Discrete Smoothing Filters for Correlated Noise

By J. D. MUSA

(Manuscript received April 18, 1963)

This paper discusses discrete, linear, time-invariant, nonrecursive, finite memory, polynomial smoothing filters for noise that is correlated from sample to sample. The wide-sense Markov process is used as a model for the noise. Analysis and synthesis of the aforementioned filters are discussed in detail and several plots are furnished. A simple method for generating discrete, wide-sense Markov noise for simulation is noted. A noise model composed of a linear combination of wide-sense Markov processes is developed and applied for the case in which the previous model is not sufficiently accurate.

I. INTRODUCTION

A discrete polynomial smoother may be defined in the following terms. Consider the random process $R(nT)$, where n is an integer and T is the period of the samples at which the process will be of interest.* The process will be thought of as comprising a desired component $\bar{R}(nT)$, and a noise component $\tilde{R}(nT)$. It will be assumed that $\bar{R}(nT)$ can be satisfactorily approximated by an rth degree polynomial in nT, $\hat{R}(nT)$. Further, assume that $\tilde{R}(nT)$ is a random process that is wide-sense stationary with respect to the sampling instants nT. The foregoing situation would occur, for example, in the tracking of a moving object whose true position could be represented as an rth degree polynomial in time, and whose measured position included a random error. We will assume that

$$E[\tilde{R}(nT)] = 0 \tag{1}$$

and denote var $[\tilde{R}(nT)]$ = var $[R(nT)]$ by σ_R^2, where E is the expected value operator of probability theory and "var" indicates "variance of".

* Symbols used throughout the paper have been collected and defined in a glossary (Section IX) for ready reference.

Let $\Phi_R(iT)$ represent the autocorrelation* function of $\tilde{R}(nT)$, where i is an integer. A discrete polynomial smoother of the pth order and rth degree is a filter which operates on $R(nT)$ in such a fashion that the output $C(nT)$ and the input $R(nT)$ are related by

$$E\{C(nT)\} = \hat{R}^{(p)}(nT + \Gamma) \tag{2}$$

for all n.† Note that the parenthetical superscript (p) denotes "pth derivative of the estimate with respect to nT." The quantity Γ represents prediction time; if Γ is negative, the operation performed is an interpolation.

We will consider linear, time-invariant smoothers which are nonrecursive and have a finite memory. These conditions may be expressed in terms of the input-output relationship

$$C(nT) = \sum_{i=0}^{N-1} W(iT)R[(n-i)T], \tag{3}$$

where the function $W(iT)$ is the weighting function or impulse response. Note that $W(iT)$ is defined only at a *finite* number of points (N points), that it is independent of the input (hence the smoother is *linear*), and that it is *invariant* with the time nT. No previous values of the output appear in (3); hence the smoother is *nonrecursive*. The latter restriction can often be circumvented, because it is frequently possible to approximate a recursive filter by a nonrecursive one.[1]

The quantity var $[C(nT)] = \sigma_C^2$ is of interest in two respects. First, we may wish to know its value, or better yet, the variance ratio

$$\mu^2 = \frac{\sigma_C^2}{\sigma_R^2}, \tag{4}$$

which is a figure of merit of the smoother. Note that μ^2 is not a function of time, since $R(nT)$ was assumed to be wide-sense stationary, and it follows that $C(nT)$ is also wide-sense stationary by the time-invariance of the smoother. Second, we may wish to find the *optimum* smoother of a class specified by p, r, Γ, N and T; i.e., we may want to determine

* In this paper, the term "autocovariance function" will be used to refer to

$$E[Z_t Z_{t+\tau}],$$

where Z_t represents a zero-mean, wide-sense stationary random process Z evaluated at time t. "Autocorrelation function" will be used to refer to the normalized autocovariance function obtained by dividing the autocovariance function by its value at $\tau = 0$.

† The 0th, 1st, and 2nd order smoothers are often referred to as position, velocity, and acceleration smoothers.

the weighting function $W(iT)$ which yields the minimum value of σ_C^2 (or μ^2) under the preceding conditions.

If $\tilde{R}(nT)$ has an autocorrelation function of arbitrary form, it may be shown, using (1), (3), and (4), that

$$\mu^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_i W_j \Phi_R[(i-j)T], \tag{5}$$

where $W_i = W(iT)$ and $W_j = W(jT)$. In general, this is a complicated expression. In previous treatments[2,3,4,5] of discrete polynomial smoothers, simplification of (5) has been achieved by assuming that the power density spectrum of the noise component of the input is white, so that

$$\Phi_R[(i-j)T] = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j). \end{cases} \tag{6}$$

This yields the simple form

$$\mu^2 = \sum_{i=0}^{N-1} W_i^2. \tag{7}$$

However, the assumption that the noise is uncorrelated from sample to sample is not justified for many physical systems because the noise is restricted in its rate of change. This is particularly true for mechanical and electromechanical systems. It will be shown that correlated noise may be represented by the wide-sense Markov process as a first-order approximation, or by a linear combination of such processes as a better approximation, with appreciable simplification of (5) still being obtained. By "represent" we refer to the approximation of one autocorrelation function or power density spectrum by another. In discussing smoothers, our primary interest is in the behavior of the generalized second moment of random processes, and further delineation of the character of these processes is not necessary.

II. WIDE-SENSE MARKOV NOISE MODEL

A rigorous definition for the wide-sense Markov process may be found in Doob.[6] It will be sufficient for our purposes to characterize the wide-sense Markov process in an alternative fashion, which Doob[7] has shown to be equivalent to the original definition. A wide-sense stationary, continuous random process will be called wide-sense Markov if it has the autocorrelation function

$$\Phi(\tau) = \exp(-\Omega\tau), \qquad \tau \geqq 0. \tag{8}$$

The quantity Ω will be called the "noise bandwidth." By using the evenness property for autocorrelation functions of real, wide-sense stationary random processes, (8) may be written as

$$\Phi(\tau) = \exp(-\Omega \mid \tau \mid). \tag{9}$$

If a wide-sense Markov random process is real and Gaussian and has zero mean, then it is also strict-sense Markov. The strict-sense Markov process is defined as a random process for which

$$\Pr[Y(t_n) \leqq \lambda \mid Y(t_1), \cdots, Y(t_{n-1})] = \Pr[Y(t_n) \leqq \lambda \mid Y(t_{n-1})] \tag{10}$$

with probability 1 for each λ, all $t_1 < \cdots < t_n$, and all n. We may say in an intuitive manner that a strict-sense Markov process is a process with a structure such that any value of the process is directly related only to the immediately preceding value.

One might consider higher-order Markov processes ("related" to several preceding values) as a better approximation for correlated noise, but it appears that using a linear combination of the simple wide-sense Markov processes gives a more manageable expression for μ^2.

For a discrete wide-sense Markov process with equally-spaced samples, we may write the autocorrelation function as

$$\Phi(\tau) = \exp(-\Omega \mid \tau \mid) Cb_T(\tau), \tag{11}$$

where Cb_T is the comb function defined by

$$Cb_T(\tau) = \sum_{i=-\infty}^{\infty} \delta(\tau - iT). \tag{12}$$

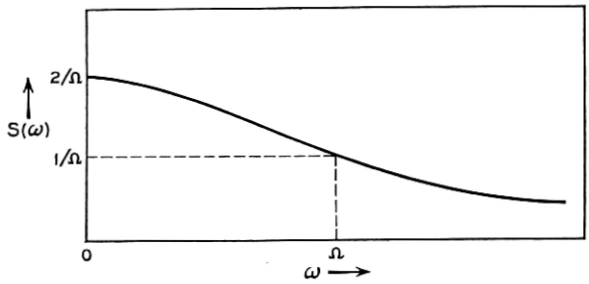

Fig. 1 — Baseband component of normalized power density spectrum for discrete wide-sense Markov process.

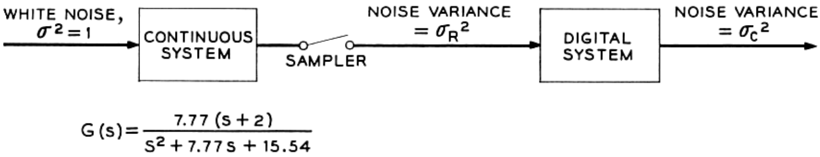

Fig. 2 — Control system used in evaluation of wide-sense Markov noise model.

The normalized† power density spectrum, obtained by Fourier transformation of (11), is

$$S(f) = \frac{2\Omega}{(2\pi f)^2 + \Omega^2} * \frac{1}{T} Cb_{1/T}(f), \tag{13}$$

where $*$ indicates convolution. The baseband component of this normalized power density spectrum is illustrated in Fig. 1. Note that the half-power point occurs at $f = \Omega/2\pi$.

