2

CR 73356

AVAILABLE TO THE PUBLIC

Properties of Sums of Pseudo-Random Variables in Feedback Shift Registers

by

Timothy J. Healy
Assistant Professor
Department of Electrical Engineering
Santa Clara, California

Final Report Submitted to the Ames Research Center,

National Aeronautics and Space Administration

Ames-Santa Clara Omnibus Agreement

Projecc No. 008

Santa Clara, California

June 1, 1969

Table of Contents

Properties of Sums of Pseudo-Random Variables in Feedback Shift Registers

0.0 Introduction

It is well-known that the output of a binary feedback shift register, such as that shown in Figure 1, is a binary pseudo-random sequence.

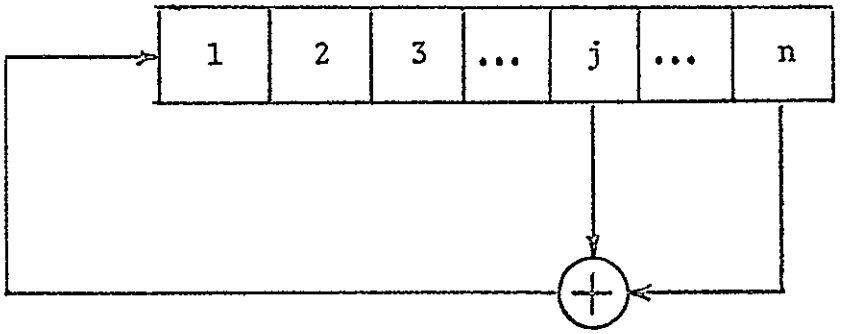


Figure 1. Binary Feedback Shift Register

The statistics of the output sequence have been studied in some detail.[1]-[9] Other researchers have studied the important delay or shifting property[10]-[13], the statistics of the output after filtering [14]-[17], the spectral density of the output[18], some practical applications[19]-[25], and finally a number of other closely related ideas[26]-[35].

This report is concerned with the statistics of the sum of the values in the shift register[36]-[38]. The problem is considered here from a number of viewpoints. We first consider in section 1 partial results obtained from a computer study. These results show a number of interesting

features. To some extent, sections 2 and 3 are intended to explain some of the features observed in section 1. The major conclusion from the computer study is that the generated distributions become extremely complex as the number of stages in the shift register and the degree of inter-dependence increase.

In section 2 we consider the problem of the sums of weighted outputs of the stages of the feedback shift register. A method of synthesizing pseudo-random distributions is developed. It is shown that several important distributions such as uniform, triangular, stairstep and bimodal, as well as an unlimited number of other less important distributions, are easily synthesized.

Section 3 is concerned with the sums of the outputs of the feedback shift register and various delayed or shifted outputs. The results of sections 2 and 3 are shown to be closely related.

In section 4 we study the problem of the infinite sum of weighted random variables. Some important relations with the problems of the first three sections are shown.

1. Computer-Generated Frequency Histograms

We are concerned in this section with the frequency histogram of the sum of values of outputs in the first M stages of a feedback shift register such as that shown in figure 1. The range on M is 0 to $2^n - 1$. Where M exceeds n, it is assumed that additional delay stages are added to the shift register outside the feedback loop. The frequencies of occurrence of different sums, over different values of M, were determined for n ranging from 3 to 11. The results are easily displayed in the form

of a "frequency histogram matrix," as seen for $n = 3$ in table 1 below.

S/M	0	1	2	3	4	5	6	7
0	7	3	1	0	0	0	0	0
1	0	4	4	3	1	0	0	0
2	0	0	2	3	3	2	0	0
3	0	0	0	1	3	4	4	0
4	0	0	0	0	0	1	3	7

Table 1. Frequency Histogram Matrix - $n = 3$

The frequency of sums S appears as a row under each value of M. For example, consider the case where we add two successive stage outputs at a time ($M = 2$). In a single complete period (length = $2^3 - 1 = 7$) the sum will be zero once, one four times and two twice.

