $\left(z\dfrac{d}{dz}\right)^\nu$ can be expanded in the form

$$\left(z\frac{d}{dz}\right)^\nu = \sum_{m=0}^{\nu} a_{m,\nu} z^m \frac{d^m}{dz^m} \qquad (\nu \geqq 0).$$

Letting

$$\boldsymbol{B} = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_P \end{bmatrix},$$

we can write (5.9) as

$$\Delta_{N-M}^T \boldsymbol{L} = \Delta_P{}^T \boldsymbol{B},$$

where the superscript T denotes matrix transpose. From (5.8) it follows that

$$\Delta_{N-M}^T \Gamma^{-1} \Delta_{N-M} \boldsymbol{\lambda} = \Delta_P{}^T \boldsymbol{B}.$$

Now, it can be shown that the $(P + 1)$-square matrix is positive definite, and has an inverse. Hence

$$\lambda = (\Delta_{N-M}^T \Gamma^{-1} \Delta_{N-M})^{-1} \Delta_P{}^T \boldsymbol{B},$$

which may be substituted in (5.8) to yield

$$\boldsymbol{L} = \Gamma^{-1} \Delta_{N-M} (\Delta_{N-M}^T \Gamma^{-1} \Delta_{N-M})^{-1} \Delta_P{}^T \boldsymbol{B}. \tag{5.10}$$

For the case where $V + 1$, the number of degrees of freedom, is small compared to $N - M$, it is convenient to obtain an alternate solution which does not require the inversion of the high order matrix Γ. We deduce from (5.3) that $E_0, E_1, \cdots, E_V$ is the solution of

$$\textstyle\sum_{r=0}^{V} \beta_{r-s} E_r = -U_s \qquad (s = 0, 1, \cdots, V),$$

where

$$U_s = \frac{1}{2\pi i} \oint_{|z|=1} H(z) \left(\frac{1}{z} - 1\right)^{P+1} B(z) \frac{dz}{z^{s+1}},$$

and

$$\beta_n = \frac{1}{2\pi i} \oint_{|z|=1} H(z) \left(\frac{1}{z} - 1\right)^{P+1} (z - 1)^{P+1} \frac{dz}{z^{n+1}}.$$

6. The behavior of σ^2 as a function of $(\gamma_1, \cdots, \gamma_k)$. In the previous section, we solved Problem 3 under the assumption that the k-tuple $(\gamma_1, \cdots, \gamma_k)$ is fixed. The transfer function of the optimum filter with $\deg K(z) = N$ is given by

$$T(z) = \frac{R(z)L(z)}{A(z)},$$

where the coefficients of $L(z)$ are obtained from (5.10). If $\gamma_i = 0$, $(0 \leqq i \leqq k)$, then $A(z) \equiv 1$, and we have the optimum finite memory filter, of specified length, which meets the requirements of Problem 2. If $\gamma_i \neq 0$ for some i, the filter (2.3) has an infinite memory, the influence of past history being determined by the positions of $(\gamma_1, \cdots, \gamma_k)$ within the unit circle. For suitable choices, it is possible to exploit the infinite memory to advantage in reducing σ^2. However, it is also possible for this quantity to become arbitrarily large for improperly chosen $(\gamma_1, \cdots, \gamma_k)$. For instance, the filter

$$x_j - (\gamma_1 + \gamma_2) x_{j-1} + \gamma_1 \gamma_2 x_{j-2} = (1 - \gamma_1)(1 - \gamma_2) y_j$$

provides an asymptotically unbiased estimate of a constant, with output noise variance

$$\sigma^2 = \frac{1 + \gamma_1 \gamma_2}{1 - \gamma_1 \gamma_2} \frac{(1 - \gamma_1)(1 - \gamma_2)}{(1 + \gamma_1)(1 + \gamma_2)}, \tag{6.1}$$

provided the input noise is uncorrelated. In (6.1), if γ_1 and γ_2 tend to nonreal conjugate points on the unit circle, or if one of them approaches -1, them σ^2 tends to infinity. On the other hand

$$\lim_{\gamma_1, \gamma_2 \to 1} \sigma^2 = 0.$$

In general, the behavior of σ^2 as a function of $(\gamma_1, \gamma_2, \cdots, \gamma_k)$ seems to be quite complicated. However, we conjecture that the most useful choices of the γ's for noise reduction are those for which they are all positive reals in $(0, 1)$. Along these lines, we have the following theorem.