Use of the wide-sense Markov noise model reduces (5) to

$$\mu^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_i W_j \alpha^{|i-j|}, \tag{14}$$

where

$$\alpha = \exp(-\Omega T) \tag{15}$$

and is called the "intersample correlation." For some weighting functions, (14) can be simplified much further by evaluating the sums, using the finite difference calculus.

As one illustration of the improvement in accuracy obtained by representing correlated noise as wide-sense Markov rather than white, consider the control system of Fig. 2. White noise is filtered by the continuous system such that the normalized power density spectrum at the input to the sampler becomes

$$S(\omega) = \frac{|G(j\omega)|^2}{\sigma^2}, \tag{16}$$

where

$$\sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\omega)|^2 \, d\omega = 4.79. \tag{17}$$

† Normalized in the sense that this is the Fourier transform of the autocorrelation function. The power density spectrum is usually defined as the Fourier transform of the autocovariance function. The normalized power density spectrum is equal to the power density spectrum divided by the variance.

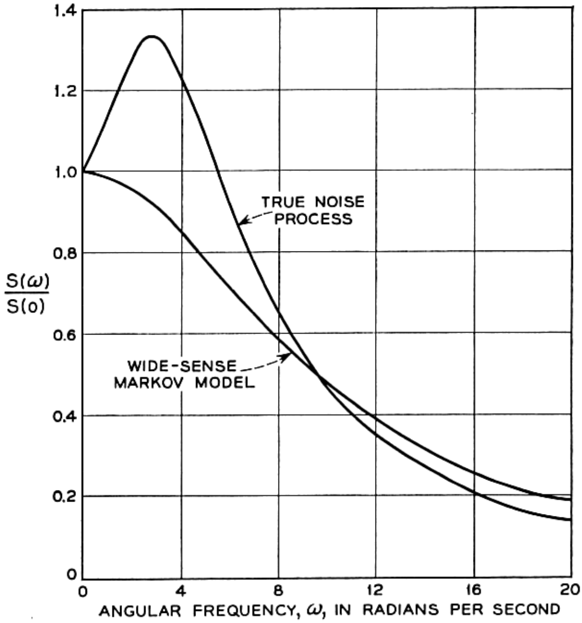

Fig. 3 — Normalized power density spectra.

Hence

$$S(\omega) = \frac{12.6(\omega^2 + 4)}{\omega^4 + 29.3\omega^2 + 241.5}. \tag{18}$$

We can fit models to the true noise process as if all processes were continuous, and following this, introduce the sampling operation. The output-input noise variance ratio μ^2 of the digital system has been computed for the case of a first-order cascaded simple averages smoother† with the following weighting coefficients:

$$W_i = \begin{cases} 0.028257 & (0 \leqq i \leqq 11) \\ 0 & (12 \leqq i \leqq 23) \\ -0.028257 & (24 \leqq i \leqq 35). \end{cases} \tag{19}$$

The true noise process has $\mu^2 = 0.0376$. Use of the wide-sense Markov model yields $\mu^2 = 0.0339$, while use of the white noise model yields $\mu^2 = 0.0192$.

† See Section V for the definition of this smoother.

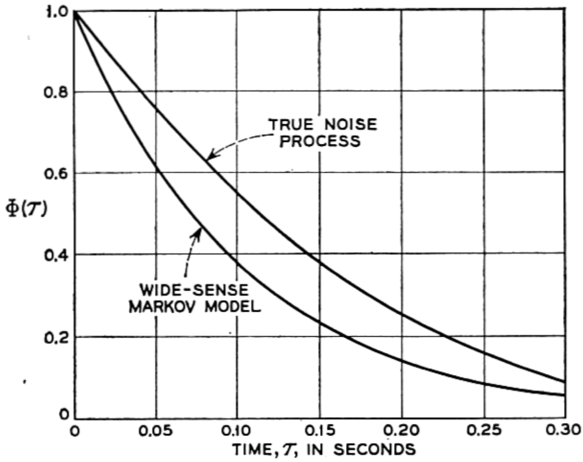

Fig. 4 — Autocorrelation functions.

The normalized power density spectra and autocorrelation functions of the true noise process and the wide-sense Markov model are illustrated in Figs. 3 and 4. The parameter Ω has been picked equal to the half-power point of the power density spectrum of the true noise process, 9.68.

III. MOMENTS OF THE WEIGHTING FUNCTION

The moments of the weighting function of a smoother are important characteristics, since the requirement (2) which specifies the desired output of the smoother is conveniently expressed in terms of them. The moments will be useful in comparing smoothers for equivalence as to meeting (2), and in determining the optimum weighting function for a class of smoothers. The qth moment M_q of the weighting function will be defined as

$$M_q = \sum_{i=0}^{N-1} (iT)^q W_i \,. \tag{20}$$

To express (2) in terms of moments, we proceed as follows. Substituting (3) and (1) in (2) we obtain

$$\sum_{i=0}^{N-1} W_i \bar{R}[(n - i)T] = \hat{R}^{(p)}(nT + \Gamma). \tag{21}$$

Now $\bar{R}(t)$ will be approximated by $\hat{R}(t)$, which may be expressed in the

Taylor series form

$$\hat{R}(t) = \sum_{q=0}^{r} \frac{\hat{R}^{(q)}(nT)}{q!} (t - nT)^q. \tag{22}$$

Substituting (22) in both sides of (21) and rearranging, we obtain

$$\sum_{q=0}^{r} \frac{(-1)^q \hat{R}^{(q)}(nT)}{q!} \sum_{i=0}^{N-1} (iT)^q W_i = \sum_{q=p}^{r} \frac{\hat{R}^{(q)}(nT)}{(q-p)!} \Gamma^{q-p}. \tag{23}$$

Considering (23) term by term, and using (20), we obtain

$$M_q = \begin{cases} 0 & (0 \leqq q < p) \\ (-1)^p p! & (q = p) \\ \dfrac{(-1)^q q!}{(q-p)!} \Gamma^{q-p} & (p < q \leqq r). \end{cases} \tag{24}$$

It should be noted that the weighting function obviously *has* moments greater than the rth; however, the condition (2) does not fix their values.

IV. OPTIMUM SMOOTHERS

By "optimum smoother" we mean that smoother of the class specified by p, r, Γ, N, and T whose weighting function yields the minimum possible value of μ^2. Optimum smoothers are often not implemented because of the amount of storage and computation required. However, they provide a standard of comparison for the systems that are implemented.

To find the weighting function of the optimum smoother of a class, the quantity μ^2 is minimized under the constraints (24), using Lagrange's method of undetermined multipliers. Blackman[8] has carried out the minimization in matrix form for a general input noise process (any autocorrelation function). The optimum smoother is specified by the matrix equation

$$W = P^{-1}A(\tilde{A}P^{-1}A)^{-1}M, \tag{25}$$

where ~ indicates "matrix transpose." The variance ratio μ^2 for the optimum smoother is given by

$$\mu^2 = \tilde{M}(\tilde{A}P^{-1}A)^{-1}M. \tag{26}$$

The matrix W is a column matrix representing the weighting function at the points $t = iT$, i.e.,

$$W = \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_{N-1} \end{bmatrix}; \tag{27}$$

P is the autocorrelation matrix of the input noise process,

$$P = \begin{bmatrix} 1 & \Phi_R(T) & \Phi_R(2T) & \cdots & \Phi_R[(N-1)T] \\ \Phi_R(T) & 1 & \Phi_R(T) & \cdots & \Phi_R[(N-2)T] \\ \Phi_R(2T) & \Phi_R(T) & 1 & \cdots & \Phi_R[(N-3)T] \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_R[(N-1)T] & \Phi_R[(N-2)T] & \Phi_R[(N-3)T] & \cdots & 1 \end{bmatrix}; \tag{28}$$

A is the "age" matrix,

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & T & T^2 & \cdots & T^r \\ 1 & 2T & (2T)^2 & \cdots & (2T)^r \\ 1 & 3T & (3T)^2 & \cdots & (3T)^r \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & (N-1)T & [(N-1)T]^2 & \cdots & [(N-1)T]^r \end{bmatrix}; \tag{29}$$

and M is the column matrix of moments,

$$M = \begin{bmatrix} M_0 \\ M_1 \\ \vdots \\ M_r \end{bmatrix}. \tag{30}$$

Unfortunately, (25) and (26) are very difficult to evaluate literally except in the simplest cases. However, they can be evaluated numerically by a digital computer. The inverse of the autocorrelation matrix for wide-sense Markov noise is readily determined to be, in literal form,

$$P^{-1} = \frac{1}{1-\alpha^2}\begin{bmatrix} 1 & -\alpha & 0 & 0 & \cdots & & 0 \\ -\alpha & 1+\alpha^2 & -\alpha & 0 & & & \\ 0 & -\alpha & 1+\alpha^2 & -\alpha & & & \vdots \\ 0 & 0 & -\alpha & 1+\alpha^2 & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & & & & & 0 \\ & & & & & 1+\alpha^2 & -\alpha \\ 0 & & \cdots & & 0 & -\alpha & 1 \end{bmatrix}. \quad (31)$$

The principal operation, aside from the matrix multiplications, is the inversion of the $(r + 1) \times (r + 1)$ matrix $\tilde{A}P^{-1}A$.

