VIII. STATISTICAL COMMUNICATION THEORY

Prof. Y. W. Lee
Prof. A. G. Bose
Prof. I. M. Jacobs
Prof. J. B. Wiesner
Dr. A. D. Hause
V. R. Algazi
R. A. Bruce
D. F. De Long, Jr.
C. L. Feldman
H. Hemami
F. J. Hernandez
D. J. Sakrison
M. Schetzen
H. L. Van Trees, Jr.
S. Weinreb
W. S. Widnall

A. WORK COMPLETED

1. SOME PROBLEMS IN NONLINEAR THEORY

The present study has been completed by M. Schetzen. In May 1961, he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science. The study will also be published as Technical Report 390.

Y. W. Lee

2. SYNTHESIS OF OPTIMUM NONLINEAR CONTROL SYSTEMS

This study was completed and presented by H. L. Van Trees, Jr. as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science, to the Department of Electrical Engineering, M.I.T., May 1961.

Y. W. Lee

3. APPLICATION OF STOCHASTIC APPROXIMATION METHODS TO SYSTEM OPTIMIZATION

This study has been completed by D. J. Sakrison. It was submitted as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science, Department of Electrical Engineering, M.I.T., May 1961, and will also be published as Technical Report 391.

A. G. Bose

4. A STUDY OF IMPULSE RESPONSES IN THE RECORDING AND REPRODUCTION OF SOUND

This study has been completed by J. Hernandez-Figueora. It was submitted as a thesis in partial fulfillment of the requirements for the degree of Master of Science, Department of Electrical Engineering, M.I.T., May 1961.

A. G. Bose

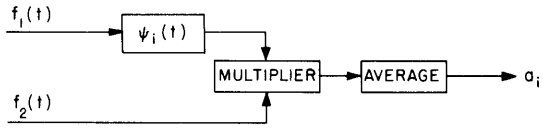


Fig. VIII-1. Circuit for the experimental determination of a.

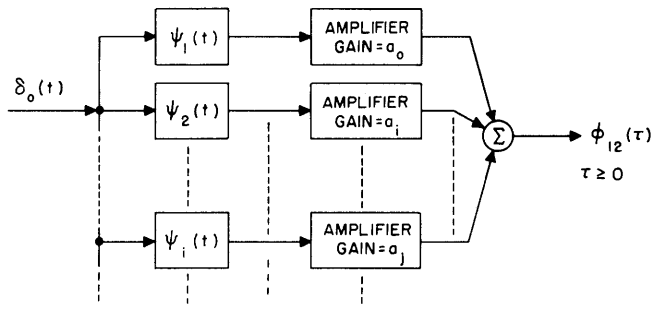


Fig. VIII-2. Network synthesis of $\phi_{12}(\tau)$.

for synthesizing the correlation function. Second, it is not possible to average over an infinite time interval. Several experiments that demonstrate this measurement technique and illustrate the errors that can occur are presented in this report.

Lampard (2) has used the orthogonal expansion method to measure the autocorrelation function of a periodic signal; and Jakowatz (3), the autocorrelation function of an aperiodic signal. In this report we shall be concerned with measurements involving random processes. In performing experiments of this nature, it is desirable to compare the measured correlation functions with the correlation functions computed analytically. This means that the random processes must be so chosen that their correlation functions can be computed. The easiest way to do this is to put white noise through two known linear systems because the crosscorrelation function of the outputs is determined by the impulse responses of the systems. In practice, it is necessary that the noise source have a spectrum that is essentially flat only over the bandwidth of the linear systems.

2. Description of Equipment

The measurement of any first-order correlation function by the orthogonal expansion method requires a group of linear systems whose impulse responses constitute a set of orthogonal functions, at least one multiplier, at least one integrator or averaging network, and a noise source. In these experiments only one multiplier and averaging network was used, so it was necessary to measure the coefficients one at a time. The Laguerre functions were chosen as the orthogonal set, primarily because of the simplicity with which they can be synthesized. Treated as system functions, the transforms

of the Laguerre functions $\{L_n(\omega)\}$ have the form

$$L_n(\omega) = \frac{\sqrt{2p}}{p + j\omega}\left(\frac{p - j\omega}{p + j\omega}\right)^n \tag{5}$$

Since the first term is common to all members of the set, we can make the network representing it the first section of the synthesis system shown in Fig. VIII-3. In cascade with this section is a phase-shift chain of identical sections (4).

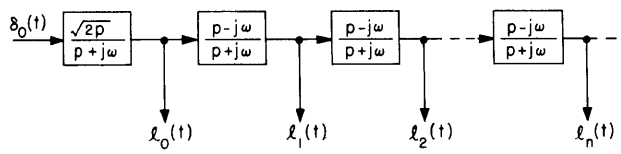


Fig. VIII-3. System for synthesizing the Laguerre functions.

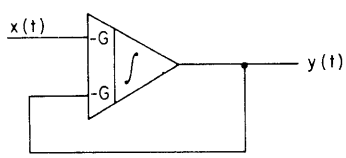


Fig. VIII-4. Computer representation of an RC lowpass filter.

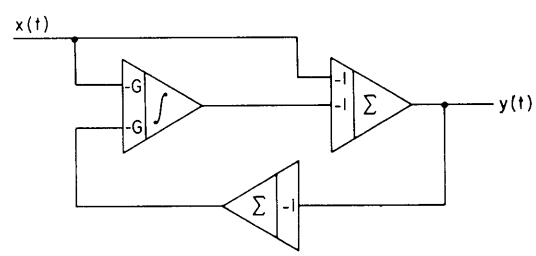


Fig. VIII-5. Computer representation of a phase-shift network.

For convenience, the Laguerre functions were synthesized on an analog computer, instead of building the appropriate networks with standard components. With operational amplifiers, all of the necessary transfer functions can be realized very easily. The form of the first section is obtained through the system shown in Fig. VIII-4. The transfer function of this system is

$$H(\omega) = \frac{G}{G + j\omega} \tag{6}$$

Observe that this differs from the desired system function by a factor of $(2/G)^{1/2}$. However, this is merely a scaling factor that can be accounted for when the measured data are scaled.

The remaining sections are each realized through the system shown in Fig. VIII-5. Writing the equations for this system, we see that

$$Y(\omega) = -\left[X(\omega) - \frac{G}{j\omega}X(\omega) + \frac{G}{j\omega}Y(\omega)\right]$$

$$Y(\omega)\left[\frac{G}{j\omega}+1\right] = X(\omega)\left[\frac{G}{j\omega}-1\right]$$

and finally

$$\frac{Y(\omega)}{X(\omega)} = H(\omega) = \frac{G - j\omega}{G + j\omega} \tag{7}$$

Since the measurements were to be made for random processes, the final piece of equipment was a noise generator. The noise source consisted of a 6D4 gas tube set in a dc magnetic field. The noise was then amplified and coupled out through a cathode follower.

3. Measurement of the Noise Spectrum

Because the random processes were to be obtained by passing white noise through linear filters, it was important for the spectrum of the noise source to be essentially constant over the bandwidth of the linear filters. Also, the power density had to be known before the measured data could be scaled for comparison with the analytic results. Therefore, before measuring any correlation functions, measurements were made to determine the spectrum of the noise source.

A quick and easy way to see if the noise is essentially white is to pass the noise through two orthogonal filters, multiply the outputs, and integrate. If the noise is white, or effectively white, the output of the integrator will stay finite. For example, consider the system shown in Fig. VIII-6, in which $\theta_1(t)$ and $\theta_2(t)$ are orthogonal functions; that is, $\int_0^\infty \theta_1(t)\,\theta_2(t)\,dt = 0$. For this system, the output is given by

$$\begin{aligned}
\overline{y(t)} &= \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)\,g(t)\,dt \\
&= \lim_{T\to\infty} \frac{1}{T}\int_0^T dt \int_0^\infty \theta_1(\sigma)\,n(t-\sigma)\,d\sigma \int_0^\infty \theta_2(\lambda)\,n(t-\lambda)\,d\lambda \\
&= \int_0^\infty \theta_1(\sigma)\,d\sigma \int_0^\infty \theta_2(\lambda)\,d\lambda \lim_{T\to\infty} \frac{1}{T}\int_0^T n(t-\sigma)\,n(t-\lambda)\,d\sigma d\lambda \\
&= \int_0^\infty \theta_1(\sigma)\,d\sigma \int_0^\infty \theta_2(\lambda)\,\phi_{nn}(\sigma-\lambda)\,d\lambda
\end{aligned} \tag{8}$$

where $\phi_{nn}(\tau)$ is the autocorrelation function of the noise source. If the noise is white with autocorrelation function $N_o\delta_o(\tau)$, where $\delta_o(\tau)$ is the Dirac delta function, Eq. 8 reduces to

$$\overline{y(t)} = N_o \int_0^\infty \theta_1(\sigma)\, \theta_2(\sigma)\, d\sigma = 0 \tag{9}$$

For the average to be zero, it is sufficient for the integral from zero to T to grow at a rate less than T. If this integral stays finite, we are assured that the average will be zero, so that in practice it is generally only necessary to integrate.