Note that the matrix is symmetrical in that the right half repeats (see Table 1), in reverse order, the left half. To see why this is so, consider two columns of symmetry designated by M and $2^n - 1 - M$. Note that these columns cover a total number of stages equal to 2^{n-1} (their sum). Also, note that when we cover all $2^n - 1$ stages the number of 1's is 2^{n-1}. Hence, obtaining a sum S in column M is equivalent to obtaining a sum $2^{n-1} - S$ in column $2^n - 1 - M$. Thus, the matrix must be symmetrical and it is necessary only to display the left half. We will also drop the $M = 0$ term since it is not of general interest. With these changes the matrices for $n = 4$ and $n = 5$ are given in tables 2 and 3.

It is of interest to briefly consider the matrix in table 3 for $n = 5$ as an example of a relatively complicated and yet manageable matrix. The

S/M	1	2	3	4	5	6	7
0	7	3	1	0	0	0	0
1	8	8	6	4	2	1	0
2	0	4	6	6	4	2	2
3	0	0	2	4	6	6	4
4	0	0	0	1	3	5	5
5	0	0	0	0	0	1	4
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0

Table 2. Frequency Histogram Matrix - n = 4

S/M	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	15	7	3	1	0	0	0	0	0	0	0	0	0	0	0
1	16	16	12	8	5	3	1	0	0	0	0	0	0	0	0
2	0	8	12	12	10	6	6	5	2	0	0	0	0	0	0
3	0	0	4	8	10	10	6	4	5	5	2	0	0	0	0
4	0	0	0	2	5	9	9	7	5	5	6	6	2	0	0
5	0	0	0	0	1	3	9	12	10	6	4	4	6	5	1
6	0	0	0	0	0	0	0	3	8	10	8	4	4	5	7
7	0	0	0	0	0	0	0	0	1	5	10	[illegible]	[illegible]	0	6
8	0	0	0	0	0	0	0	0	0	0	1	5	9	9	5
9	0	0	0	0	0	0	0	0	0	0	0	[illegible]	[illegible]	5	9
10	0	0	0	0	0	0	0	0	0	0	0	[illegible]	[illegible]	1	3

Table 3. Frequency Histogram Matrix - n = 5

first n columns follow, except for one zero term the well-known and easily expressed binomial frequency law.[4] (See section 2) A fairly concise expression is also available[4] for the $n + 1$ column. For greater values of M no easily handled explicit expression is known. The problem is that if $M > n$ the summed terms become dependent and their frequency histograms highly complex. One of the objectives of sections 2 and 3 is to explain how this dependency arises and what kinds of distributions or frequency histograms arise in this and other kinds of summing techniques.

We do not have space to present frequency histograms for larger values of n. It is of value, however, to observe that as n increases and M increases the histograms become increasingly complex and apparently without order. This reflects the high degree of interdependency of terms. Consider, as an example, the histogram $H(s)$ for $n = 8$ and $M = 67$.

s	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
H(s)	3	16	27	28	8	11	17	26	33	26	19	18	20	2	1

This is the 67^{th} column of the $n = 8$ matrix. It is clearly very complex. There is no apparent order or structure in this matrix.

2.0 Weighted Sums from Feedback Shift Registers

In this section we consider a weighted sum of outputs from a feedback shift register. The basic system is shown in figure 2. It will be assumed that the feedback connections are such that maximal-length binary sequences of length $2^n - 1$ are generated.