Blackman[8] has evaluated (25) and (26), assuming that the noise is wide-sense Markov, for zero prediction time smoothers with $p = 0$, $r = 0$ and $p = 1$, $r = 1$. For the former,

$$W_i = \begin{cases} \dfrac{1}{N-(N-2)\alpha} & (i = 0, N-1) \\[2ex] \dfrac{1-\alpha}{N-(N-2)\alpha} & (i = 1, 2, \cdots, N-2) \end{cases} \quad (32)$$

and

$$\mu^2 = \frac{1+\alpha}{N-(N-2)\alpha}. \quad (33)$$

For the latter,

$$W_i = \begin{cases} \dfrac{3}{T}\dfrac{(1+\eta)[1+\eta(N-2)]}{(N-1)\{[1+\eta(N-1)][2+\eta(N-1)]+[1-\eta^2]\}} & (i = 0) \\[2ex] \dfrac{6}{T}\dfrac{\eta^2(N-1-2i)}{(N-1)\{[1+\eta(N-1)][2+\eta(N-1)]+[1-\eta^2]\}} & (i = 1, 2, \cdots, N-2) \\[2ex] -\dfrac{3}{T}\dfrac{(1+\eta)[1+\eta(N-2)]}{(N-1)\{[1+\eta(N-1)][2+\eta(N-1)]+[1-\eta^2]\}} & (i = N-1) \end{cases} \quad (34)$$

and

$$\mu^2 = \frac{1}{T^2} \frac{12\eta}{(N-1)\{[1+\eta(N-1)][2+\eta(N-1)]+[1-\eta^2]\}}, \qquad (35)$$

where

$$\eta = \frac{1-\alpha}{1+\alpha}. \qquad (36)$$

These optimum weighting functions and variance ratios have been plotted in a normalized form in Figs. 5, 6, 7, and 8. The ordinates for the 1st order, 1st degree smoother are given in terms of the smoothing interval $T_S = (N-1)T$. The curves are plotted for the parameter $B = \Omega T_S$, which may be thought of as a noise-smoother "bandwidth ratio." The asymptotes for the above curves, as $N \to \infty$ (with T_S and Ω fixed), are derived in Appendix A.

Let us consider the behavior of these curves from a physical viewpoint. For wide-sense Markov noise, the noise autocorrelation function is positive and monotonically decreasing with time. Hence, if the number of samples smoothed, N, is increased with the smoothing interval

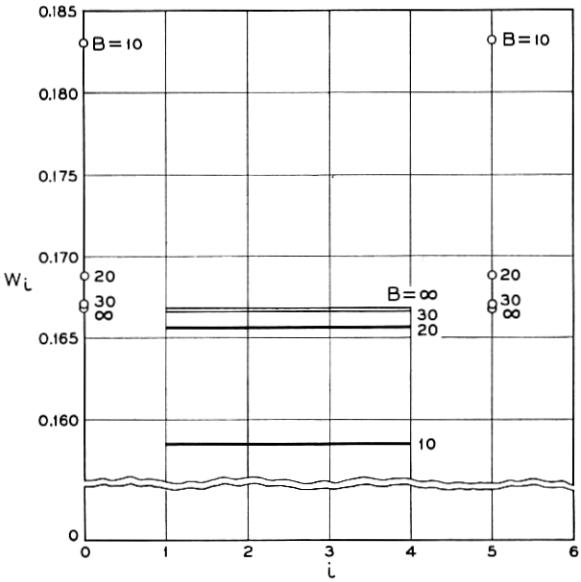

Fig. 5 — Optimum weighting function: 0th order, 0th degree smoother ($\Gamma = 0$, $n = 6$).

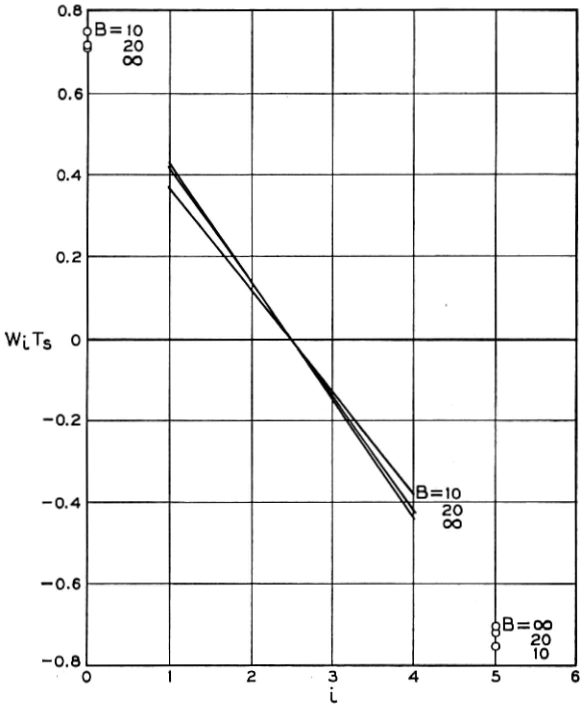

Fig. 6 — Optimum weighting function: 1st order, 1st degree smoother ($\Gamma = 0$, $n = 6$).

T_S and the noise characteristics remaining fixed, the intersample correlation will increase. Although each additional sample provided to the smoother gives additional information, the information added eventually approaches zero due to the increasing correlation. Now a smoother can reduce its variance ratio only by obtaining more information about the noise or by making better use of the information it already has. An optimum smoother makes the best use of the information available to it. Consequently, the variance ratio of an optimum smoother operating on a signal which includes wide-sense Markov noise (or any noise whose autocorrelation function is positive and decreases monotonically with time) must approach a constant as N increases.

V. CASCADED SIMPLE AVERAGES SMOOTHERS

Cascaded simple averages smoothers are a class of smoothers developed by R. B. Blackman.[1] A cascaded simple averages smoother of sth order

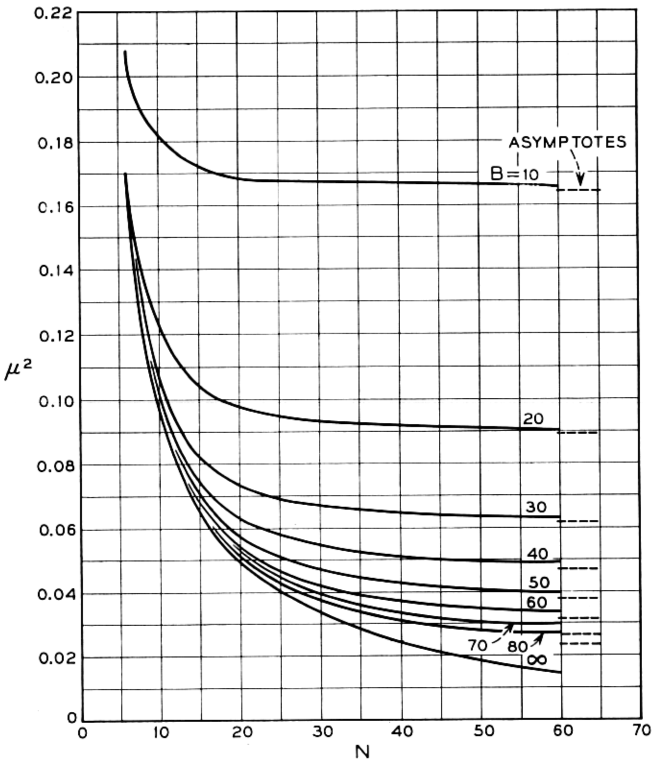

Fig. 7 — Noise variance ratio: optimum 0th order, 0th degree smoother.

approximates an optimum (with respect to white noise) sth order, sth degree, zero prediction time smoother. It may also be used to approximate smoothers that have been optimized with respect to wide-sense Markov noise. The approximation involves using only the values K, $-K$, and 0 for the weighting coefficients, where K is some constant. This smoothing method reduces the amount of storage and the number of arithmetic operations required, at the cost of a slight increase in μ^2 over the optimum method.