THEOREM 1. *Let $\Phi(z)$, defined by (2.10), be analytic in the annulus $\rho < |z| < 1/\rho$, with $0 \leqq \rho < 1$, and $C(z)$, (3.1 and 3.2), have the property that*

$$C^{(\nu)}(1) = 0 \qquad (\nu < r), \tag{6.2 a}$$

$$C^{(r)}(1) \neq 0, \tag{6.2 b}$$

where $0 \leqq r < P$, while $A(z) = (1 - \gamma z)^k$, with $0 < \gamma < 1$. Then, the output variance, $\sigma^2(\gamma)$, of the optimum filter obtained in § 5 is of the form

$$\sigma^2(\gamma) = O_1(\gamma)(1 - \gamma)^{2r+1} + O_2(\gamma)(1 - \gamma)^{2k+2r-2P} \tag{6.3}$$

in $\rho < \gamma < 1$. The functions $O_1(\gamma)$ and $O_2(\gamma)$ remain bounded as γ increases toward unity.

Proof. We first consider the case where the degree of $K(z)$ is required to be $M + P$, which leads to $E(z) \equiv 0$. Let $\sigma_1^2(\gamma)$ be the output variance for this case. Then

$$\sigma_1^2(\gamma) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{\Phi(z)R(z)R(1/z)}{(1 - \gamma z)^k(1 - \gamma/z)^k} B(z)B(1/z) \frac{dz}{z}. \tag{6.4}$$

Recalling that

$$B(z) = \sum_{n=0}^{P} \frac{b_n(\gamma)}{n!} (z - 1)^n,$$

where

$$b_n(\gamma) = \frac{d^n}{dz^n} \left[\frac{A(z)C(z)}{R(z)} \right] \Bigg|_{z=1},$$

we can write (6.4) as

$$\begin{aligned} \sigma_1^2(\gamma) &= \frac{1}{2\pi i} \sum_{\mu,\nu=0}^{P} \frac{b_\mu(\gamma)b_\nu(\gamma)}{\mu!\nu!} \\ &\quad \cdot \oint_{|z|=1} \frac{\Phi(z)R(z)R(1/z)}{(1 - \gamma z)^k(1 - \gamma/z)^k} (z - 1)^\mu \left(\frac{1}{z} - 1\right)^\nu \frac{dz}{z}. \end{aligned} \tag{6.5}$$

Defining

$$e_n = \frac{1}{n!} \frac{d^n}{dz^n} \left(\frac{C(z)}{R(z)} \right) \Bigg|_{z=1},$$

we have

$$\frac{b_n(\gamma)}{n!} = \sum_{m=0}^{n} (-\gamma)^m \binom{k}{m} e_{n-m}(1-\gamma)^{k-m} \qquad (n = 0, 1, \cdots, P).$$

From (6.2a),

$$\frac{b_n(\gamma)}{n!} = 0 \qquad (n < r),$$

$$\frac{b_n(\gamma)}{n!} = (1-\gamma)^{k+r-n}\beta_n(\gamma) \qquad (r \leqq n \leqq P),$$

where $\beta_n(\gamma)$ is continuous (and possibly vanishes) at $\gamma = 1$. From this, and the fact that $\Phi(z)$ is analytic in $\rho < |z| < 1/\rho$, we can write (6.5) as

$$\begin{aligned} \sigma_1^2(\gamma) = \sum_{\mu,\nu=r}^{P} \frac{\beta_\mu(\gamma)\beta_\nu(\gamma)}{2\pi i} (1-\gamma)^{2k+2r-\mu-\nu} \\ \cdot \oint_{|z|=\rho_1} \frac{\psi_{\mu,\nu}(z)(1-z)^{\mu+\nu}}{(1-\gamma z)^k(z-\gamma)^k}\, dz \\ + \sum_{\mu,\nu=r}^{P} \frac{\beta_\mu(\gamma)\beta_\nu(\gamma)}{2\pi i} (1-\gamma)^{2k+2r-\mu-\nu} \oint_{C_\gamma} \frac{\psi_{\mu,\nu}(z)(1-z)^{\mu+\nu}}{(1-\gamma z)^k(z-\gamma)^k}\, dz, \end{aligned} \tag{6.6}$$