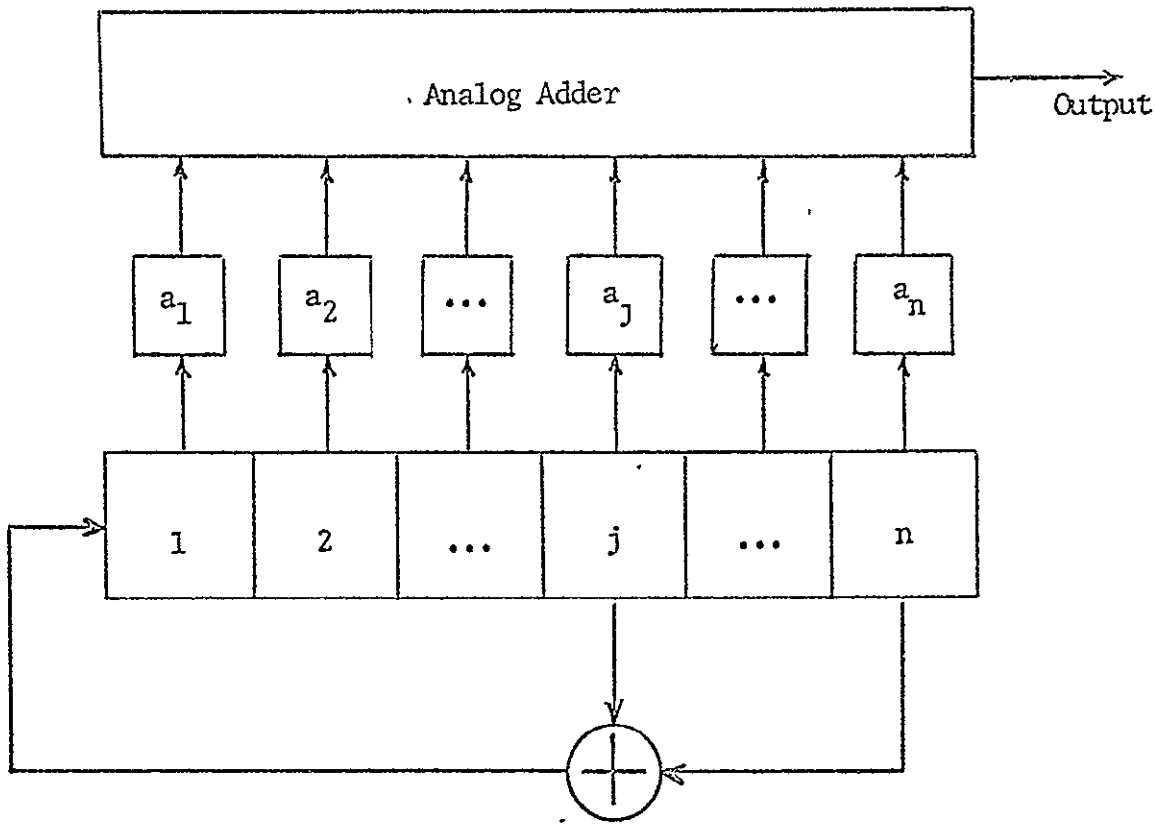


Figure 2. Weighted Sums

The required connections[5] will not be considered here. The n stages of the feedback shift register can have values 0 or 1 (arbitrary units) during any specified clocking period. As the shift register cycles through a complete period, each stage has a sequence of 0's and 1's following a well-known[1] and definite pseudo-statistical pattern.

The term "pseudo" is used here and in other parts of the paper to indicate that the outputs or variables appear to be random and thus capable of a statistical description. Actually, we are dealing with a finite-state-machine with a prescribed initial state and defining relationship, and the

output is thus known deterministically. Nevertheless, it is of great value to treat the output as random in some sense. We will adopt the practice here of speaking of probability distributions even though a term such as "deterministic frequency histogram" might be more correct.

2.1 The Binomial Distribution

In the diagram of figure 2 the weighting factors, a_j, can take any integer value. Consider first the case where $a_j = 1$ for all j from 1 to n. In this case, we are simply adding the number of outputs of value 1. It will be shown in section 3, when we discuss the generating polynomial, that the outputs of the stages are independent of each other. Since the stages can have two possible values, the sum of n stages must have a binomial probability distribution. Actually, there is one state, the all-zero state, which cannot occur. The result is that the distribution is binomial except for one term. With the probability of 0 or 1 equal to a half, the distribution of the sum becomes:[4],[7]

$$P(s) = \binom{n}{s}\frac{1}{2^n - 1} \qquad \text{for } 1 \leq s \leq n \tag{1}$$

where $\binom{n}{k}$ stands for the combinations of n things taken k at a time.

In much of the discussion to follow the all-zero state will be assumed possible. This simplifies calculations significantly. It will be removed when final results are presented. When the all-zero state is assumed the distribution is the familiar binomial. A series of binomial distributions for increasing values of n is easily presented by means of the Pascal triangle as shown in figure 3.[39]

								1													n =
							1		1												n =
						1		2		1											n =
					1		3		3		1										n =
				1		4		6		4		1									n =
			1		5		10		10		5		1								n =
		1		6		15		20		15		6		1							n =
	1		7		21		35		35		21		7		1						n =
1		8		28		56		70		56		28		8		1					n =
	•		•		•		•		•		•		•		•		•				•