The weighting functions of cascaded simple averages smoothers of 0th, 1st, and 2nd orders are as follows (respectively):

$$W_i = \frac{1}{N}, \tag{37}$$

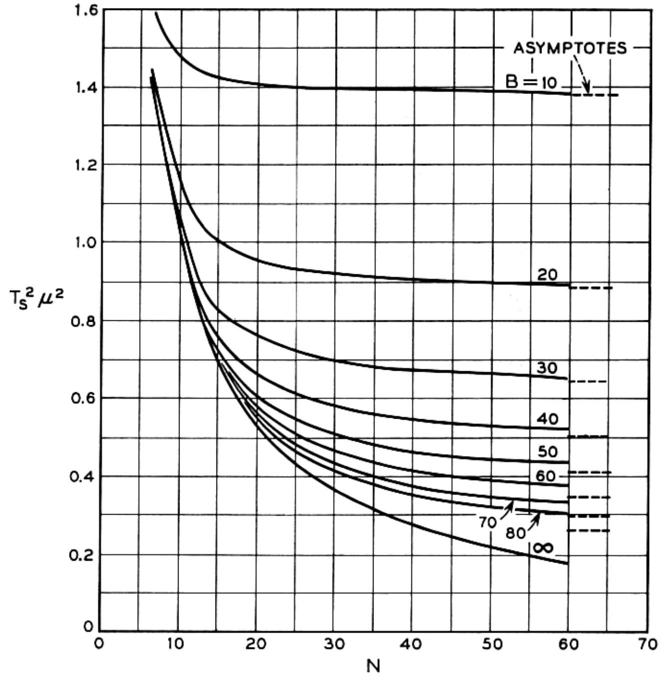

Fig. 8 — Normalized noise variance ratio: optimum 1st order, 1st degree smoother

$$W_i = \begin{cases} \dfrac{4.5(N-1)}{N^2T_S} & \left(0 \leqq i \leqq \dfrac{N}{3} - 1\right) \\ 0 & \left(\dfrac{N}{3} \leqq i \leqq \dfrac{2}{3}N - 1\right) \\ -\dfrac{4.5(N-1)}{N^2T_S} & \left(\dfrac{2}{3}N \leqq i \leqq N - 1\right), \end{cases} \tag{38}$$

and

$$W_i = \begin{cases} \dfrac{36(N-1)^2}{N^3T_S^2} & \left(0 \leqq i \leqq \dfrac{N}{6} - 1\right) \\ 0 & \left(\dfrac{N}{6} \leqq i \leqq \dfrac{N}{3} - 1\right) \\ -\dfrac{36(N-1)^2}{N^3T_S^2} & \left(\dfrac{N}{3} \leqq i \leqq \dfrac{2}{3}N - 1\right) \\ 0 & \left(\dfrac{2}{3}N \leqq i \leqq \dfrac{5}{6}N - 1\right) \\ \dfrac{36(N-1)^2}{N^3T_S^2} & \left(\dfrac{5}{6}N \leqq i \leqq N - 1\right), \end{cases} \tag{39}$$

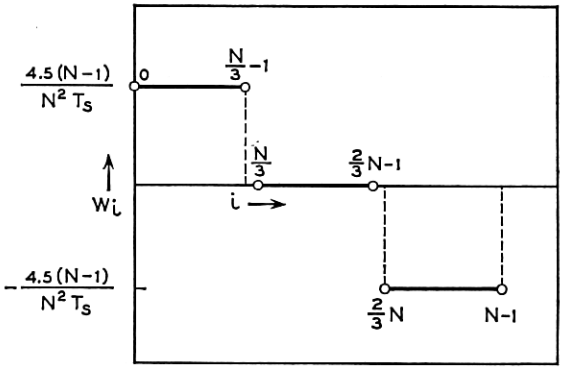

Fig. 9 — Weighting function for 1st order cascaded simple averages smoother.

where N is a multiple of 3 in (38) and a multiple of 6 in (39). The weighting functions for 1st and 2nd order smoothers are plotted in Figs. 9 and 10, respectively.

The variance ratios for 0th, 1st, and 2nd order cascaded simple averages smoothers for a wide-sense Markov noise input are, respectively:

$$\mu^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha^{|i-j|}, \tag{40}$$

$$\mu^2 = \left[\frac{4.5(N-1)}{N^2 T_S}\right]^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} sgn\, W_i \; sgn\, W_j \, \alpha^{|i-j|}, \tag{41}$$

and

$$\mu^2 = \left[\frac{36(N-1)^2}{N^3 T_S^2}\right]^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} sgn\, W_i \; sgn\, W_j \, \alpha^{|i-j|}, \tag{42}$$

where

$$sgn\, W_i = \begin{cases} -1 & (W_i < 0) \\ 0 & (W_i = 0) \\ 1 & (W_i > 0). \end{cases} \tag{43}$$

By use of the finite difference calculus, (40), (41), and (42) may be simplified to

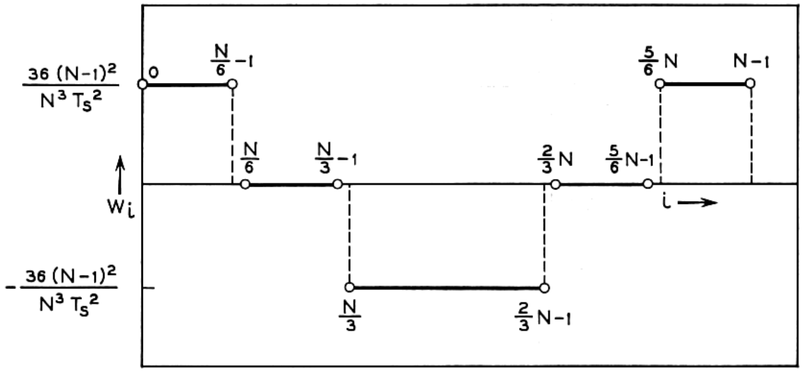

Fig. 10 — Weighting function for 2nd order cascaded simple averages smoother.

$$\mu^2 = \frac{1 + \alpha}{N(1 - \alpha)} + \frac{2\alpha(\alpha^N - 1)}{[N(1 - \alpha)]^2}, \tag{44}$$

$$\mu^2 = 2\left(\frac{4.5}{T_S}\right)^2\left(\frac{N-1}{N}\right)^2 \cdot \left\{\frac{1 + \alpha}{3N(1 - \alpha)} - \frac{\alpha(\alpha^N - 2\alpha^{2N/3} - \alpha^{N/3} + 2)}{[N(1 - \alpha)]^2}\right\}, \tag{45}$$

and

$$\mu^2 = 2\left(\frac{36}{T_S^2}\right)^2\left(\frac{N-1}{N}\right)^4 \cdot \left\{\frac{1 + \alpha}{3N(1 - \alpha)} + \frac{\alpha(\alpha^N - 2\alpha^{5N/6} - \alpha^{2N/3} + 2\alpha^{N/2} + 3\alpha^{N/3} - 3)}{[N(1 - \alpha)]^2}\right\}, \tag{46}$$

respectively.

In Figs. 11, 12, and 13, the variance ratios have been plotted in normalized form for 0th, 1st, and 2nd order cascaded simple averages weighting functions, respectively. The ordinates are μ^2, $T_S^2\mu^2$, and $T_S^4\mu^2$, respectively. The curves are plotted for the noise-smoother "bandwidth ratio" $B = \Omega T_S$. The asymptotes for the above smoothers as $N \to \infty$ (with T_S and Ω fixed) are derived in Appendix A. Note that the expressions simplify appreciably for larger values of B, the exponential terms becoming negligible.