for $\rho < \rho_1 < \gamma < 1$, C_γ being a small circle about $z = \gamma$, contained in $\rho_1 < |z| < 1$, and

$$\psi_{\mu,\nu}(z) = (-1)^\mu z^{k-\nu-1}\Phi(z)R(z)R(1/z).$$

The integrals in the first sum of (6.6) are continuous functions of γ near $\gamma = 1$. Hence, the first sum can be written in the form of the second term of (6.3). For the second sum, since $\psi_{\mu,\nu}(z)$ is regular in C_γ,

$$\oint_{C_\gamma} \frac{\psi_{\mu,\nu}(z)(1-z)^{\mu+\nu}}{(1-\gamma z)^k(z-\gamma)^k}\, dz = \frac{1}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}\left[\frac{\psi_{\mu,\nu}(z)(1-z)^{\mu+\nu}}{(1-\gamma z)^k}\right]\Bigg|_{z=\gamma},$$

which can be shown by simple differentiation to be of the form

$$(1-\gamma)^{\mu+\nu-2k+1}J_{\mu,\nu}(\gamma),$$

where $J_{\mu,\nu}(\gamma)$ is continuous in $\rho < \gamma < 1/\rho$. Substituting this result in the second sum of (6.6), we conclude that it is of the form given by the first term of (6.3), which proves the theorem for $E(z) \equiv 0$.

For the general case, note that (5.3) implies that

$$\frac{1}{2\pi i}\oint_{|z|=1} H(z)\left(\frac{1}{z}-1\right)^{P+1}[B(z) + (z-1)^{P+1}E(z)]E\left(\frac{1}{z}\right)\frac{dz}{z} = 0,$$

and now, from (5.2)

$$\sigma^2(\gamma) = \sigma_1^2(\gamma) - \frac{1}{2\pi i}\oint_{|z|=1} H(z)(z-1)^{P+1}\left(\frac{1}{z}-1\right)^{P+1} E(z)E\left(\frac{1}{z}\right)\frac{dz}{z}.$$

But the integral is positive, and we conclude that

$$0 < \frac{\sigma^2(\gamma)}{(1-\gamma)^q} < \frac{\sigma_1^2(\gamma)}{(1-\gamma)^q} \qquad (0 < \gamma < 1),$$

for any q; in particular, $q = \min[2r + 1, 2k + 2r - 2P]$, which completes the proof of the theorem.

As a necessary condition for the variance to approach zero, we have:

THEOREM 2. *Let $\sigma^2(\gamma_1, \gamma_2, \cdots, \gamma_k)$, defined in the set*

$$R = \{(\gamma_1, \cdots, \gamma_k) \mid |\gamma_i| < 1, \quad a_i \text{ real}, \quad (i = 1, \cdots, k)\},$$

denote the output variance of the optimum filter obtained in §5, for $A(z)$ given by (2.1). Let $(\hat{\gamma}_1, \cdots, \hat{\gamma}_k)$ be a point in the closure of R such that

$$\lim \sigma^2(\gamma_1, \cdots, \gamma_k) = 0$$

as $(\gamma_1, \cdots, \gamma_k)$ approaches $(\hat{\gamma}_1, \cdots, \hat{\gamma}_k)$ over some path in R. Then $\hat{A}(z) = (1 - \hat{\gamma}_1 z) \cdots (1 - \hat{\gamma}_k z)$ is of the form

$$\hat{A}(z) = (1 - z)^{P-r+1}\hat{\hat{A}}(z).$$

Proof. Let $(\gamma_{1n}, \cdots, \gamma_{kn})$, $(n = 0, 1, \cdots)$, be a sequence of vectors in R, tending to $(\hat{\gamma}_1, \cdots, \hat{\gamma}_k)$, such that $\lim_{n\to\infty} \sigma^2(\gamma_{1n}, \cdots, \gamma_{kn}) = 0$. From (2.11) we can write

$$\sigma^2(\gamma_{1n}, \cdots, \gamma_{kn}) = \frac{1}{2\pi}\int_0^{2\pi} \Phi(e^{i\theta}) \mid T(e^{i\theta}; \gamma_{1n}, \cdots, \gamma_{kn}) \mid^2 d\theta,$$

where $T(z; \gamma_{1n}, \cdots, \gamma_{kn})$ is the transfer function of the optimum filter associated with $(\gamma_{1n}, \cdots, \gamma_{kn})$. Since Φ is positive on the unit circle (being essentially a power spectral density), it follows that