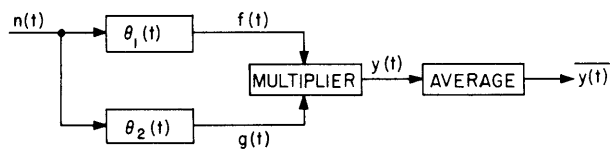


Fig. VIII-6. System for testing a noise source.

While it is true that the average will be zero for white noise inputs, we cannot say that the noise is white when such a test is made with only one pair of orthogonal filters. For such a measurement, there is the possibility that the autocorrelation functions of the noise and the orthogonal filters are related in such a way that the average is zero when the noise is not white. However, if the same measurements are made by using

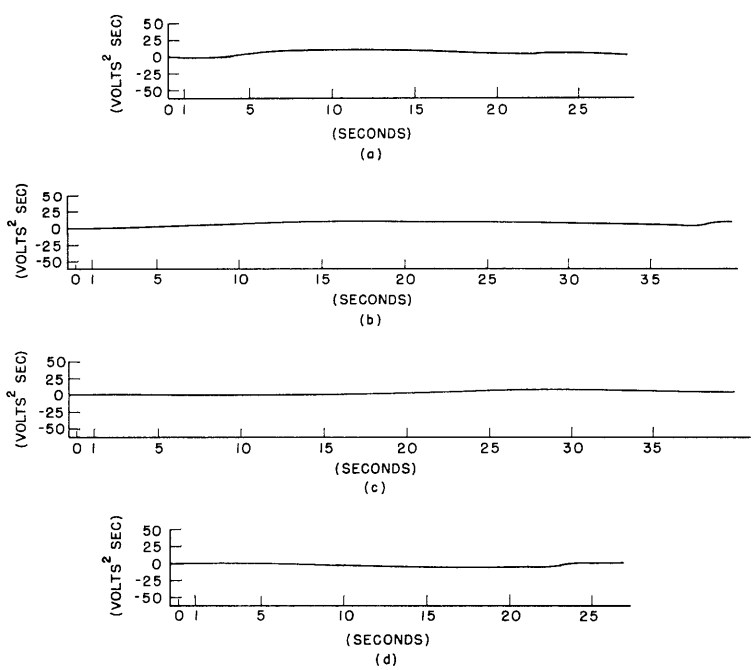


Fig. VIII-7. Output of the integrator in Fig. VIII-6 for: (a) $\theta_1(t) = \ell_o(t)$ and $\theta_2(t) = \ell_1(t)$; (b) $\theta_1(t) = \ell_o(t)$ and $\theta_2(t) = \ell_4(t)$; (c) $\theta_1(t) = \ell_3(t)$ and $\theta_2(t) = \ell_5(t)$; (d) $\theta_1(t) = \ell_2(t)$ and $\theta_2(t) = \ell_4(t)$.

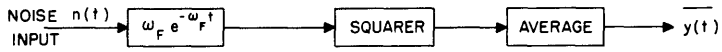


Fig. VIII-8. System for measuring power density spectrum.

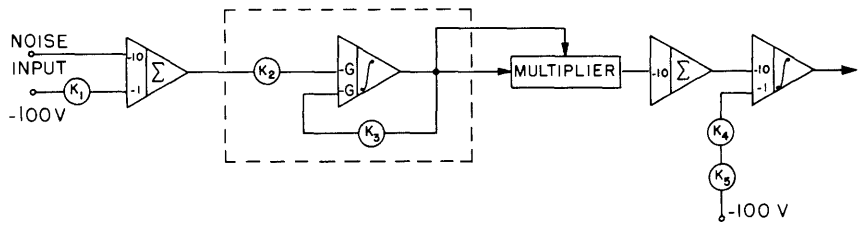


Fig. VIII-9. Computer program for measuring power density spectrum.

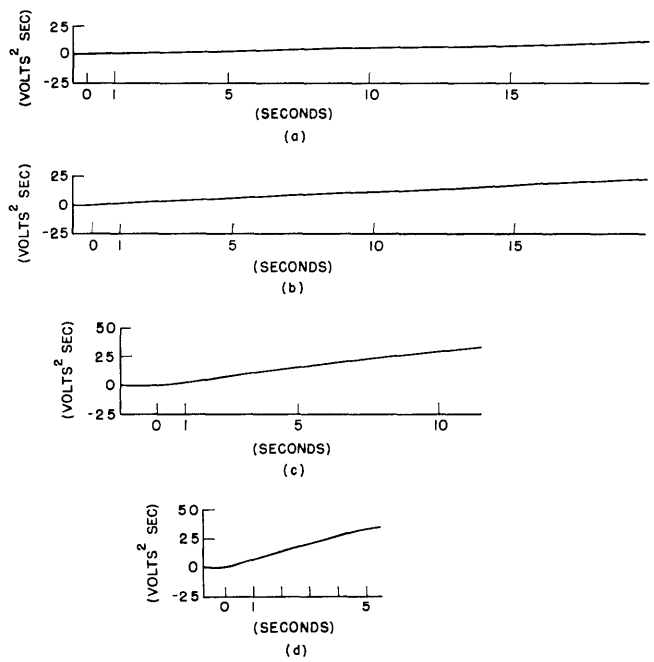


Fig. VIII-10. Measurements of the power density spectrum of the noise source. (a) ω_F = 25 rad/sec. (b) ω_F = 50 rad/sec. (c) ω_F = 100 rad/sec. (d) ω_F = 200 rad/sec.

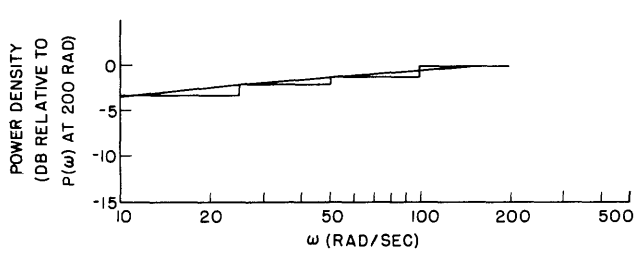


Fig. VIII-11. Measured power density spectrum of the noise generator.

several different pairs of orthogonal filters and the average in each case is zero, we can say with some confidence that the noise is white over the bandwidth of the linear systems.

The Laguerre functions were used as the orthogonal filters to make this test. A cutoff frequency of 100 rad/sec was chosen for these filters; that is, the parameter p in Eq. 5 was 100. Measurements were made with 4 different pairs of Laguerre filters and the output of the integrator is shown in Fig. VIII-7 for each pair. In every case the output never exceeded 12 volts^2 sec, so the average does approach zero as the averaging time increases. Therefore, we can expect the noise to be white at least for the frequency range 0-100 rad/sec.

Measurements of the noise spectrum were made with the use of the system shown in Fig. VIII-8, which is essentially the same method as that discussed by Lee (4). The noise voltage is applied to an RC lowpass filter with cutoff frequency ω_F, and the output of the filter is then squared and averaged. This average is given by

$$\overline{y(t)} = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt \int_0^\infty \omega_F\, e^{-\omega_F \sigma}\, n(t-\sigma)\, d\sigma \int_0^\infty \omega_F\, e^{-\omega_F \lambda}\, n(t-\lambda)\, d\lambda$$

$$= \omega_F^2 \int_0^\infty \int_0^\infty e^{-\omega_F(\sigma+\lambda)}\, \phi_{nn}(\sigma-\lambda)\, d\sigma d\lambda \tag{10}$$

From these measurements, the noise was assumed to be effectively white noise with autocorrelation function $N_o \delta_o(\tau)$. Substituting this function in Eq. 10 and carrying out the integration, we obtain the average

$$\overline{y(t)} = \frac{N_o}{2}\, \omega_F \tag{11}$$

The actual measurements were made by synthesizing the lowpass filter on the analog computer, and a multiplier was used in place of a squarer. The complete system is shown in Fig. VIII-9.

Measurements were made for cutoff frequencies of 25, 50, 100, and 200 rad/sec. Several records of the output of the integrator were obtained for each cutoff frequency, and a typical set is shown in Fig. VIII-10. The mean of the measured values was used to obtain a step-function approximation to the power density spectrum. The result is plotted in Fig. VIII-11 in decibels relative to the power density in the range 100-200 rad/sec. From this curve we observe that the spectrum is not flat, but drops off at low frequencies. Nevertheless, the measurements made with orthogonal filters indicate that the noise can be considered white without introducing appreciable error. The value of N_o for a cutoff frequency of 100 rad/sec was found to be 1.05×10^{-3} volts^2.

4. Correlation Function of Two Random Processes

In this experiment the crosscorrelation function of two random processes was measured for positive values of τ. Specifically, the following problem was considered.