The behavior of these variance ratio curves is somewhat different from those for the optimum smoother. They do not necessarily decrease

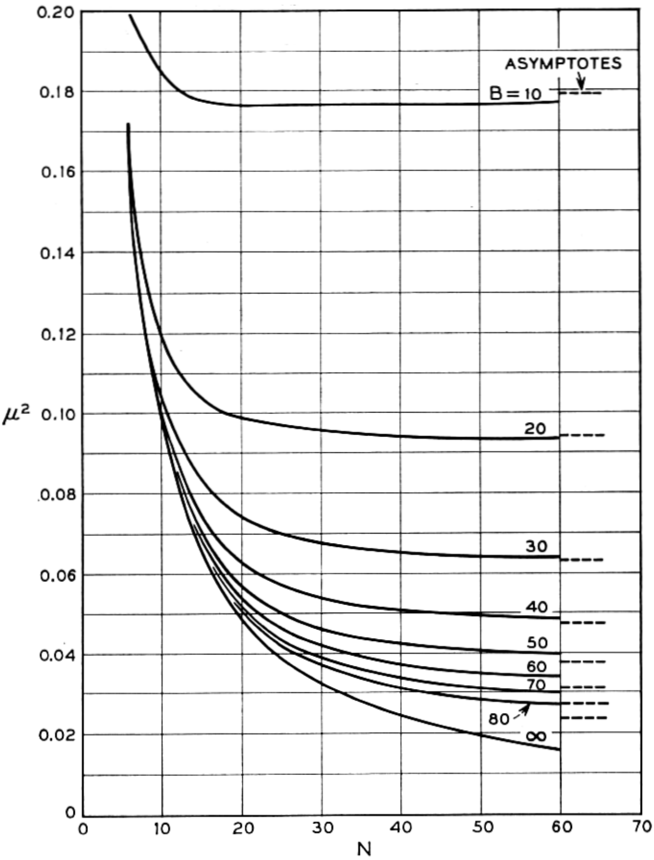

Fig. 11 — Noise variance ratio: 0th order cascaded simple averages smoother.

monotonically with N, even though they have asymptotes similar to the optimum curves. This is due to the fact that the smoothers are not optimum, and therefore the information about the noise is not necessarily utilized in the best manner. Consequently, as N increases, change in variance ratio may be due to changes in the *utilization* of the information available as well as changes in the information available, and the change cannot be readily predicted.

Note that the curves for all three orders of smoothers (Figs. 11, 12, and 13) either have a minimum at some finite value of N or approach a minimum as $N \to \infty$. These minima are more or less broad. In specifying a smoother, it is advantageous to choose the lowest value of N for

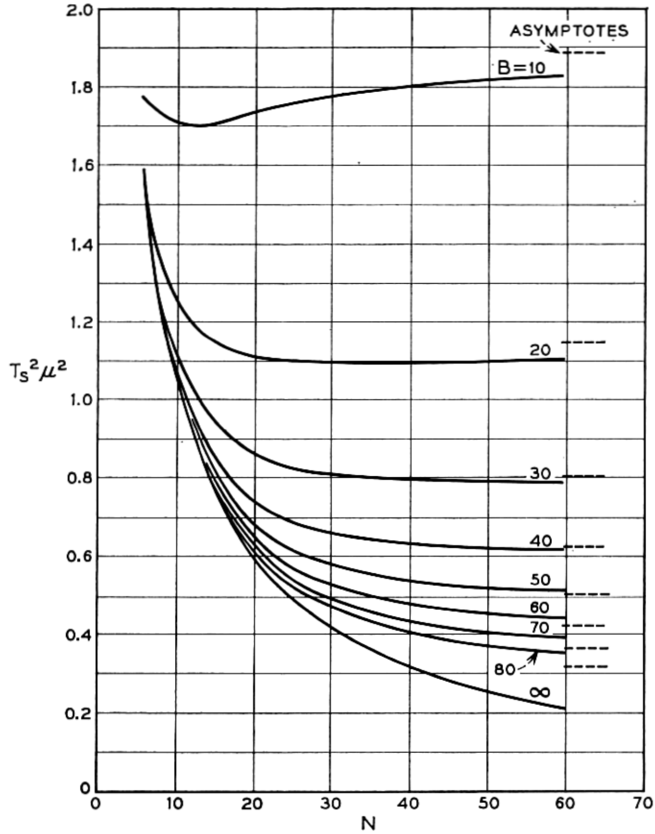

Fig. 12 — Normalized noise variance ratio: 1st order cascaded simple averages smoother.

which μ^2 is reasonably close to the minimum. Note that the neighborhood of the minimum variance ratio as a function of N is reached at lower values of N as B decreases (intersample correlation α increases for fixed T_S). This is reasonable physically, since the value of smoothing a larger number of samples decreases as these samples become more highly correlated.

VI. SYNTHESIS OF POLYNOMIAL SMOOTHERS

In general, the polynomial smoothers we have been discussing are classified by the parameters p, r, Γ, N, and T.† It would be convenient

† The optimum smoother is also classified by the parameter α.

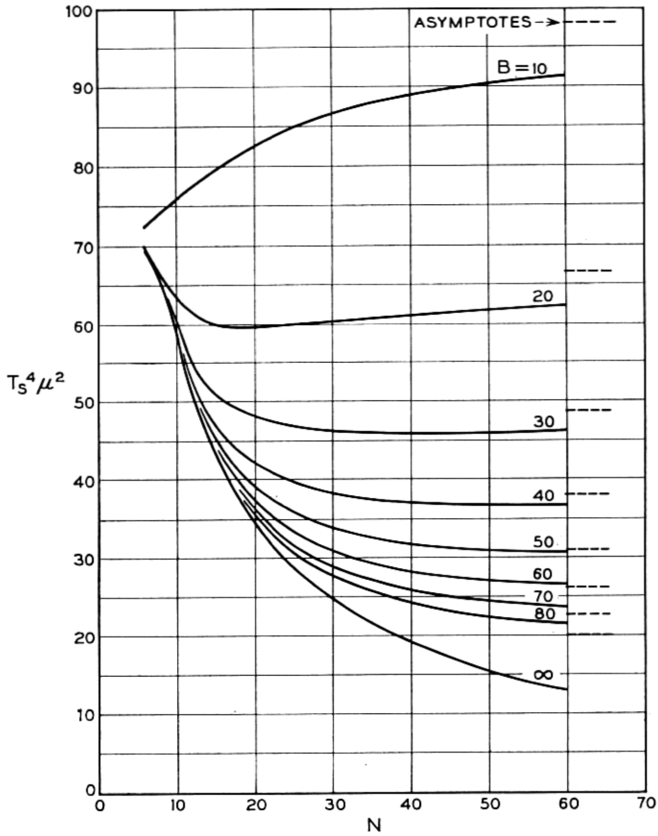

Fig. 13 — Normalized noise variance ratio: 2nd order cascaded simple averages smoother.

to be able to synthesize the smoother *in terms of* sth order, sth degree, zero prediction time components, where $p \leqq s \leqq r$. Note that the components are functions of s, N, and T only; hence their characteristics could be specified fairly simply. Further, several smoothers with different parameters p, r, and Γ but the same N and T could be synthesized with common components by weighting these components differently. Finally, the above breakdown permits any polynomial smoother of the class considered in this paper to be constructed from cascaded simple averages components. The derivation and procedures discussed in this section are valid for discrete polynomial smoothers in general and are not restricted to optimum smoothers or to particular input noise power density spectra.

Consider the linear combination of sth order, sth degree, zero prediction time components shown in Fig. 14. Let W_{si} represent the value of

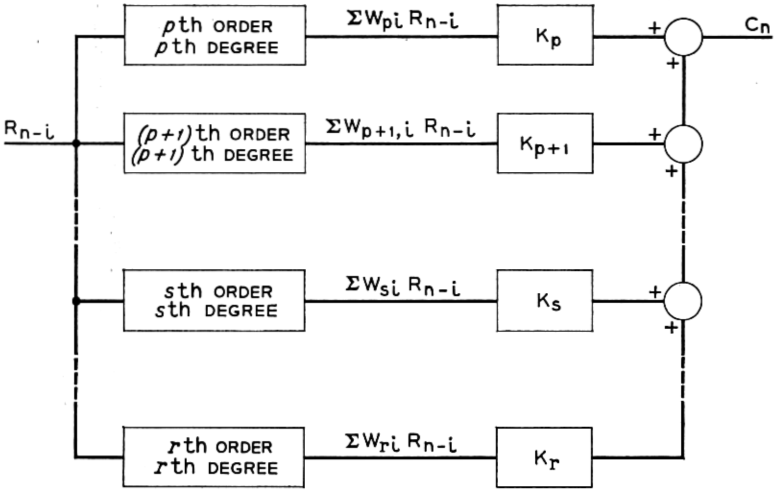

Fig. 14 — Synthesis of pth order, rth degree smoother from sth order, sth degree components.

the weighting function of the sth component at the sample with age iT. Let M_{sq} be the qth moment of the weighting function of the sth component. From Fig. 14 it will be seen that the "over-all" weighting function W_i of the entire smoother is related to the component weighting functions by

$$W_i = \sum_{s=p}^{r} K_s W_{si} . \tag{47}$$

Now, using (47) and (20),

$$M_q = T^q \sum_{i=0}^{N-1} i^q W_i = T^q \sum_{s=p}^{r} K_s \sum_{i=0}^{N-1} i^q W_{si} = \sum_{s=p}^{r} K_s M_{sq} . \tag{48}$$

From (24) we obtain

$$M_{sq} = \begin{cases} (-1)^q q! & (s = q) \\ 0 & (s > q). \end{cases} \tag{49}$$

Substituting (49) in (48) we get

$$M_q = \sum_{s=p}^{q-1} K_s M_{sq} + K_q (-1)^q q!. \tag{50}$$

If the linear combination of components is to be equivalent to the pth order, rth degree smoother, then (24) must be satisfied. It follows that we must have

$$K_s = \begin{cases} 1 & (s = p) \\ \dfrac{\Gamma^{s-p}}{(s-p)!} - \dfrac{(-1)^s}{s!} \sum_{u=p}^{s-1} K_u M_{us} & (p < s \leqq r). \end{cases} \quad (51)$$

It should be noted that a linear combination of optimum components will not necessarily be optimum unless the outputs of the components at a common time are uncorrelated.