$$\lim_{n\to\infty} T(e^{i\theta}; \gamma_{1n}, \cdots, \gamma_{kn}) = 0 \tag{6.7}$$

almost everywhere. Let f_j be an arbitrary polynomial of degree not exceeding P, and g_j be defined by (3.1). For every n, we can write

$$\textstyle\sum_{m=0}^{k} a_m(\gamma_{1n}, \cdots, \gamma_{kn}) g_{j-m} = \sum_{r=0}^{N} K_r(\gamma_{1n}, \cdots, \gamma_{kn}) f_{j-r}, \tag{6.8}$$

for $j \geqq \max[k, N]$. Holding j fixed, but arbitrary, and noting that (6.7) implies that $\lim_{n\to\infty} K_r(\gamma_{1n}, \cdots, \gamma_{kn}) = 0$ for $r = 0, 1, \cdots, N$, we can pass to the limit in (6.8) to obtain

$$\textstyle\sum_{m=0}^{k} a_m(\hat{\gamma}_1, \cdots, \hat{\gamma}_k) g_{j-m} = 0 \qquad (j \geqq \max[k, N]).$$

However, g_j can be an arbitrary polynomial of degree $P - r$, so that the conclusion of Theorem 2 follows from the fact that

$$\sum_{\nu=0}^{P-r+1} (-1)^\nu \binom{P - r + 1}{\nu} g_{j-\nu} = 0$$

is the irreducible equation satisfied by the class of such polynomials.

A word of caution is in order. As $(\gamma_1, \cdots, \gamma_k)$ approaches the boundary of R, the recursion formula (2.3) becomes less desirable from two points of view. First, the transient error may not die out rapidly enough, and secondly, the growth of computational errors becomes more bothersome, despite the fact that the overall computational error remains bounded, provided the totality of individual errors has a finite bound.

For instance, suppose one attempts to design a filter of the form

$$\sum_{\nu=0}^{P+1} (-1)^\nu \binom{P+1}{\nu} g_{j-\nu} = \sum_{r=0}^{N} K_r f_{j-r},$$

to perform a smoothing operation on a polynomial of degree P or less. (For an example of a filter of a type similar to this, see Blum [3]). It is easy to show that round off errors committed at each step can excite an error which grows in order of magnitude like a polynomial of degree $P + 1$. In many cases, this error will completely mask the desired solution.

7. Examples. In this section we develop some examples to illustrate the techniques given above. Unfortunately, even in the simple cases to be discussed, the expressions for the filter coefficients are cumbersome in form, although they can be easily computed for specific values of the parameters involved.

Example 1. As a first example, we derive coefficients $K_0(\gamma), K_1(\gamma), \cdots, K_N(\gamma)$ such that if

$$y_j = f(jh) + N_j, \tag{7.1}$$

where $f(t)$ is a polynomial of degree one and $\{N_j\}$ is uncorrelated zero-mean noise, then the sequence $\{x_j\}$, defined by

$$x_j - 2\gamma x_{j-1} + \gamma^2 x_{j-2} = \textstyle\sum_{r=0}^{N} K_r(\gamma) y_{j-r} \tag{7.2}$$

is an asymptotically unbiased, minimum-variance estimate of $\{g_j\}$, defined in §3. It is to be understood that the variance is minimized for that class of filters of the form (7.2), and that it is a function of γ and N. The parameter γ is assumed to lie in the interval $0 < \gamma \leqq 1$.

Since second and high order differences of a first degree polynomial vanish we can write

$$g_j = C_0 f_j + C_1(f_j - f_{j-1}).$$

Hence, the most general form of $C(z)$ for this example is

$$C(z) = C_0 + C_1(1 - z),$$

while

$$A(z) = (1 - \gamma z)^2,$$

$$\Phi(z) = R(z) = 1.$$

The function $B(z)$ required in (5.1) is

$$B(z) = (1-\gamma)^2 C_0 - [(1-\gamma)^2 C_1 + 2\gamma(1-\gamma)C_0](z-1).$$

The elements of the matrix $\Gamma = (\alpha_{r-s})$ are given by

$$\alpha_{r-s} = \frac{1}{2\pi i}\oint_{|z|=1} \frac{1}{(1-\gamma z)^2(1-\gamma/z)^2}\frac{dz}{z^{r-s+1}} \qquad (r, s = 0, 1, \cdots, N).$$