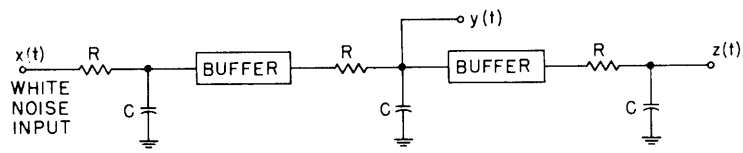


Fig. VIII-12. System for obtaining random processes with a given crosscorrelation function from white Gaussian noise.

Suppose we have the system shown in Fig. VIII-12. The input to this system is white noise with autocorrelation function $\delta_o(\tau)$, and the crosscorrelation function of y(t) and z(t) is to be determined. For this system we can write

$$y(t) = \int_{-\infty}^{\infty} h_2(\sigma)\, x(t-\sigma)\, d\sigma \tag{12}$$

and

$$z(t) = \int_{-\infty}^{\infty} h_3(\sigma)\, x(t-\sigma)\, d\sigma \tag{13}$$

where $h_2(t)$ characterizes the system containing two RC lowpass filters in cascade, and $h_3(t)$ characterizes the total system of three cascaded RC filters. These impulse responses are given by the relations

$$h_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{\frac{1}{RC}}{\frac{1}{RC} + j\omega}\right]^2 e^{-j\omega t}\, d\omega = \left(\frac{1}{RC}\right)^2 t\, e^{-t/RC} \tag{14}$$

and

$$h_3(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{\frac{1}{RC}}{\frac{1}{RC} + j\omega}\right]^3 e^{-j\omega t}\, d\omega = \frac{1}{2}\left(\frac{1}{RC}\right)^3 t^2\, e^{-t/RC} \tag{15}$$

The crosscorrelation function of y(t) and z(t) is, then,

$$\phi_{yz}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \int_{-\infty}^{\infty} h_2(\sigma_1)\, x(t-\sigma_1)\, d\sigma_1 \int_{-\infty}^{\infty} h_3(\sigma_2)\, x(t-\sigma_2+\tau)\, d\sigma_2$$

$$= \int_{-\infty}^{\infty} h_2(\sigma_1)\, d\sigma_1 \int_{-\infty}^{\infty} h_3(\sigma_2)\, d\sigma_2 \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t-\sigma_1)\, x(t+\tau-\sigma_2)\, dt$$

$$= \int_{-\infty}^{\infty} h_2(\sigma_1)\, d\sigma_1 \int_{-\infty}^{\infty} h_3(\sigma_2)\, \phi_{xx}(\tau+\sigma_1-\sigma_2)\, d\sigma_2 \tag{16}$$

where $\phi_{xx}(\tau)$ is the autocorrelation function of the white noise input. If we substitute the expression for this autocorrelation function in Eq. 16 and integrating over σ_2, the correlation function becomes

$$\phi_{yz}(\tau) = \int_{-\infty}^{\infty} h_2(\sigma_1)\, h_3(\sigma_1+\tau)\, d\sigma_1 \tag{17}$$

Substituting Eqs. 14 and 15 in Eq. 17 and performing the integration, we obtain the expression for the correlation function:

$$\phi_{yz}(\tau) = \begin{cases} \frac{1}{8RC}\left\{\left(\frac{\tau}{RC}\right)^2 + 2\left(\frac{\tau}{RC}\right) + 3/2\right\} e^{-\tau/RC} \\ \frac{1}{8RC}\left\{\frac{3}{2} - \left(\frac{\tau}{RC}\right)\right\} e^{-\tau/RC} \end{cases} \tag{18}$$

In order to measure $\phi_{yz}(\tau)$ by the orthogonal expansion method, we must choose a set with which to make the expansion. For this, the Laguerre functions were used, with the parameter p in Eq. 5 equal to 1/RC. This was done to obtain a rapid convergence. In fact, with this choice the correlation function $\phi_{yz}(\tau)$ can be represented exactly by the first three members of the Laguerre set for $\tau \geqslant 0$, and by the first two members for $\tau < 0$. For illustrative purposes, we shall consider only the expansion of $\phi_{yz}(\tau)$ for positive values of τ. The coefficients for this expansion can be computed with the use of Eq. 4, and the results are

$$\begin{aligned} a_o &= \frac{3}{16}\left(\frac{2}{RC}\right)^{1/2} \\ a_1 &= \frac{1}{8}\left(\frac{2}{RC}\right)^{1/2} \\ a_3 &= \frac{1}{32}\left(\frac{2}{RC}\right)^{1/2} \\ a_i &= 0 \qquad i > 2 \end{aligned} \tag{19}$$

Note that the analytic solutions contain the parameter RC. For a better comparison, the experimental results were also expressed in terms of this parameter. This was

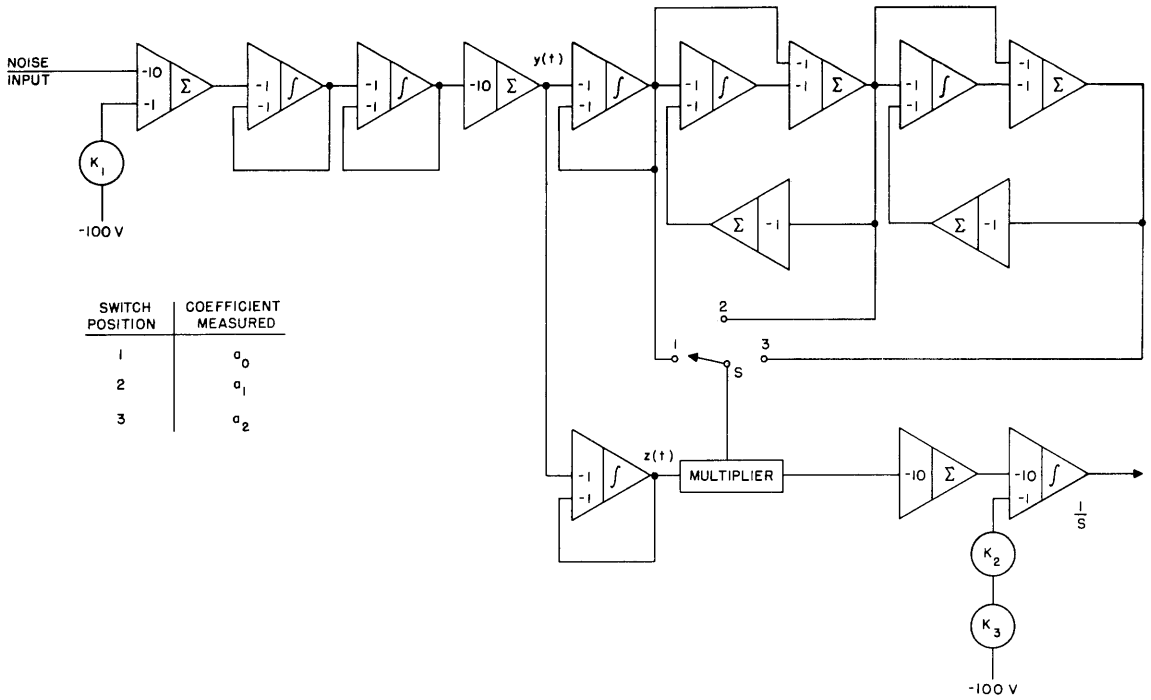


Fig. VIII-13. Computer program for measuring the coefficients of the crosscorrelation function of the random processes in Fig. VIII-12. (All integrators have transfer function 100/S, unless otherwise noted.)

accomplished by making the measurements for a convenient value of the product RC, and then converting the results to a form containing this parameter. The particular RC product used was 0.01, and the complete computer program for measuring the coefficients is shown in Fig. VIII-13.

The measurements for each coefficient were made several times and a typical set is shown in Fig. VIII-14. In order to obtain the desired coefficients, the measured averages had to be scaled according to the constants involved in making the measurements. The measured coefficients were found to be random variables as expected, and the mean and variance of the measurements are given in Table VIII-1 along with the computed coefficients for an RC product of 0.01 and an averaging time of 2.5 seconds.

From previous work it is known that for the averaging system used in the measurements, the mean value of the measured coefficients should be the true coefficient. From the values in Table VIII-1 it is seen that on a percentage basis the difference between these values is approximately 13 per cent for a_0 and 7 per cent for a_1 and a_2.