The synthesis procedure proper consists of finding a smoother or set of component smoothers which produces the desired output (2) with the least total error ϵ compatible with a simple implementation. In the case of recursive smoothers, stability must be considered; the latter topic is adequately covered in standard texts on control theory.[9] The total error ϵ is given by

$$\epsilon = [\sigma_C^2 + \epsilon_T^2]^{\frac{1}{2}}, \quad (52)$$

where ϵ_T is the truncation error

$$\epsilon_T = \bar{R}^{(p)}(nT + \Gamma) - \hat{R}^{(p)}(nT + \Gamma). \quad (53)$$

Alternatively, using (21), we may write (53) as

$$\epsilon_T = \bar{R}^{(p)}(nT + \Gamma) - \sum_{i=0}^{N-1} W_i \bar{R}[(n - i)T]. \quad (54)$$

Synthesis involves the choice of type of filter (optimum, cascaded simple averages, etc.) and the selection of r, N and T. For convenience, the parameters will be selected in the alternate form r, N, and T_s.

The selection of r is based on the requirement that $r \geqq p$ and the direction of change in ϵ as r increases. Now σ_C^2 increases and ϵ_T decreases (in general) with increasing r. The rate of increase of σ_C^2 with r is such that smoothers with $r > 2$ are seldom used in practice.

The selection of N and T_s will be a trial-and-error process based on achieving a near minimum in ϵ while keeping N as small as possible (for simpler implementation). In the case of a set of components, each component may have a different value of T_s provided the values of T are the same. Figs. 7, 8, 11, 12, and 13 will be useful in calculating σ_C^2. When calculating the over-all output noise of a set of component smoothers, it will be useful to know that the noise outputs of 0th, 1st, and 2nd order cascaded simple averages smoothers are all mutually uncorrelated, though this is not true for all orders.[1]

The problems involved in estimating truncation error have been discussed by Hamming[10] in some detail. We will make the simplifying assumption that the truncation error ϵ_T of an rth degree smoother may be approximated by using (54) with $\bar{R}(t)$ considered as an $(r + 1)$th degree polynomial. Thus $\bar{R}(t)$ may be expressed

$$\bar{R}(t) = \sum_{q=0}^{r+1} \frac{\bar{R}^{(q)}(nT)}{q!} (t - nT)^q. \tag{55}$$

Substituting (55) into (54), and using (22) and (26), we obtain

$$\epsilon_T = \bar{R}^{(r+1)}(nT) \left[\frac{\Gamma^{r-p+1}}{(r-p+1)!} - (-1)^{r+1} \frac{M_{r+1}}{(r+1)!} \right]. \tag{56}$$

Blackman[11] has calculated the $(r + 1)$th moments of 0th, 1st, and 2nd order cascaded simple averages smoothers as $\frac{1}{2}T_s$, $-T_s$, and $3T_s$, respectively. Hence, the truncation errors for these smoothers may be calculated from (56) as $\frac{1}{2}T_s\bar{R}^{(r+1)}(nT)$.

VII. GENERATION OF DISCRETE WIDE-SENSE MARKOV NOISE FOR SIMULATION

It is frequently desired to simulate the performance of discrete smoothing filters and perhaps larger discrete systems of which they may be a part. Standard techniques are available for simulating discrete white noise by generation of a sequence of uncorrelated pseudo-random numbers.[12,13] It is relatively easy to generate discrete wide-sense Markov noise from such a sequence, due to the simple correlation structure of the wide-sense Markov process. The foregoing is another advantage in using the wide-sense Markov model to represent correlated noise.

Let $\{Y_n\}$ be the desired discrete wide-sense Markov noise and $\{X_n\}$ be a sequence of uncorrelated random numbers of zero mean and unit variance. Then Y_n may be generated as

$$Y_1 = \sigma X_1, \tag{57}$$

$$Y_n = \alpha Y_{n-1} + \sigma\sqrt{1 - \alpha^2}\, X_n, \qquad (n > 1), \tag{58}$$

where σ^2 is the variance and α the intersample correlation of the wide-sense Markov noise.

Since Y_n is in effect a linear combination of the X_{n-i}, $i = 0, \cdots, n - 1$, it follows that if the X_{n-i} are jointly Gaussian, then the Y_n are jointly Gaussian.

VIII. DISCRETE SMOOTHING FILTERS BASED ON A MODEL USING A LINEAR COMBINATION OF WIDE-SENSE MARKOV PROCESSES

In some cases the simple wide-sense Markov noise model may not be a sufficiently accurate representation of a physical noise process. A better model may be obtained by approximating the known or assumed noise process by a linear combination of wide-sense Markov noise processes. We may approximate the autocorrelation function by wide-sense Markov autocorrelation functions, or, equivalently, we may approximate the normalized power density spectrum by wide-sense Markov normalized power density spectra. For the purpose of making the preceding approximations, we can work with the process as though it were continuous, later introducing the sampling operation. In a parallel to the use of Fourier series to analyze the behavior of a complicated waveform in a linear system, the wide-sense Markov autocorrelation functions may be used to analyze the behavior of a complicated correlated random noise process in a linear discrete system, by applying the principle of superposition. It is possible to synthesize discrete smoothers using this more complex model.

Further, discrete random noise of arbitrary power density spectrum may be generated in an approximate manner for simulation purposes by a suitable linear combination of wide-sense Markov noise components. In the preceding applications, the use of the wide-sense Markov noise components is simpler and more efficient than use of the actual noise process.

There are two types of approximations that can be made. One is a cut-and-try type of approximation in which one tries various linear combinations of wide-sense Markov noise components with the bandwidths of the components not necessarily being integral multiples of some fundamental bandwidth. The other approach is to use a linear combination of orthonormal functions of wide-sense Markov components. In the latter approach, the bandwidths of the components are integral multiples of a fundamental component. In either case, we may write

$$\Phi(\tau) = \sum_{v=1}^{z} A_v \exp(-\Omega_v \mid \tau \mid) \tag{59}$$

or

$$S(\omega) = \sum_{v=1}^{z} A_v \frac{2\Omega_v}{\Omega_v^2 + \omega^2}. \tag{60}$$

Note that the sum of the coefficients A_v must be equal to 1. In the orthonormal approximation,

$$\Omega_v = v\Omega \tag{61}$$

and the A_v will have a definite form. This is shown in the following section.

8.1 *Orthonormal Approximation*

Laning and Battin[14] and Lee[15] have developed orthonormal approximations for an arbitrary autocorrelation function and an arbitrary normalized power density spectrum. These approximations are in terms of components which will be recognized as wide-sense Markov auto-

TABLE I—VALUES OF COEFFICIENTS c_{kv}

k	c_{k1}	c_{k2}	c_{k3}	c_{k4}	c_{k5}
1	1				
2	2	−3			
3	3	−12	10		
4	4	−30	60	− 35	
5	5	−60	210	−280	126

correlation functions and normalized power spectra, respectively. We shall develop the approximation in somewhat different form.

The set of functions

$$\Phi_k(\tau) = \sum_{v=1}^{k} c_{kv} \sqrt{k\Omega} \exp(-v\Omega |\tau|) \tag{62}$$

can be made orthonormal on the interval $-\infty < \tau < \infty$ by proper choice of the coefficients c_{kv}. These coefficients are listed in Table I for values of k up to 5.