Figure 3. The Pascal Triangle

Once again n is the number of stages. Thus, for n = 5 stages we can use the n = 5 row to find the binomial distribution of the sum of outputs in the five stages (including the all-zero state). To use the Pascal triangle for n = 5 we first normalize by dividing each term in the n = 5 row by the sum of the terms, which is $32(2^5)$. The first term on the far left is the sum equals zero term. Thus, the probability that the sum equals zero is 1/32. The probability that the sum equals one is 5/32, etc. The complete distribution is:

S	0	1	2	3	4	5
P(s)	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

2.2 Sums of random variables - Discrete Convolution

We wish now to consider the very interesting case where the a's are

not all one. In the work which is to follow we make use of a mathematical approach and a nomenclature which are not too common. We will, therefore, review these ideas first.

We will write the probability distribution as a sequence of unnormalized numbers. For example, the binomial distribution for $n = 5$ is:

[1 5 10 10 5 1]

This is intended to convey exactly the same information as the distribution at the end of section 2.1.

Recall now that we are basically interested in obtaining the analog sum of some independent random variables (in the stages of the feedback shift register.) It is well-known (see for example [40], Chapter 15) that the probability distribution of the sum of two random variables is the convolution of the two original probability distributions. It is particularly easy to convolve two discrete probability distributions using the sequence format shown above. (see [41] or chapter 3 of [40]). As an example let us find the distribution of the sum of binomial random variables for $n = 3$ and $n = 4$. From the Pascal triangle we see that the correct sequences are:

[1 3 3 1] for $n = 3$

[1 4 6 4 1] for $n = 4$

There are various ways to mechanically carry out this convolution. Perhaps the simplest is to set up a format much like that of conventional multiplication. The only difference is that with this approach we do not carry tens.

```
                    1     4     6     4     1
                          1     3     3     1
                   ------------------------------
                    1     4     6     4     1

              3    12    18    12     3
        3    12    18    12     3
  1     4     6     4
 ------------------------------------------------
  1     7    21    35    35    21     7     1
```

Using the symbol * for convolution we write the above problem in the form:

$$[\,1\ \ 4\ \ 6\ \ 4\ \ 1\,] * [\,1\ \ 3\ \ 3\ \ 1\,] = [\,1\ \ 7\ \ 21\ \ 35\ \ 35\ \ 21\ \ 7\ \ 1\,]$$

(We might note that the solution to this convolution problem is a sequence equivalent to the binomial distribution for $n = 7$. The property of obtaining the distribution for order n by convolving any two distributions whose orders add to n holds in general for the binomial family, as well as for a number of other families. It does not hold for all families of distributions.)

2.3 The Uniform Distribution

With the above technique we can now obtain the distribution of the sum of weighted outputs from stages of the feedback shift register.

As a first example consider the simple case where $n = 2$, $a_1 = 1$ and $a_2 = 2$. The weighted output of the first stage is 0 or 1 with equal probability of 1/2 (we are assuming the all-zero state for the present). This distribution is expressed as:

$$[\,1\ \ 1\,]$$

The weighted output of the second stage is 0 or 2 with equal probability which we express as:

[1 0 1]

That is, zero appears with probability 1/2, one with probability zero and two with probability 1/2. All terms below zero and above the greatest term specified in the sequence occur with probability zero. To find the distribution for both stages, out of the analog adder, we convolve the above to obtain:

[1 1] * [1 0 1] = [1 1 1 1]

This says that the sums 0, 1, 2 and 3 occur with equal probability 1/4. This result can be checked by means of the following table. Call the state of the two stages k_1 and k_2.

k_1	k_2	a_1k_1	a_2k_2	$a_1k_1 + a_2k_2$
0	0	0	0	0
0	1	0	2	2
1	0	1	0	1
1	1	1	2	3

Thus we see that each term appears just once for the four distinct and equally probable combinations of states k_1 and k_2. We can now make use of the fact that the all zero state is not possible and the maximal-length sequence has a length $2^2 - 1 = 3$. Without the all zero state our distribution becomes

[0 1 1 1]

Now let us add a third stage with weight factor 4. First we go back and restore the all zero state until the problem is completed. The distribution of the weighted third stage is:

$$[\,1 \quad 0 \quad 0 \quad 0 \quad 1\,]$$

That is, the values 0 and 4 occur with equal probability 1/2, and the values 1 , 2 and 3 occur with probability zero. To find the distribution of the sum for three stages we convolve the distribution of the third stage with that of the first two.

$$[\,1 \quad 1 \quad 1 \quad 1\,] * [\,1 \quad 0 \quad 0 \quad 0 \quad 1\,] = [\,1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1\,]$$

and removing the all zero term we have

$$[\,0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1\,]$$

which says that the integers one through seven occur with equal probability.