By using several different values of the measured coefficients, the correlation function was synthesized by means of the computer program illustrated

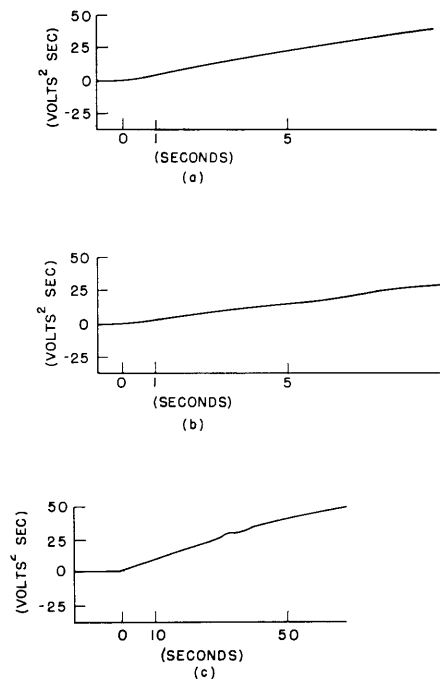


Fig. VIII-14. Coefficient measurements for the correlation function of filtered noise. Integrator output for measurement of: (a) a_o; (b) a_1; (c) a_2.

in Fig. VIII-15. Several comments should be made concerning this system. The program includes all of the first six Laguerre functions, even though only three were used in this experiment. However, all six will be used later, so provisions were made for the six Laguerre functions to be available with both polarities. For plotting purposes, an RC product of 1 was used. The input RC section when excited with a step function has the same effect as if the voltage transfer function of this section were $1/(1+j\omega)$ and an impulse is applied to the network. The step excitation was chosen because it is easier to obtain a good approximation to a step than to an impulse.

Table VIII-1.

Coefficient	Measured Values		Computed Value
	Mean	Variance	
a_o	2.35	0.09	2.65
a_1	1.64	0.023	1.77
a_2	0.41	0.022	0.442

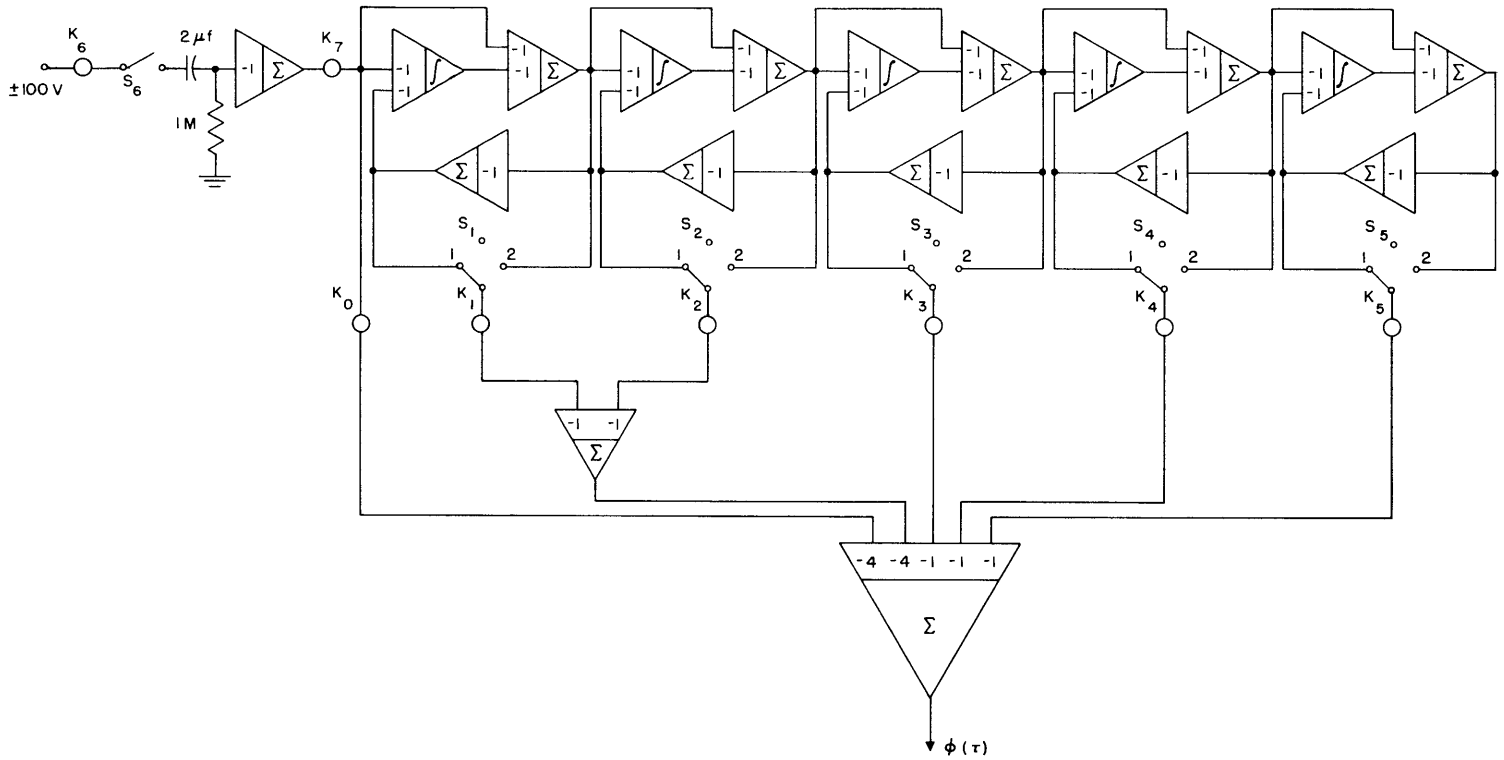


Fig. VIII-15. System for synthesizing the correlation functions. (All integrators have a transfer function, 1/S. The gain of K_7 is 0.14 for all synthesis programs.)

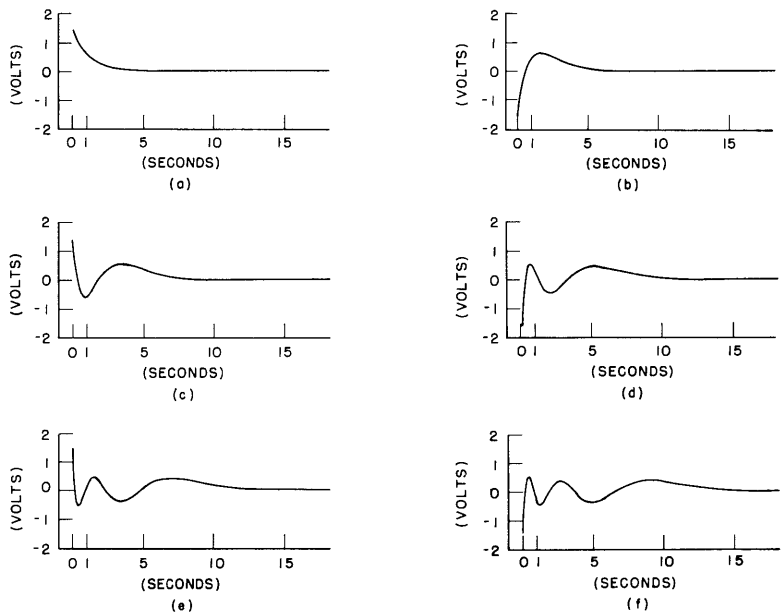


Fig. VIII-16. The first six Laguerre functions used to synthesize correlation functions. (a)-(f): $\ell_0(t)$; $\ell_1(t)$; $\ell_2(t)$; $\ell_3(t)$; $\ell_4(t)$; $\ell_5(t)$.

The first six Laguerre functions are shown in Fig. VIII-16 for reference. With the use of the measured coefficients, the correlation function was synthesized, and some of the results are shown in Fig. VIII-17. Since the problem was scaled for plotting, axes have been added to each record to indicate the true scales. The results shown in Fig. VIII-17a and 17b were obtained with some of the various measured coefficients. In Fig. VIII-17c, the correlation function is plotted by using the mean values of the measured coefficients. In order to illustrate what can happen when an expansion is truncated, Fig. VIII-17d shows the expansion with only two members of the set used. Finally, the true correlation function is shown in Fig. VIII-17e. In all of these, including the truncated expansion, the general shape is very much the same. The peak occurs in almost the same place, and its height does not change appreciably. The most significant difference is the value at the origin.

5. Extraction of a Pulse from Noise

One technique for detecting a known signal in the presence of noise is crosscorrelating the signal plus noise with a replica of the signal. If the signal is denoted s(t) and the noise n(t), this crosscorrelation function is given by

$$\phi(\tau) = \overline{[s(t)+n(t)]\, s(t+\tau)} = \overline{s(t)\, s(t+\tau)} + \overline{n(t)\, s(t+\tau)} = \phi_{ss}(\tau) + \overline{n(t)}\;\overline{s(t)}$$

If the noise has zero mean, this reduces to

$$\phi(\tau) = \phi_{ss}(\tau) \tag{20}$$

That is, the crosscorrelation function is simply the autocorrelation function of the signal.