This matrix can be inverted by inspection, as follows. If the polynomials

$$\beta_r(z) = \sum_{m=0}^{N} \beta_{r,m} z^m \qquad (r = 0, 1, \cdots, N)$$

are defined by

$$\beta_0(z) = (1-\gamma z)^2,$$

$$\beta_1(z) = z(1-2\gamma/z)(1-\gamma z)^2,$$

$$\beta_r(z) = z^r(1-\gamma/z)^2(1-\gamma z)^2 \quad (r = 2, 3, \cdots, N-2),$$

$$\beta_{N-1}(z) = z^{N-1}(1-\gamma/z)^2(1-2\gamma z),$$

$$\beta_N(z) = z^N(1-\gamma/z)^2,$$

it is easy to show that

$$\sum_{m=0}^{N} \beta_{r,m}\alpha_{m-s} = \frac{1}{2\pi i}\oint_{|z|=1} \frac{\beta_r(z)}{(1-\gamma z)^2(1-\gamma/z)^2}\frac{dz}{z^{s+1}} = \delta_{rs}$$

and this, coupled with the fact that $\alpha_{m-s} = \alpha_{s-m}$, implies that

$$(\Gamma^{-1})_{r,s} = \beta_{r,s} \qquad (r, s = 0, 1, \cdots, N).$$

Referring to (5.7), and noting that $M = 0$, and $L_s = K_s$, we obtain, after some manipulation,

$$K_r = \lambda_0\beta_r(1) + \lambda_1\beta_r'(1) \qquad (r = 0, 1, \cdots, N).$$

(The "prime" indicates differentiation with respect to z.) The quantities λ_0 and λ_1 are determined by requiring that $K(1) = B(1)$ and $K'(1) = B'(1)$, where

$$K(z) = \sum_{r=0}^{N} K_r z^r.$$

This leads to a pair of simultaneous linear equations, whose solution is

$$\begin{pmatrix}\lambda_0\\ \lambda_1\end{pmatrix} = \frac{(1-\gamma)}{S_0(\gamma)S_2(\gamma) - S_1^2(\gamma)}\begin{pmatrix} S_2(\gamma) & -S_1(\gamma)\\ -S_1(\gamma) & S_0(\gamma)\end{pmatrix} \cdot \begin{pmatrix}(1-\gamma) & 0\\ -2\gamma & -(1-\gamma)\end{pmatrix}\begin{pmatrix}C_0\\ C_1\end{pmatrix},$$

where

$$S_0(\gamma) = \sum_{r=0}^{N} \beta_r(1) = (N+1)(1-\gamma)^4 + 4\gamma(1-\gamma)^3,$$

$$S_1(\gamma) = \sum_{r=0}^{N} r\beta_r(1) = \sum_{r=0}^{N} \beta_r'(1) = \frac{N(N+1)}{2}(1-\gamma)^4 + 2N\gamma(1-\gamma)^3,$$

$$S_2(\gamma) = \sum_{r=0}^{N} r\beta_r'(1)$$

$$= \frac{N(N+1)(2N+1)}{6}(1-\gamma)^4$$

$$+ 2\gamma(1-\gamma)[N^2(1-\gamma)^2 + N(1-\gamma^2) + \gamma(1+\gamma)].$$

The expression for $\sigma^2(\gamma)$ is quite complicated. However, Theorem 1 guarantees that $\sigma^2(\gamma)/(1-\gamma)$ remains bounded as γ approaches unity, while if $C_0 = 0$, $\sigma^2(\gamma)/(1-\gamma)^2$ remains bounded. On the other hand, setting $\gamma = 0$ yields a finite-memory minimum-variance filter of the type studied in [3] and [5].

Example 2. For this example, we derive a filter of the form

$$x_j = \textstyle\sum_{r=0}^{N} K_r y_{j-r}$$

which provided the unbiased, minimum-variance estimate of $\{g_j\}$ when $\{y_j\}$ has the form (7.1), with $\{f_j\}$ being sampled values of a first degree polynomial and

$$E(N_j) = 0, \qquad E(N_j \cdot N_{j+m}) = \rho^{|m|}$$

where $0 < \rho < 1$. This type of noise spectrum would arise from sampling a Gaussian-Markoffian process with zero mean and unit average power.