If we add fourth stage with weight 8 the distribution suggests that the first 15 integers are equally likely to occur. Appropriate weightings cause this pattern to be continued. Thus we come to the most interesting conclusion that a uniform pseudo-random distribution can be generated by a feedback shift register with suitable weights. It is apparent that the proper weight for the j^{th} stage is 2^{j-1} if a uniform distribution is desired.

2.4 Some Additional Non-Binomial Distributions

The possible variations suggested by this basic approach are endless. We mention here only some of the most important.

Consider a six stage register with weights $a_1 = a_2 = 1$, $a_3 = a_4 = 2$,

$a_5 = a_6 = 4$. We have just seen that the combination of weights 1, 2, 4 produces a distribution.

[1 1 1 1 1 1 1 1]

But now we have two such distributions so we convolve them together to obtain:

[1 1 1 1 1 1 1 1] * [1 1 1 1 1 1 1 1]

= [1 2 3 4 5 6 7 8 7 6 5 4 3 2 1]

The result is a triangular (Simpson) distribution. (The pertinent weights and final distributions are summarized for this case and other important cases in table 4.)

A "stairstep" distribution is obtained from a combination of weights such as 1, 2, 4, 4 which give.

[1 1 1 1 1 1 1 1] * [1 0 0 0 1]

= [1 1 1 1 2 2 2 2 1 1 1 1]

Finally, a "Bimodal" distribution is obtained from weights such as 1,1,1,1,5 which give:

[1 4 6 4 1] * [1 0 0 0 0 1]

= [1 4 6 4 1 1 4 6 4 1]

The peak of the second mode can be extended out to any value by simply increasing the last weight. This bimodal technique can be used on any type of distribution by simply adding a stage with a weight equal to the distance the second mode is to be shifted.

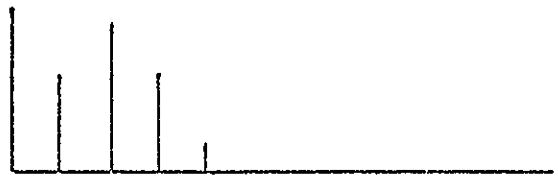
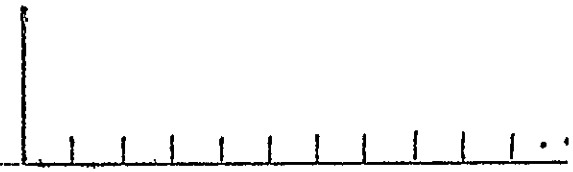
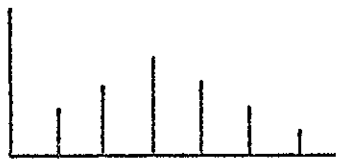
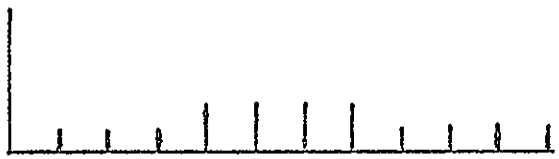
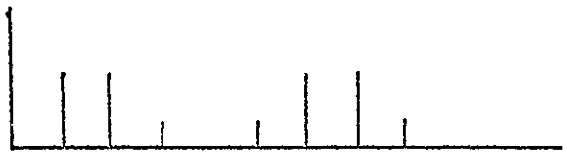
Name	Weights				Distribution	Histogram
	a_1	a_2	a_3	a_4		
Binomial	1	1	1	1	[0 4 6 4 1]	
Uniform	1	2	4	8	[0 1 1 1 1 1 1 1 1 1 1 ...]	
Triangular	1	2	1	2	[0 2 3 4 3 2 1]	
Stairstep	1	2	4	4	[0 1 1 1 2 2 2 2 1 1 1 1]	
Bimodal	1	1	1	5	[0 3 3 1 0 1 3 3 1]	

Table 4. Synthesis of Distributions of Pseudo-Random Variables

The preceding examples are summarized in table 4. Necessary weights and the resulting distributions are shown. The histogram is simply a pictorial representation of the distribution.