For this experiment the signal was found to be a 1-volt pulse of 1-sec duration, and the noise was white noise whose peak-to-peak voltage was approximately 25 volts. The crosscorrelation function is, then, the autocorrelation function of the pulse which is known to be

$$\phi_{ss}(\tau) = \begin{cases} 1 - |\tau| & |\tau| \leq 1 \\ 0 & |\tau| > 1 \end{cases} \tag{21}$$

Thus we can compute the coefficients for comparison with the measured values. Again, the Laguerre functions were chosen as the orthogonal set, and the first six members

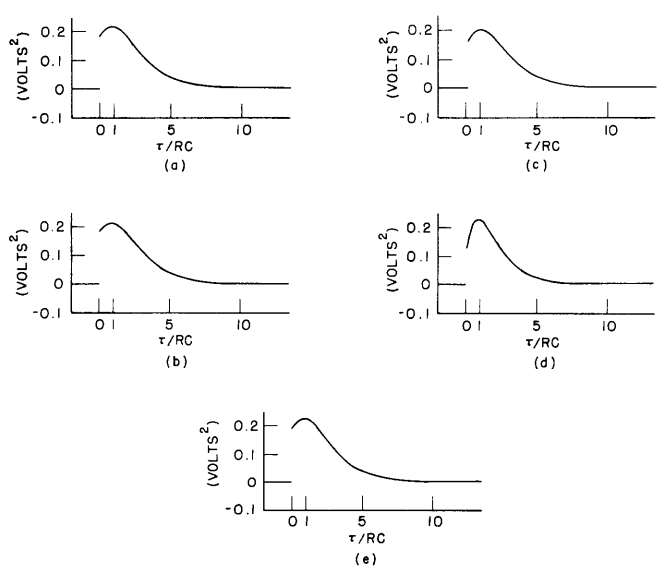
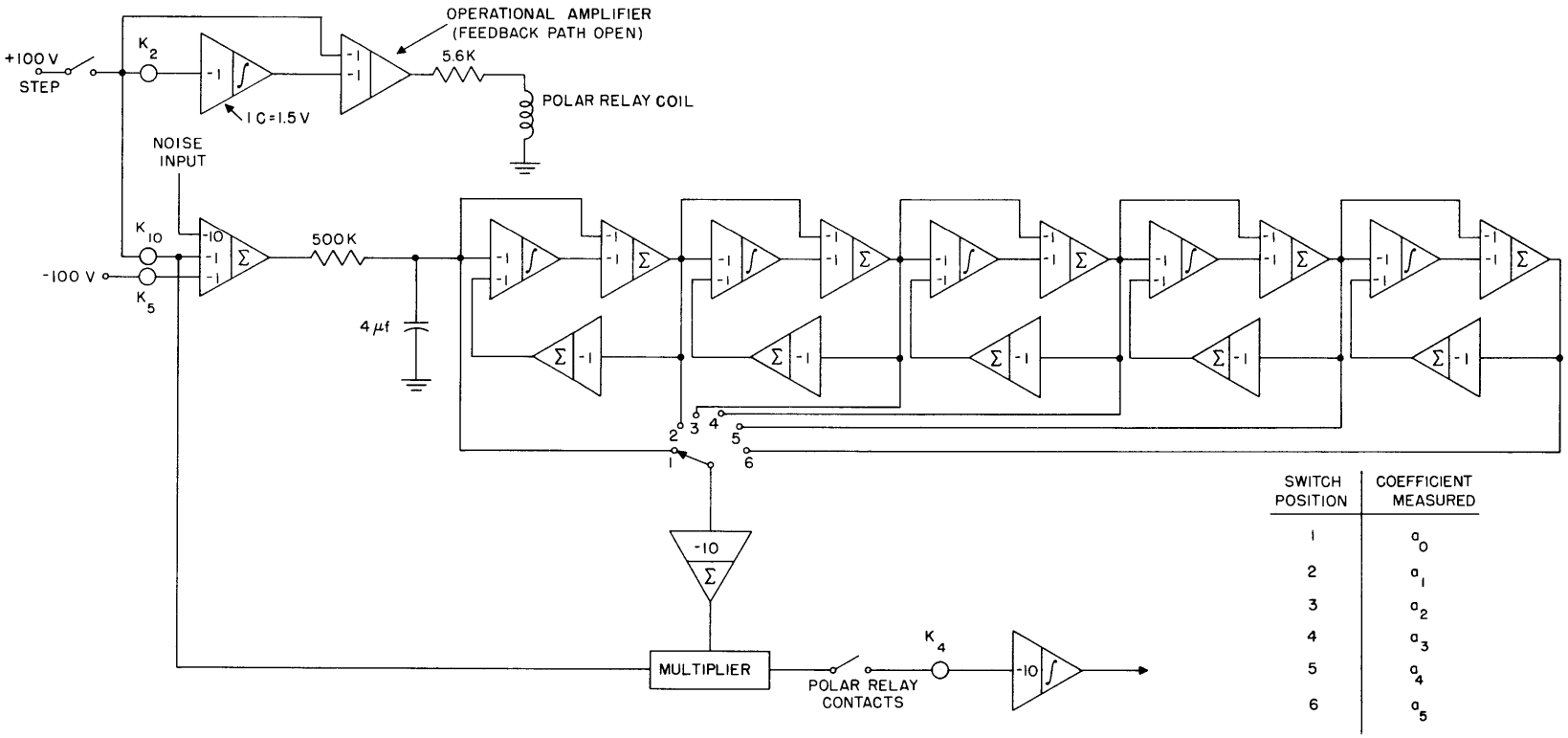


Fig. VIII-17. Measured correlation function of filtered noise normalized to $RC\phi_{yz}(\tau/RC)$. (a) $a_o = 0.25$; $a_1 = 0.169$; $a_2 = 0.043$. (b) $a_o = 0.25$; $a_1 = 0.165$; $a_2 = 0.042$. (c) $a_o = 0.235$; $a_1 = 0.166$; $a_2 = 0.041$. (d) $a_o = 0.242$; $a_1 = 0.166$; $a_2 = 0.$ (e) $a_o = 0.265$; $a_1 = 0.177$; $a_2 = 0.044$.

were used in the expansion. In order for the expansion to converge rapidly, the parameter p in Eq. 5 was chosen to be 1. For this set, direct computation shows that the coefficients are:



SWITCH POSITION	COEFFICIENT MEASURED
1	a_0
2	a_1
3	a_2
4	a_3
5	a_4
6	a_5

Fig. VIII-18. System for measuring the coefficients of the correlation function of a pulse in noise. (All integrators have transfer function, 1/S.)

$a_o = 0.52$ $\qquad$ $a_3 = 0.0156$

$a_1 = -0.229$ $\qquad$ $a_4 = -0.0382$

$a_2 = 0.068$ $\qquad$ $a_5 = 0.048$

In programming this problem for the computer we found that there were not enough amplifiers and integrators on the computer for the complete simulation. The first Laguerre function was therefore built with standard components. The complete system for measuring the coefficients is shown in Fig. VIII-18. The values of R and C used in the first Laguerre function were necessary to compensate for the loading caused by the input impedance of the operational amplifiers.

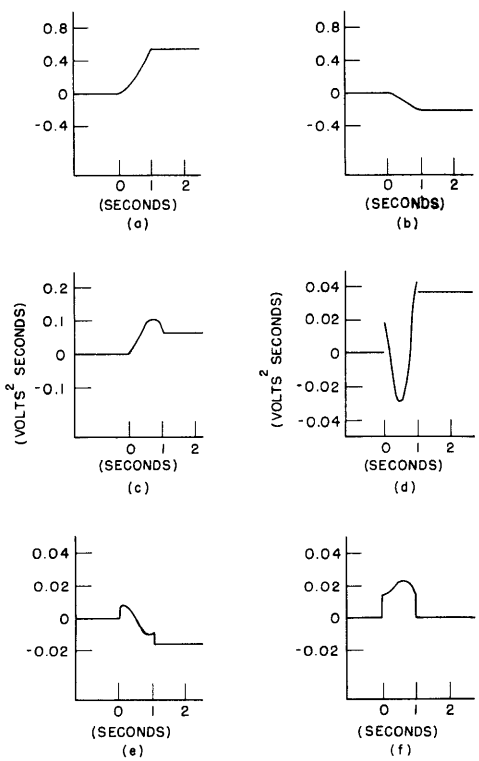


Fig. VIII-19. Measurement of the coefficients of the correlation function of a pulse in noise: (a) a_o; (b) a_1; (c) a_2; (d) a_3; (e) a_4; (f) a_5.

Another aspect of the program is the manner in which the signal, or pulse, was obtained. In the ideal measuring system, one input to the multiplier is nonzero only when the pulse is present. The output of the multiplier then behaves in the same manner. The multiplier used in these measurements, however, had a small nonzero output with one input grounded. To prevent this from affecting the measurements, the integrator was placed in the circuit only when the signal was present. This was achieved by an arrangement that made use of a bipolar relay. For example, with the stepping

Table VIII-2.

Coefficient	Measured Values		True Value
	Mean	Variance	
a_o	0.52	0.0021	0.52
a_1	-0.22	0.0005	-0.226
a_2	0.058	$6.3\ (10^{-6})$	0.068
a_3	0.0216	$8.7\ (10^{-6})$	0.0156
a_4	-0.017	.00012	-0.0382
a_5	0.004	$11\ (10^{-6})$	0.048

switch open, the initial condition on the integrator following potentiometer K_2 insures that the integrator is out of the circuit. When the step is applied, a pulse is started into the system and the integrator is placed in the circuit by the action of the high-gain amplifier. After a time determined by the gain of potentiometer K_2, the polarity of the input to the high-gain amplifier is reversed, and the integrator is removed from the circuit that simulates the end of a pulse.