We have

$$A(z) = R(z) = 1$$

$$\Phi(z) = \frac{1-\rho^2}{(1-\rho z)(1-\rho/z)}$$

$$B(z) = C(z).$$

In a manner similar to that of Example 1, we find that

$$(\Gamma^{-1})_{r,s} = \beta_{r,s},$$

where the polynomials $\{\beta_r(z)\}$ are given by

$$\beta_0(z) = \frac{1-\rho z}{1-\rho^2}$$

$$\beta_r(z) = z^r \frac{(1-\rho z)(1-\rho/z)}{1-\rho^2} \qquad (r = 1, 2, \cdots, N-1),$$

$$\beta_N(z) = \frac{z^N(1-\rho/z)}{1-\rho^2}.$$

The required coefficients are

$$K_r = \lambda_0\beta_r(1) + \lambda_1\beta'(1),$$

where

$$\begin{pmatrix}\lambda_0\\ \lambda_1\end{pmatrix} = \frac{1}{S_0(\rho)S_2(\rho) - S_1^2(\rho)}\begin{pmatrix} S_2(\rho) & -S_1(\rho)\\ -S_1(\rho) & S_0(\rho)\end{pmatrix}\begin{pmatrix}C_0\\ C_1\end{pmatrix}$$

and

$$S_0(\rho) = (N+1)\frac{1-\rho}{1+\rho} + \frac{2\rho}{1+\rho},$$

$$S_1(\rho) = \frac{N(N+1)}{2}\frac{1-\rho}{1+\rho} + \frac{N\rho}{1+\rho},$$

$$S_2(\rho) = \frac{N(N+1)(2N+1)}{6}\frac{1-\rho}{1+\rho} + \frac{N^2\rho}{1+\rho} + \frac{N\rho}{1-\rho^2}.$$

Example 3. This example illustrates the synthesis of a digital filter which annihilates an undesirable sinusoidal oscillation. Specifically, suppose

$$y(t) = f(t) + B\sin(\omega t + \phi),$$

where $f(t)$ is a second degree polynomial, ω is a known, nonzero frequency, and B and ϕ are arbitrary, but fixed numbers. Let $y_n = y(nh)$. We will derive a filter of the form

$$x_n = \textstyle\sum_{r=0}^{4} K_r y_{n-r}$$

such that

$$x_n = f_{n-s}. \tag{7.3}$$

Thus, the filter will pass the polynomial component with a lag, and destroy the unwanted sinusoid.

The second requirement will be satisfied if $K(z)$ is of the form

$$K(z) = (1 - 2z\cos\omega h + z^2)L(z)$$

from Lemma 3. The operation denoted by (7.3) results in the transfer function $C(z) = z^2$.

For this case, we see from the arguments in §4 that $L(z) = B(z)$, and $L(z)$ is equal to the first three terms of the Taylor expansion

$$\frac{z^2}{1 - 2z\cos\omega h + z^2} = \frac{1}{4\sin^2(\omega h/2)}\left[1 + (z-1) - \frac{(z-1)^2}{4\sin^2(\omega h/2)} + \cdots\right].$$

Then

$$K(z) = R(z)B(z) = z^2 - \frac{(z-1)^4}{16 \sin^4 (\omega h/2)},$$

from which the coefficients $K_0, \cdots, K_4$ can be obtained by expansion about $z = 0$.

REFERENCES

1. M. Blum, *An extension of the minimum mean square prediction theory for sampled input signals*, I. R. E. Transactions on Information Theory, Vol. IT-2, No. 3, (1956), pp. 176–184.
2. M. Blum, *On exponential digital filters*, J. Assoc. Comput. Mach., 6 (1959), pp. 283–304.
3. M. Blum, *Fixed memory least squares filters*, I. R. E. Transactions on Information Theory, Vol. IT-3, No. 3, (1957), pp. 178–182.
4. W. M. Brown, *Analysis of discrete linear systems*, this Journal, 5 (1957), pp. 206–224.
5. K. R. Johnson, *Optimum linear, discrete filterings of signals containing a nonrandom component*, I. R. E. Transactions on Information Theory, Vol. IT-2, (1956), pp. 49–55.
6. A. B. Lees, *Interpolation and extrapolation of sampled data*, I. R. E. Transactions on Information Theory, Vol. IT-2, (1956), pp. 12–17.