Once again it should be emphasized that the possible variations to the above are unlimited. Any number of different distributions can be generated simply by changing the weights, a_j .

Although the above technique for synthesizing pseudo-distributions is very attractive it is not without limitation. Consider the uniform distribution, generated by weights 2^{j-1} . The range of this pseudo-random variable is from 1 to $2^n - 1$ where n is again the number of stages. Thus, for example, this simple circuit cannot be connected to yield a pseudo-random variable uniform between 0 and 9 . The latter would be necessary to develop a decimal range pseudo-random number generator.

These results have been reported on in the literature[42] [43].

2.5 Some Thoughts on Range and Smoothness

In this section we are interested first in the range of the pseudo-random variable S , the output from the circuit of figure 2, and second on the effect which weighting has on "smoothness" of the output distribution. These factors are of course inter-related. Note that we define the range here as the number of possible outcomes or values of S , rather than the difference between S_{max} and S_{min} .

The range of S can be important for a number of applications as it indicates the number of different output levels of the device. Aggarwal[44], for example, has suggested a frequency generator which generates a number of frequencies equal to the number of different levels of S .

If we consider the circuit of figure 2 with unity weights, the range

on the sum is n since the sum can be any number from 1 to n, with a distribution following (almost) the binomial law. If we let the weights be 2^{j-1}, we obtain a uniform distribution with range $2^n - 1$, as we saw in section 2.3. This latter range is obviously the maximum range for the circuit of figure 2, since after $2^n - 1$ outputs the periodic sequenc repeats. Likewise n is the minimum range. Hence, the range r_s has bounds:

$$n \leq r_s \leq 2^n - 1$$

where the weighting sequence for the lower limit is $[\, 1, 1, 1, \cdots 1 \,]$, and for the upper limit $[1, 2, 4, 8, \cdots 2^{n-1}\,]$. For any weighting sequ $[\, a_1, a_2, \cdots, a_n \,]$ the range r_s is simply the number of different numl which can be obtained by adding any combinations of the weights taken anyw from 1 to n at a time.

The analysis in sections 2.3 and 2.4 raises some interesting question related to the Central Limit Theorem tendencies of sums of random variable The Central Limit Theorem says that the distribution of the sum of continu random variables tends to be Gaussian (or normal) as the number of summed variables increases, under certain conditions. A discrete approximation to the Gaussian is the Binomial distribution. The accuracy of the approxi increases as the order of the binomial increases. The Gaussian distributi is in a sense the "smoothest" possible continuous distribution. In the sa general sense the Binomial distribution is the smoothest discrete distribu Summed discrete random variables tend to be Binomial under conditions analogous to those necessary for the Central Limit Theorem.

Consider what happens when we use weights of unity in all stages of t circuit of figure 2. The distribution of each stage is $[\, 1 \;\; 1 \,]$ and the

distribution of the sum for an n-stage circuit is binomial of order n.

Consider now what happens as we start to use non-unity weights. Let $n = 3$ and find the distribution for a series of weight combinations as shown in table 5. Assume the all-zero state.

a_1	a_2	a_3	0	1	2	3	4	5	6	7
1	1	1	1	3	3	1	0	0	0	0
1	2	1	1	2	2	2	1	0	0	0
1	2	2	1	1	2	2	1	1	0	0
1	2	3	1	1	1	2	1	1	1	0
1	2	4	1	1	1	1	1	1	1	1

Table 5. Effect of Weights on "Smoothness"

As the weights increase the range increases and the "smoothness" decreases. The ordinary (normal?) tendency under summing is for the distribution to become increasingly smooth. However, the tendency of the weights is to emphasize individual distributions. These two tendencies tend to offset each other. If the weights are chosen as 2^0, 2^1, 2^2, 2^3 .. as in the last line of table 5, the two tendencies cancel each other and we retain the uniform distribution.