Since the correlation function that is being measured is the correlation function of aperiodic signals, the coefficients are measured by integrating instead of averaging. Five measurements were made for each coefficient, and a typical set is shown in Fig. VIII-19. For the purpose of comparison, the mean and variance of the measured coefficients are listed in Table VIII-2, together with the true coefficients. It is seen that the more significant coefficients compare very well, while there is some discrepancy in the smaller values. This difference will not result in a significant change in the shape of the correlation function, since it is small compared with the first two terms.

By using these coefficients, the correlation function was synthesized with the use of the system shown in Fig VIII-15, and some of the results are shown in Fig. VIII-20. The curves in Fig. VIII-20a and 20b were made with some of the various measured coefficients used, and Fig. VIII-20e is the correlation function corresponding to the mean values of the measured coefficients. For comparison, the correlation function was synthesized by using the computed coefficients, and it is shown in Fig. VIII-20f. Finally, the effects of truncating the expansion are illustrated in Fig. VIII-20c and 20d. Only the first four Laguerre functions were used in Fig. VIII-20c; and in Fig. VIII-20d, only the first three were used. The most noticeable effect of truncation is that the correlation function decays less rapidly.

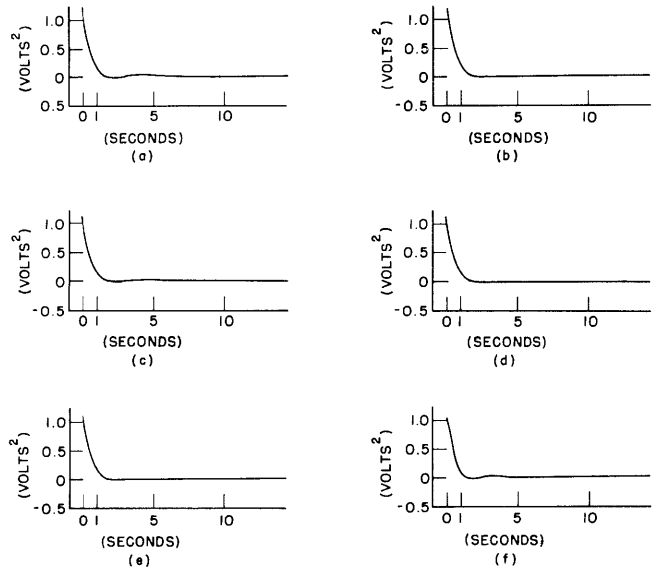


Fig. VIII-20. Correlation function of a pulse in noise with a pulse.

6. The Effects of Using More Members of a Set than Are Necessary

For the most part, one generally thinks of the errors that can be introduced by truncating an expansion and attempting to use as many members of a set as possible. There are cases in which the use of more and more terms in the expansion is not desirable. For example, if a correlation function can be represented exactly by N members of a set, then using more than N terms increases the integral-square error and contributes no additional information concerning the true correlation function. The experiment described in this section was designed to illustrate the effects of using more terms than are necessary.

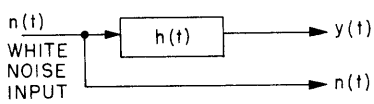
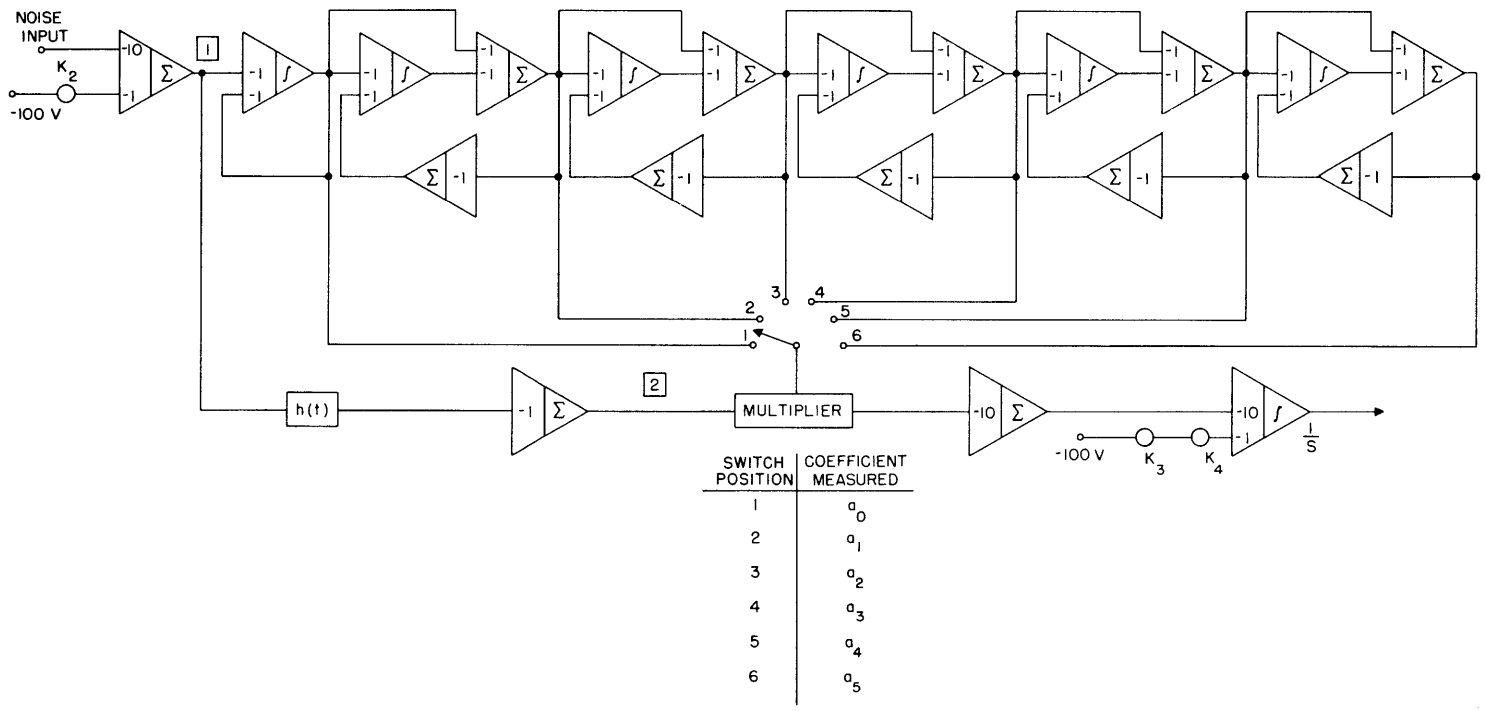


Fig. VIII-21. Linear system with white noise as its input.

This experiment consists of measuring the impulse response of an RC lowpass filter with white noise present. By choosing the Laguerre functions properly, this impulse can be expanded exactly by the first Laguerre function. All of the first six members of the set were used, however, so that five of the terms were unnecessary. Also, the impulse response of any realizable linear system is zero for negative time. It is of interest to see how well zero can be expanded, since it is possible for two random processes to be uncorrelated.

The impulse response of a linear system can be measured by putting white noise into the system and crosscorrelating the input with the output. For example, consider the



SWITCH POSITION	COEFFICIENT MEASURED
1	a_0
2	a_1
3	a_2
4	a_3
5	a_4
6	a_5

Fig. VIII-22. System for measuring the coefficients for the expansion of the impulse response of a linear system.

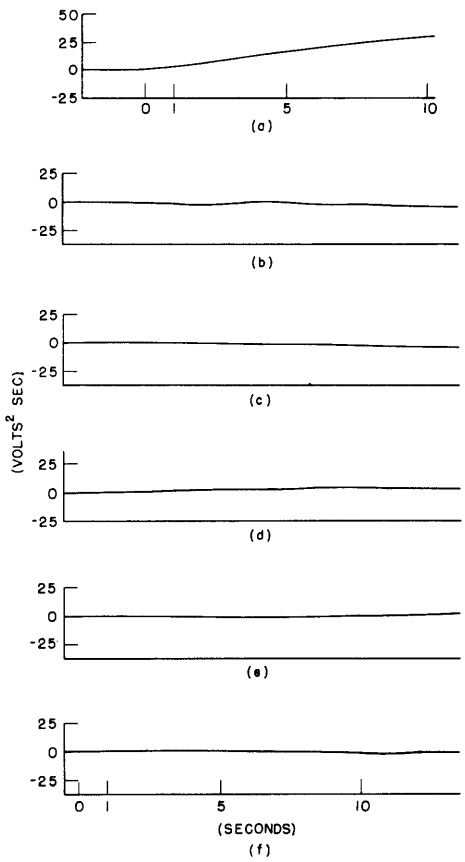


Fig. VIII-23. Measurement of coefficients of impulse response for $\tau \geqslant 0$. Integrator output for measurement of: (a) a_o; (b) a_1; (c) a_2; (d) a_3; (e) a_4; (f) a_5.