These functions may be used to form an orthogonal expansion of any piecewise continuous even function (and hence any piecewise continuous autocorrelation function) on the interval $-\infty < \tau < \infty$. We may write

$$\Phi(\tau) = \sum_{k=1}^{\infty} a_k \Phi_k(\tau), \tag{63}$$

where

$$a_k = \int_{-\infty}^{\infty} \Phi(\tau)\Phi_k(\tau)\, d\tau. \tag{64}$$

If we take z terms of the series expansion and denote the corresponding partial sum $\hat{\Phi}(\tau)$, we may group terms to obtain

$$\Phi(\tau) \approx \hat{\Phi}(\tau) = \sum_{v=1}^{z} A_v \exp(-v\Omega |\tau|), \tag{65}$$

where

$$A_v = \sum_{k=v}^{z} a_k c_{kv} \sqrt{k\Omega}. \tag{66}$$

The coefficients A_v for $z \leqq 5$ are given by (note that if $z < 5$, then $a_k = 0$ for $k > z$)

$$A_1 = \sqrt{\Omega}\,[a_1 + 2\sqrt{2}\,a_2 + 3\sqrt{3}\,a_3 + 4\sqrt{4}\,a_4 + 5\sqrt{5}\,a_5], \tag{67}$$

$$A_2 = -3\sqrt{\Omega}\,[\sqrt{2}\,a_2 + 4\sqrt{3}\,a_3 + 10\sqrt{4}\,a_4 + 20\sqrt{5}\,a_5], \tag{68}$$

$$A_3 = 10\sqrt{\Omega}\,[\sqrt{3}\,a_3 + 6\sqrt{4}\,a_4 + 21\sqrt{5}\,a_5], \tag{69}$$

$$A_4 = -35\sqrt{\Omega}\,[\sqrt{4}\,a_4 + 8\sqrt{5}\,a_5], \tag{70}$$

$$A_5 = 126\sqrt{\Omega}\,[\sqrt{5}\,a_5]. \tag{71}$$

Now let $S_k(\omega)$ be the Fourier transform of $\Phi_k(\tau)$. Then

$$S_k(\omega) = \sum_{v=1}^{k} c_{kv}\sqrt{k\Omega}\,\frac{2v\Omega}{(v\Omega)^2 + \omega^2}. \tag{72}$$

It can be shown, using Parseval's theorem, that the set of functions $\{(1/\sqrt{2\pi})S_k(\omega)\}$ is orthonormal on the interval $-\infty < \omega < \infty$. Hence we may expand any piecewise continuous even function (and thus any piecewise continuous normalized power density spectrum) on this interval. We may write

$$S(\omega) = \sum_{k=1}^{\infty} a_k S_k(\omega), \tag{73}$$

where

$$a_k = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)S_k(\omega)\,d\omega. \tag{74}$$

From Parseval's theorem it will be seen that the a_k in (74) are the same as those in (64). If we take z terms of the series expansion and denote the corresponding partial sum $\hat{S}(\omega)$, we may group terms as before to obtain

$$S(\omega) \approx \hat{S}(\omega) = \sum_{v=1}^{z} A_v\,\frac{2v\Omega}{(v\Omega)^2 + \omega^2}, \tag{75}$$

where the A_v are given by (67) through (71). Note that (65) and (75) form a Fourier transform pair. Thus, if we have approximated a noise process in terms of normalized power density spectra or autocorrelation functions, the alternative approximation can be immediately obtained. The quantity $v\Omega$ represents the half-power point for each wide-sense Markov component.

Simple rules for the selection of Ω for a particular expansion cannot be established; it is a matter of judgment and perhaps trial and error. The fact that it is the half-power point of the fundamental component of the approximation may be of some help. Also, note that as $\tau \to \infty$, $\hat{\Phi}(\tau) \to A_1 \exp(-\Omega\,|\,\tau\,|)$. We might choose Ω that $\hat{\Phi}(\tau)$ and $\Phi(\tau)$ approach zero at the same rate. However, matching autocorrelation functions by means of their tails is not necessarily a desirable approach.

8.2 *System Analysis and Synthesis*

The noise variance ratio of a linear discrete system for which the input noise autocorrelation function has been approximated by a linear combination of wide-sense Markov autocorrelation functions may be obtained by substituting (59) in (5). We have

$$\mu^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{v=1}^{z} W_i W_j A_v \exp\left(-\Omega_v T \mid i - j \mid\right). \tag{76}$$

Now let

$$\alpha_v = \exp\left(-\Omega_v T\right). \tag{77}$$

Then

$$\mu^2 = \sum_{v=1}^{z} A_v \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_i W_j \alpha_v^{\mid i-j \mid} \right]. \tag{78}$$

The expression in brackets represents the noise variance ratio of the linear discrete system when the vth wide-sense Markov component of the noise is the input. Thus, it is clear that the principle of superposition can be used to find the total noise variance ratio. Figs. 7, 8, 11, 12, and 13 may be applied to the wide-sense Markov components individually.

Use of the linear combination noise model will not be profitable in determining the optimum smoother of a class. There is a matrix inversion required [refer to (25)] which is more easily performed directly with the actual autocorrelation matrix. One should keep in mind that the noise variance ratio of a digital smoother is relatively insensitive to departures of the weighting function from the optimum. Hence a smoother optimized for the simple wide-sense Markov model may be satisfactory.

The synthesis of a polynomial smoother based on the linear combination noise model follows the method of Section VI, except that calculation of ϵ is somewhat more difficult, since σ_C^2 must be calculated using (78) and the relevant plots. It should be noted that the estimate of ϵ_T may not be sufficiently accurate to justify the use of the linear combination noise model. One should consider whether or not the simple wide-sense Markov model might be satisfactory.

8.3 *Noise Generation for Simulation*

Discrete stationary random noise of arbitrary autocorrelation function $\Phi(\tau)$ and variance σ^2 may be approximately generated as a linear combination of independent, wide-sense Markov components. Let

$\hat{Z}_i$ represent the ith sample of the approximating linear combination

$$\hat{Z}_i = \sum_{v=1}^{z} b_v Y_{vi}, \tag{79}$$

where Y_{vi} is the ith sample of the vth wide-sense Markov component. This vth component is generated (refer to Section VII) as

$$Y_{v1} = X_{v1}, \tag{80}$$

$$Y_{vi} = \alpha_v Y_{v,i-1} + \sqrt{1 - \alpha_v^2}\, X_{vi}, \qquad (i > 1). \tag{81}$$

Care must be taken that the normalized uncorrelated random numbers X_{vi} are generated in z similar but mutually uncorrelated sequences $\{X_{1i}\}, \{X_{2i}\}, \ldots, \{X_{zi}\}$ to ensure that the sequences $\{Y_{1i}\}, \{Y_{2i}\}, \ldots, \{Y_{zi}\}$ are mutually uncorrelated. Note that each Y_{vi} will have zero mean and unit variance.

To evaluate the coefficients b_v, approximate the autocorrelation function of the arbitrary random noise process by a linear combination of wide-sense Markov components. Thus, from (59) and (77) we obtain

$$\Phi(\,|\,i - j\,|\,T) \approx \hat{\Phi}(\,|\,i - j\,|\,T) = \sum_{v=1}^{z} A_v \alpha_v^{|i-j|}. \tag{82}$$

Now since the Z process is stationary

$$\Phi(\,|\,i - j\,|\,T) = \frac{\operatorname{cov}(\hat{Z}_i, \hat{Z}_j)}{\operatorname{var}(\hat{Z}_i)} = \frac{\sum_{v=1}^{z} b_v^2 \operatorname{cov}(Y_{vi}, Y_{vj})}{\sigma^2} = \frac{\sum_{v=1}^{z} b_v^2 \alpha_v^{|i-j|}}{\sigma^2}. \tag{83}$$

We have set $\operatorname{var}(\hat{Z}_i) = \sigma^2$ since the arbitrary process and its approximation must be matched in variance. Now, equating terms of (82) and (83), we obtain

$$b_v = \sigma \sqrt{A_v}. \tag{84}$$

Thus,

$$\hat{Z}_i = \sigma \sum_{v=1}^{z} \sqrt{A_v}\, Y_{vi}. \tag{85}$$

IX. GLOSSARY OF SYMBOLS

A_v = Coefficient in approximation of power density spectrum or autocorrelation function by linear combination of wide-sense Markov components

B = ΩT_S = noise-smoother bandwidth ratio for fundamental wide-sense Markov component