It is of interest to consider the sum which leads to the uniform distribution in more detail. First, we recall that the condition for the Central Limit Theorem, the Lindeberg-Feller condition, can be expressed in simple form as.

$$\lim_{n\to\infty} \frac{\sigma_j^2}{\sum_{j=1}^{n} \sigma_j^2} = 0 \qquad (2)$$

where σ_j^2 is the variance of the j^{th} random variable.

This condition is essentially applicable to the discrete case as well to the continuous case to the extent that the binomial distribution approximates the Gaussian. Let us consider the limit in (2) for the case where the weights are 2^0, 2^1, 2^2, 2^3, $\cdots$. The corresponding distributions are: [1 1], [1 0 1], [10001], $\cdots$ and the corresponding variances are 1/4, 1, 4, 16, $\cdots$ or 4^{j-2} where $j = 1, 2, 3, \cdots$.

The ratio of the variance of the j^{th} variable to the sum of the variances of the first n variables is:

$$R_j = \frac{4^{j-2}}{\sum_{j=1}^{n} 4^{j-2}}$$

$$= \frac{4^j}{\sum_{1=1}^{n} 4^j}$$

$$= \frac{3x4^j}{4^{n+1}-1} \qquad (3)$$

The largest variance appears for $j = n$ in which case the ratio is:

$$R_n = \frac{3}{4-1/4n} \qquad (4)$$

And the limit is:

$$\lim_{n\to\infty} R_n = 0.75 \qquad (5)$$

Thus, for this case, the Lindeberg-Feller conditions are not met. This is, of course, exactly as we would have expected. A similar analysis for

series of weights 1, 2, 3, 4, ... indicates the limit is zero. Thus, these weights are not large enough to offset the central tendency of the summed random variables. We will return to this question briefly when we discuss infinite sums in section 4.

2.6 The Autocorrelation Function

In the preceding work, it was convenient to let the states take values 0 and 1 . In our study of the autocorrelation function, we will let these values be -1 and +1 . We are interested in the stream of values out of the analog adder in figure 2. This can be thought of as a stochastic process with a new value occurring at each clocking time. We will designate this sequence of output values as:

$$[\, b_1, b_2, \cdots, b_k, \cdots, b_{2^n-1} \,]$$

We will only consider 2^n-1 terms since the sequence is periodic and repeats after 2^n-1 terms. We define the autocorrelation function as:

$$R(m) = \sum_{k=1}^{2^n-1} b_k \, b_{k+m} \quad , \quad m = 0,1,2,\cdots \tag{6}$$

(Note that because the sequence is periodic $b_{k + 2^n-1} = b_k$.)

This is also equal to the number of agreements (same sign) minus the number of disagreements when the original sequence is compared with itself shifted by m units.

We consider first the case where all the a's except one in figure 2 equal zero. The one exception is equal to 1 . For this case, it is known[1] that the autocorrelation function is:

$$R_1(m) = \begin{cases} 2^n-1 & m=0 \\ -1 & m\neq 0 \end{cases} \qquad (7)$$

This function which also has period 2^n-1 , is plotted in figure 4. Note that we are considering m as a discrete integer variable. The graph represents an integer as all of the space between the integer and the next higher integer.

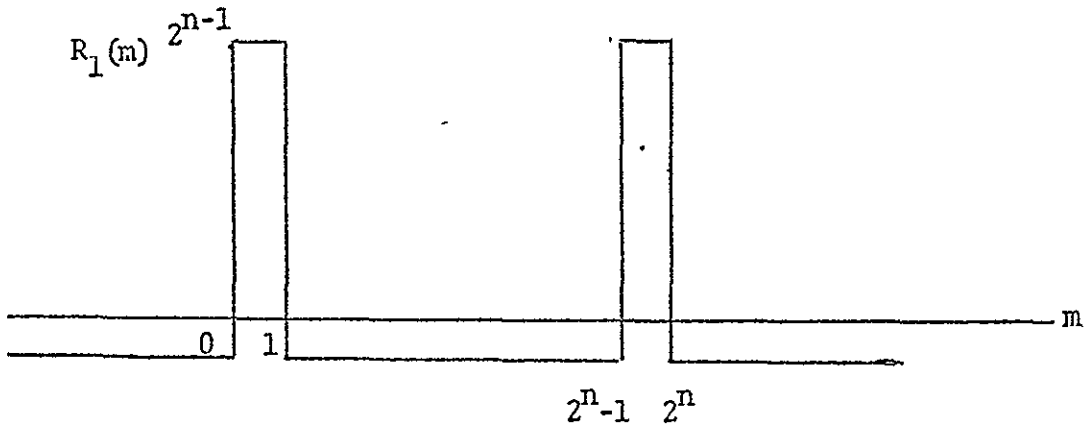


Figure 4: Autocorrelation Function of Maximal-Length Pseudo-Random Sequence

From figure 4 we see that as the length or period of the sequence (2^n-1) increases the autocorrelation becomes increasingly concentrated at $m = 0$.