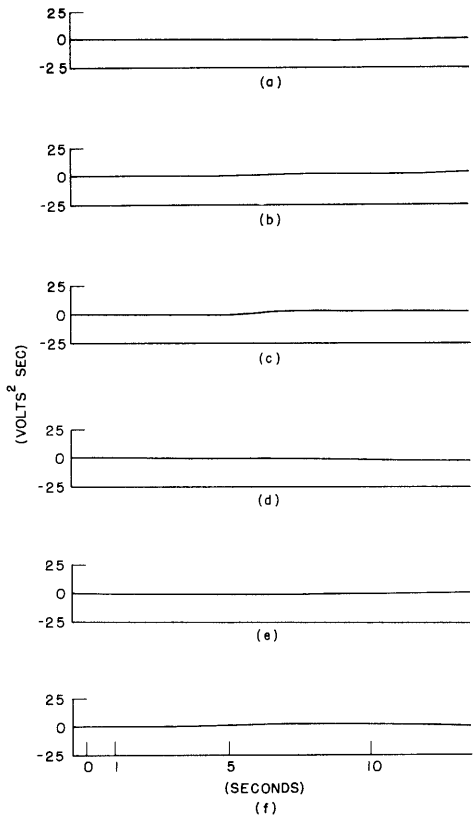


Fig. VIII-24. Measurement of coefficients of impulse response for $\tau < 0$. (a)-(f) Integrator output for measurement of a_o, a_1, a_2, a_3, a_4, a_5.

system shown in Fig. VIII-21. Let n(t) be white noise whose autocorrelation function is $N_o\delta_o(\tau)$. The crosscorrelation function of the input with the output is given by

$$\phi_{ny}(\tau) = \overline{n(t)\, y(t+\tau)} = \lim_{T\to\infty} \frac{1}{T}\int_0^T n(t)\, dt \int_{-\infty}^{\infty} h(\sigma)\, n(t+\tau-\sigma)\, d\sigma$$

$$= \int_{-\infty}^{\infty} h(\sigma)\, d\sigma \lim_{T\to\infty} \frac{1}{T}\int_0^T n(t)\, n(t+\tau-\sigma)\, dt$$

$$= \int_{-\infty}^{\infty} h(\sigma)\, \phi_{nn}(\tau-\sigma)\, d\sigma = N_o \int_{-\infty}^{\infty} h(\sigma)\, \delta_o(\tau-\sigma)\, d\sigma$$

$$= N_o h(\tau) \tag{23}$$

The impulse response is given by the relation

$$h(\tau) = \frac{1}{N_o} \phi_{ny}(\tau) \tag{24}$$

The linear system chosen for this experiment had an impulse response equal to $100\, e^{-100t}$, and the Laguerre functions used for the expansion had the same cutoff frequency. Therefore, for positive values of τ, the coefficient of the first Laguerre function should be 7.07, and all others zero. For negative values of τ, all of the coefficients should be zero.

The system used to measure the coefficients is shown in Fig. VIII-22. As it appears, the coefficients of the expansion for positive argument are being determined. For negative argument, the coefficients are measured with the same system but with the leads marked $\boxed{1}$ and $\boxed{2}$ interchanged.

As in the previous experiments, the various coefficients were measured several times, and examples of these measurements are given in Figs. VIII-23 and VIII-24. The measurements for the expansion of $h(+\tau)$ are indicated in Fig. VIII-23, and those for the expansion of $h(-\tau)$ in Fig. VIII-24. Again, the measured coefficients were random variables, and the mean values of these measurements are given in Table VIII-3 for an averaging time of 10 seconds. For this averaging time, the coefficients that should have been zero were found to have a finite nonzero value. However, in each of these measurements the output of the integrator never exceeded 5 volts^2 seconds. This means that even though these coefficients were not zero, their values approach zero as the averaging time increases.

The correlation function was synthesized by using the coefficients given in Table VIII-3 and the system shown in Fig. VIII-15. The results of this synthesis are

Table VIII-3.

Coefficient	Mean of Measured Values	
	For $\tau \geq 0$	For $\tau < 0$
a_o	7.35	0.125
a_1	-0.40	0.55
a_2	-0.775	0.25
a_3	0.325	-0.30
a_4	0.075	-0.113
a_5	-0.275	0.215

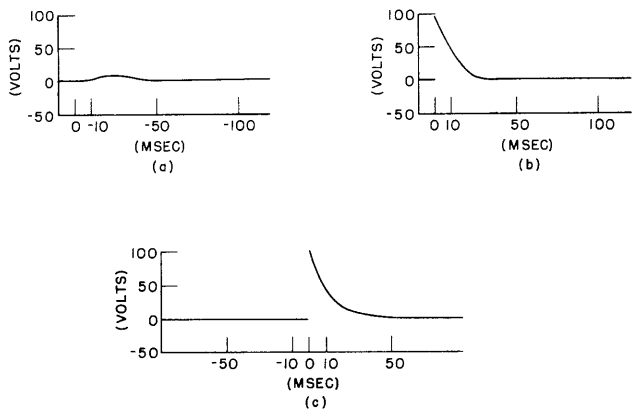


Fig. VIII-25. Impulse response measured with white noise. (a) Impulse response for $\tau < 0$. (b) Impulse response for $\tau \geqslant 0$. (c) True impulse response.

shown in Fig. VIII-25, and the scales given take into account the scaling used in the computer.

The function in Fig. VIII-25a is the measured impulse response for $\tau < 0$. Since the coefficients for this expansion were not zero, there is some response for negative time. The impulse response measured for positive time is given in Fig. VIII-25b. These results illustrate what can happen when more members of a set are used than are necessary to represent the correlation function. In this case, the extra members caused some distortion after approximately four time constants. By using a much longer averaging time, this distortion can be made arbitrarily small. For comparison, the true impulse response is included in Fig. VIII-25c.

As more members of a set are used to determine a correlation function for a fixed averaging time, the integral-square error resulting from finite time averaging will increase. This means that if we attempt to obtain a better estimate of the correlation function by using many terms in the expansion, the integral-square error can become extremely large. One way in which this can be prevented is by increasing the averaging time in proportion to the number of terms added to the expansion. This suggests that a rule of thumb to follow is to maintain the ratio N/T constant, where N is the number of terms used, and T is the averaging time.

R. A. Bruce

References

1. M. Schetzen, Measurement of correlation functions, Quarterly Progress Report No. 59, Research Laboratory of Electronics, M.I.T., April 15, 1960, pp. 93-98.

2. D. G. Lampard, A New Method of Determining Correlation Functions of Stationary Time Series, Monograph No. 104, Proc. Inst. Elect. Engrs. (London), Part IV, Aug. 16, 1954.

3. C. V. Jakowatz, Simulation of orthogonal functions, Proc. Symposium on Active Networks and Feedback Systems (Polytechnic Press, Polytechnic Institute of Brooklyn, New York, 1960).

4. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960).

C. MUTUAL INFORMATION IN SIGNAL REPRESENTATION

The reconstruction of a message variable x from an observed signal z is a typical problem in statistical communication theory. Wiener (1), and others, have shown that x can be estimated by a polynomial (or some other function) in the variables $z_1, \ldots, z_n$, which have been selected to provide an adequate linear representation of z and its past. We suggest that a criterion of adequacy for this representation be obtained by comparing the average mutual information between x and the set $z_1, \ldots, z_n$ with the average mutual information between x and z. The outstanding feature of such a criterion is its independence of the polynomial (or other function) that comprises the nonlinear part of the estimator. This independence may simplify analysis in the representation problem, although it is not likely to provide maximum efficiency or economy. We shall consider some simple examples showing that the desired simplicity need not always be nullified by difficulties in computing average mutual information. In the rest of this report we assume that x and $z_1, \ldots, z_n$ possess a joint Gaussian distribution for any n, and that $z_1, \ldots, z_n$ is capable of representing z and its past arbitrarily well for sufficiently large n.

Let $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ be a set of Gaussian variables. If the first subset is called U, the second V, and the whole set U, V, the average mutual information (in natural units) between U and V is given by

$$I(U;V) = \frac{1}{2} \ln \frac{|\Lambda_U| \, |\Lambda_V|}{|\Lambda_{U,V}|} \tag{1}$$

where $|\Lambda_W|$ is the determinant of the covariance matrix Λ_W of variables in the set W (2). It is quite possible that Eq. 1 has appeared in published works on information theory; at any rate, the proof is a straightforward computation. In our case, the set U consists of x (we also denote the set containing only x by x), and V is the set $z_1, \ldots, z_n$, denoted Z_n. Under the further assumption that x and all z_i have the same variance, Eq. 1 reduces to

$$I(x;Z_n) = \frac{1}{2} \ln \frac{\left|P_{Z_n}\right|}{\left|P_{x,Z_n}\right|} \tag{2}$$

where P is a correlation matrix corresponding to Λ (3).