C = output signal of smoother

E = expected value operator = $\int_{-\infty}^{\infty} dF$

f = dF = probability density function

M = matrix of moments of weighting coefficients

M_q = $\sum_{i=0}^{N-1} i^q T^q W_i$ = qth moment of weighting function

n = present sample

N = number of samples operated on by nonrecursive discrete smoother

p = order of smoother

P = autocorrelation matrix

r = degree of smoother

R = total input signal to smoother

R_{m-i} = total input signal evaluated at $t = (m - i)T$

$\bar{R}$ = desired component of input signal to smoother

$\bar{R}^{(p)}$ = pth derivative of desired component of input signal

$\hat{R}$ = polynomial approximation to desired component of input signal

$\tilde{R}$ = noise component of input signal

$S(\omega)$ = power density spectrum (normalized sense, $\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\, d\omega = 1$); Fourier transform of $\Phi(\tau)$.

t,τ = time variables (seconds)

T = sampling interval (seconds)

T_S = smoothing interval (seconds)

W_i = weighting function of digital filter evaluated at $t = iT$

$X(t)$ = white noise process

$Y(t)$ = wide-sense Markov noise process

$Z(t)$ = general noise process

$\hat{Z}(t)$ = approximation to general noise process

α = $\exp(-\Omega T)$ = intersample correlation for fundamental wide-sense Markov component

α_v = $\exp(-\Omega_v T)$ = intersample correlation for vth wide-sense Markov component

Γ = prediction time

ϵ = total output error of smoother
ϵ_T = truncation error
μ = σ_C/σ_R = output-input standard deviation ratio
σ^2 = noise variance
σ_C^2 = output noise variance
σ_R^2 = input noise variance
$\Phi(\tau)$ = autocorrelation function
ω = angular frequency variable (radians/sec)
Ω = bandwidth of fundamental wide-sense Markov power density spectrum (radians/sec)
Ω_ν = bandwidth of νth wide-sense Markov power density spectrum component (radians/sec)

X. ACKNOWLEDGMENTS

The author wishes to express his appreciation to all those who, directly or indirectly, have contributed to this paper. Particular thanks are due to N. Sharko, M. Iwama, C. E. Paul, V. O. Mowery, J. A. Cochran, and J. P. O'Donohue of Bell Telephone Laboratories, W. S. Spracklen of Texas Instruments, and D. L. Williams of the University of Michigan for their constructive comments. He is most grateful to Miss C. Ziegler and Miss M. Devine of Bell Telephone Laboratories for the programming and many of the illustrations required for this paper.

APPENDIX

Asymptotic Behavior of Smoothers

We will consider the behavior of μ^2 as $N \to \infty$ with Ω and T_S fixed. We shall first find the limits of two expressions which will be needed in finding the limits of the larger noise variance ratio expressions:

$$\begin{aligned}\lim_{N\to\infty} \alpha^{aN+b} &= \lim_{N\to\infty} \exp\left[-\Omega T(aN+b)\right] = \lim_{N\to\infty} \exp\left[-\frac{\Omega T_S(aN+b)}{N-1}\right] \\ &= \lim_{N\to\infty} \exp\left[-\frac{B(aN+b)}{N-1}\right] = \exp(-aB),\end{aligned} \tag{86}$$

$$\begin{aligned}\lim_{N\to\infty} N(1-\alpha) &= \lim_{N\to\infty} \frac{1-\exp\left[-B/(N-1)\right]}{1/N} \\ &= \lim_{N\to\infty} \frac{N^2 B \exp\left[-B/(N-1)\right]}{(N-1)^2} = B.\end{aligned} \tag{87}$$

A.1 *Optimum Smoother—0th Order, 0th Degree*

We have, using (33)

$$\lim_{N\to\infty} \mu^2 = \lim_{N\to\infty} \frac{1+\alpha}{N-(N-2)\alpha} = \lim_{N\to\infty} \frac{1+\alpha}{N(1-\alpha)+2\alpha} = \frac{2}{B+2}. \quad (88)$$

A.2 *Optimum Smoother—1st Order, 1st Degree*

We have, using (35) and (36):

$$\lim_{N\to\infty} T_s^2\mu^2 = \lim_{N\to\infty} \frac{12(1+\alpha)[N(1-\alpha)-1+\alpha]}{[N(1-\alpha)+2\alpha][N(1-\alpha)+1+3\alpha]+4\alpha}$$
$$= \frac{24B}{(B+2)(B+4)+4} = \frac{24B}{B^2+6B+12}. \quad (89)$$

A.3 *Zeroth Order Cascaded Simple Averages Smoother*

Using (44) we have

$$\lim_{N\to\infty} \mu^2 = \lim_{N\to\infty} \left\{\frac{1+\alpha}{N(1-\alpha)} + \frac{2\alpha(\alpha^N-1)}{[N(1-\alpha)]^2}\right\}$$
$$= \frac{2}{B} + \frac{2}{B^2}[\exp(-B)-1] = \frac{2}{B^2}[\exp(-B)+B-1]. \quad (90)$$

A.4 *First Order Cascaded Simple Averages Smoother*

We have, using (45),

$$\lim_{N\to\infty} T_s^2\mu^2 = \lim_{N\to\infty} 2(4.5)^2\left(\frac{N-1}{N}\right)^2$$
$$\cdot\left\{\frac{1+\alpha}{3N(1-\alpha)} - \frac{\alpha(\alpha^N-2\alpha^{2N/3}-\alpha^{N/3}+2)}{[N(1-\alpha)]^2}\right\}$$
$$= \frac{40.5}{B^2}\left\{-\exp(-B)+2\exp(-2B/3)\right.$$
$$\left.+\exp(-B/3)+\frac{2}{3}B-2\right\}. \quad (91)$$

A.5 *Second Order Cascaded Simple Averages Smoother*

Using (46) we have

$$\lim_{N\to\infty} T_S{}^4\mu^2 = \lim_{N\to\infty} 2(36)^2 \left(\frac{N-1}{N}\right)^4 \left\{\frac{1+\alpha}{3N(1-\alpha)} + \frac{\alpha(\alpha^N - 2\alpha^{5N/6} - \alpha^{2N/3} + 2\alpha^{N/2} + 3\alpha^{N/3} - 3)}{[N(1-\alpha)]^2}\right\}$$

$$= \frac{2592}{B^2}\left\{\exp(-B) - 2\exp(-5B/6) - \exp(-2B/3) + 2\exp(-B/2) + 3\exp -B/3) + \frac{2}{3}B - 3\right\}. \quad (92)$$

REFERENCES

1. Blackman, R. B., Smoothing and Prediction of Time Series by Cascaded Simple Averages, 1960 IRE Convention Record, **8,** Part 2, March 21–24, pp. 47–54. Also published as Bell System Monograph 3678.
2. Lees, A. B., Interpolation and Extrapolation of Sampled Data, IRE Trans. on Information Theory, **IT-2,** 1, March, 1956, pp. 12–17.
3. Johnson, K. R., Optimum, Linear, Discrete Filtering of Signals Containing a Nonrandom Component, IRE Trans. on Information Theory, **IT-2,** 2, June, 1956, pp. 49–55.
4. Blum, M., An Extension of the Minimum Mean Square Prediction Theory for Sampled Input Signals, IRE Trans. on Information Theory, **IT-2,** 3, September, 1956, pp. 176–184.
5. Blum, M., On the Mean Square Noise Power of an Optimum Linear Discrete Filter Operating on Polynomial Plus White Noise Input, IRE Trans. on Information Theory, **IT-3,** 4, December, 1957, pp. 225–231.
6. Doob, J. L., *Stochastic Processes*, Wiley, New York, 1953, p. 90.
7. Doob, J. L., op. cit., pp. 233–234.
8. Blackman, R. B., *Linear Data Smoothing and Prediction in Theory and Practice*, to be published.
9. Ragazzini, J. R., and Franklin, G. F., *Sampled-data Control Systems*, McGraw-Hill, New York, 1958.
10. Hamming, R. W., *Numerical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1962, pp. 143–152.
11. Blackman, R. B., *Linear Data Smoothing and Prediction in Theory and Practice.*
12. Taussky, Olga, and Todd, J., Generation and Testing of Pseudo-Random Numbers, *Symposium on Monte Carlo Methods*, Wiley, New York, 1956, pp. 15–28.
13. Chartres, B. A., An Exact Method of Generating Random Normal Deviates, T. R. No. 5, Contract No. DA-30-069-SC-78130, Division of Applied Mathematics, Brown University, March, 1959.
14. Laning, J. H., and Battin, R. H., *Random Processes in Automatic Control*, McGraw-Hill, New York, 1956, pp. 381–394.
15. Lee, Y. W., *Statistical Theory of Communication*, Wiley, New York, 1960, pp. 460–469.