In certain classes of problems Eq. 2 may be further simplified. For example, suppose that the correlation coefficients depend only on "spacing," that is,

$$\rho_{ij} = f(|i-j|) \tag{3}$$

where ρ_{ij} is the correlation coefficient of z_i and z_j for $i > 0$, and ρ_{oj} is the correlation coefficient of x and z_j. It is clear that, in this case, $P_{x, Z_n} = P_{Z_{n+1}}$. If P_{Z_k} is put into triangular form (4) by the modified Gauss-Jordan reduction (rows not multiplied by a constant that would make diagonal elements one and would also multiply the determinant by that constant), and if the ith diagonal element so formed is called d_i, then $\left|P_{Z_k}\right| = d_1 \ldots d_k$. Thus, Eq. 2 becomes

$$I(x;Z_n) = \frac{1}{2} \ln \frac{1}{d_{n+1}} \tag{4}$$

and the calculation of average mutual information is reduced to the triangularization of a correlation matrix.

For a very simple example, let the correlation coefficients of Eq. 3 be specified by

$$f(k) = \begin{cases} 1, & k = 0 \\ \frac{1}{2}, & k = 1 \\ 0, & k > 1 \end{cases}$$

The ith diagonal element of the triangularized correlation matrix is readily found to be

$$d_i = \frac{i+1}{2i} \tag{6}$$

The limit of this expression for large i is

$$d_\infty = \frac{1}{2} \tag{7}$$

It is therefore easy to determine how well $z_1, \ldots, z_n$ represents z for the purpose of estimating x in terms of how close $I(x;Z_n)$ comes to its limiting value, or, how close $\frac{1}{2} \ln \frac{2n+2}{n+2}$ comes to $\frac{1}{2} \ln 2$.

Consider next the more general case, for which

$$f(k) = \begin{cases} 1, & k = 0 \\ c, & k = 1 \\ 0, & k > 1 \end{cases} \tag{8}$$

Notice that this agrees with the preceding example for $c = \frac{1}{2}$. If $|c| > \frac{1}{2}$, it will always be possible to find an i large enough so that $d_i < 0$, in which case the distribution is not Gaussian. This fact may be proved from a careful study of the recursion formula

$$d_{i+1} = 1 - \left(c^2 / d_i\right) \tag{9}$$

where $d_1 = 1$. There is no simple expression for d_i, but this is not a great practical difficulty. The important point is that we can find the limit expression,

$$d_\infty = \frac{1}{2} + \left(\frac{1}{4} - c^2\right)^{1/2} \tag{10}$$

where $|c| \leq \frac{1}{2}$. This agrees with Eq. 7 for $|c| = \frac{1}{2}$. When $c = 0$, Eq. 10 yields $d_\infty = 1$. Then Eq. 4 gives $I(x;Z_\infty) = 0$, which agrees with the fact that independent variables have zero mutual information.

Finally, let us suppose that

$$f(k) = \begin{cases} \frac{m-k}{m}, & 0 \leq k \leq m \\ 0, & k > m \end{cases} \tag{11}$$

where m and k are both integers. Notice that this agrees with Eq. 5 when $m = 2$. It is certainly possible to establish a recursive procedure for computing the d_i corresponding to a given m, although a direct formulation of d_i will probably be difficult. We conjecture that the limiting expression, for any m, is

$$d_\infty = \frac{1}{m} \tag{12}$$

The rigorous proof that the limit exists and is given by Eq. 12 has thus far eluded us, although a convincing information theory argument makes existence of the limit quite plausible.

A. D. Hause

References

1. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1958).

2. W. B. Davenport, Jr. and W. L. Root, An Introduction to the Theory of Random Signals and Noise (McGraw-Hill Book Company, Inc., New York, 1958), p. 152.

3. Ibid., loc. cit.

4. E. A. Guillemin, The Mathematics of Circuit Analysis (The Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1949), pp. 3-4.

D. POWER ABSORBED BY A NONLINEAR TWO-TERMINAL NETWORK WITH A WHITE GAUSSIAN INPUT

The average power absorbed by a two-terminal linear or nonlinear network, as depicted in Fig. VIII-26, is the average of the product of the voltage, e(t), and the

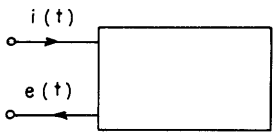


Fig. VIII-26. A two-terminal linear or nonlinear network.

current, i(t). The network is passive if, for any excitation, the average power absorbed is positive. That is, if

$$P = \overline{e(t)\ i(t)} \geqslant 0 \tag{1}$$

in which the bar indicates that the time average is taken. In terms of the crosscorrelation function, $\phi_{ie}(\tau) = \overline{i(t)\ e(t+\tau)}$, the average power absorbed can be expressed as

$$P = \phi_{ie}(0) = \int_{-\infty}^{\infty} \Phi_{ie}(\omega)\ d\omega \tag{2}$$

in which $\Phi_{ie}(\omega)$ is the Fourier transform of $\phi_{ie}(\tau)$:

$$\Phi_{ie}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{ie}(\tau)\ e^{-j\omega\tau}\ d\tau \tag{3}$$

Since $\phi_{ie}(\tau)$ is a real function of τ, we observe that the real part of $\Phi_{ie}(\omega)$, $\mathrm{Re}\{\Phi_{ie}(\omega)\}$, is an even function of ω, and the imaginary part of $\Phi_{ie}(\omega)$, $\mathrm{I_m}\{\Phi_{ie}(\omega)\}$, is an odd function of ω. Thus, Eq. 2 can be written as

$$P = 2 \int_{0}^{\infty} \mathrm{Re}\{\Phi_{ie}(\omega)\}\ d\omega \tag{4}$$

If the network is linear, then the current can be expressed as

$$i(t) = \int_{-\infty}^{\infty} k(\sigma)\ e(t-\sigma)\ d\sigma \tag{5}$$

in which k(t) is the current response to an impulse of voltage. That is,

$$K(\omega) = \int_{-\infty}^{\infty} k(t)\ e^{-j\omega t}\ dt \tag{6}$$

is the input admittance of the network. In terms of $K(\omega)$, we can express $\Phi_{ie}(\omega)$ as

$$\Phi_{ie}(\omega) = \Phi_{ee}(\omega)\ K(\omega) \tag{7}$$

in which $\Phi_{ee}(\omega)$ is the power density spectrum of the voltage, e(t). Thus, for a linear passive network,

$$\mathrm{Re}\{\Phi_{ie}(\omega)\} = \Phi_{ee}(\omega)\ \mathrm{Re}\{K(\omega)\} \geqslant 0 \tag{8}$$

for all ω. This result follows because $\Phi_{ee}(\omega)$ is a real, positive function of ω, and, for a passive network, $\mathrm{Re}\{K(\omega)\} \geq 0$. If the network is not linear, then it is possible that $\mathrm{Re}\{\Phi_{ie}(\omega)\} < 0$ over some range of frequencies; although from Eq. 4, the positive area must be larger than the negative area if the system is passive. In this report, we shall present some results that we have obtained for systems in which $\mathrm{Re}\{\Phi_{ie}(\omega)\} \geq 0$, for all ω. Systems for which this is not true will be discussed in a later report.

We note, first, that $\mathrm{Re}\{\Phi_{ie}(\omega)\}$ is the Fourier transform of the even part of $\phi_{ie}(\tau)$. Thus, since $\mathrm{Re}\{\Phi_{ie}(\omega)\} \geq 0$, the even part of $\phi_{ie}(\tau)$ is a function of the class P (1). Now, in terms of the orthogonal functionals described by Wiener (2), we can express the current in terms of the voltage as

$$i(t) = \sum_{n=0}^{\infty} G_n[h_n, e(t)] \tag{9}$$

If the voltage is a Gaussian white-noise process for which $\phi_{ee}(\tau) = K\mu(\tau)$, then (see Lee and Schetzen (3))

$$\phi_{ie}(\tau) = Kh_1(\tau) \tag{10}$$

and from Eq. 3,

$$\Phi_{ie}(\omega) = \frac{K}{2\pi} H_1(\omega) \tag{11}$$

Thus, from Eq. 4, for a Gaussian white-noise excitation, the power absorbed by a nonlinear network is determined solely by the real part of $H_1(\omega)$. Furthermore, if $\mathrm{Re}\{\Phi_{ie}(\omega)\} \geq 0$, we require that $\mathrm{Re}\{H_1(\omega)\} \geq 0$, for all ω. This means that, for such a condition, $H_1(\omega)$ must be realizable as the admittance of a linear passive network.

M. Schetzen

References

1. M. Schetzen, Fourier transforms of positive functions, Quarterly Progress Report No. 52, Research Laboratory of Electronics, M.I.T., Jan. 15, 1959, pp. 79-86.

2. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1958).

3. Y. W. Lee and M. Schetzen, Measurement of the kernels of a nonlinear system by crosscorrelation, Quarterly Progress Report No. 60, Research Laboratory of Electronics, M.I.T., Jan. 15, 1961, pp. 118-130.