Now we wish to study this autocorrelation function for the stream of values out of the analog adder in figure 2 for some non-trival weightings. As an example let $a_1 = 1$, $a_2 = 1$ and all other weights be zero. Then the sequence out of the generator is:

$$[\, b_1 + b_2,\ b_2 + b_3,\ \cdots,\ b_k + b_{k+1},\ \cdots,\ b_{2^n-1} + b_1 \,]$$

with autocorrelation function:

$$R_2(m) = \sum_{k=1}^{2^n-1} (b_k + b_{k+1})(b_{k+m} + b_{k+m+1})$$

$$= \sum_{k=1}^{2^n-1} (b_k b_{k+m} + b_{k+1} b_{k+1+m} + b_{k+1} b_{k+m})$$

$$= 2R_1(m) + R_1(m-1) + R_1(m+1) \tag{8}$$

Thus we have established the very interesting and useful, though hardly surprising, result that the autocorrelation function of the sum of sequences is a sum of the individual sequence autocorrelation functions. This discrete result is analogous to the well-known[45] expression for the autocorrelation of a sum of continuous random variables:

$$R_{x+y}(\tau) = R_{xx}(\tau) + R_{xy}(\tau) + R_{yx}(\tau) + R_{yy}(\tau) \tag{9}$$

Consider now the sum of p sequences obtained by letting $a_1 = a_2 = \cdots = a_p = 1$ and all other a_j's are zero. The autocorrelation function is:

$$R_p(m) = pR_1(m) + (p-1)R_1(m-1)$$

$$+ (p-1)R_1(m+1) + (P-2)R_1(m-2) + \cdots$$

$$+ R_1(m-p+1) + R_1(m+p+1) \tag{10}$$

This result is a fairly straightforward extension of the $p = 1$ and $p = 2$ cases. It simply indicates the number of cross-terms of given delay obtained in the multiplication:

$$(b_k + b_{k+1} + \cdots + b_{k+1+m})(b_{k+m} + b_{k+m+1} + \cdots + b_{k+m+p-1})$$

Equation (10) reduces to the following simple relation for the autocorrelation of the sum of p successive unity-weighted pseudo-random sequences.

$$R_p(m) = \begin{cases} (P-|m|)2^n - P^2 \ , & m = 0 \ , \pm 1 \ , \pm 2 \ , \cdots \ , + P \\ -P^2 \ , & m = P \ , P+1 \ , \cdots \ , 2^n - 1 - P \end{cases} \quad (11)$$

This function is plotted in figure 5 for $n = 4$ and $p = 1,2,3,4,5,6,7$ and 8. The effect of summing p sequences is seen to be to introduce "memory" into the output process. The process "remembers" over a number of clocking pulses equal to P, the number of successive sequences added. The effect is analogous to passing the original sequence through a low-pass filter. In fact the system is actually a non-recursive digital filter.

It is, of course, not necessary to add successive stages or to use unit (or equal) weights. We consider first the sum of non-successive sequences and second the sum of sequences with different weights.

Let $a_1 = a_3 = 1$ and all other weights be zero.

$$R_{13}(m) = \sum_{k=1}^{2^n-1} (b_k + b_{k+2})(b_{k+m} + b_{k+m+2}) \quad (12)$$

$$= 2R_1(m) + R_1(m-2) + R_1(m+2)$$

And, similarly:

$$R_{1q}(m) = 2R_1(m) + R_1(m-q) + R_1(m+q) \quad (13)$$

is the autocorrelation function of the sum of a sequence and the q^{th} later sequence. The effect of delaying the weighting to the q^{th} stage is to delay the "memory" in the process. As an example $R_{14}(m)$ is plotted in

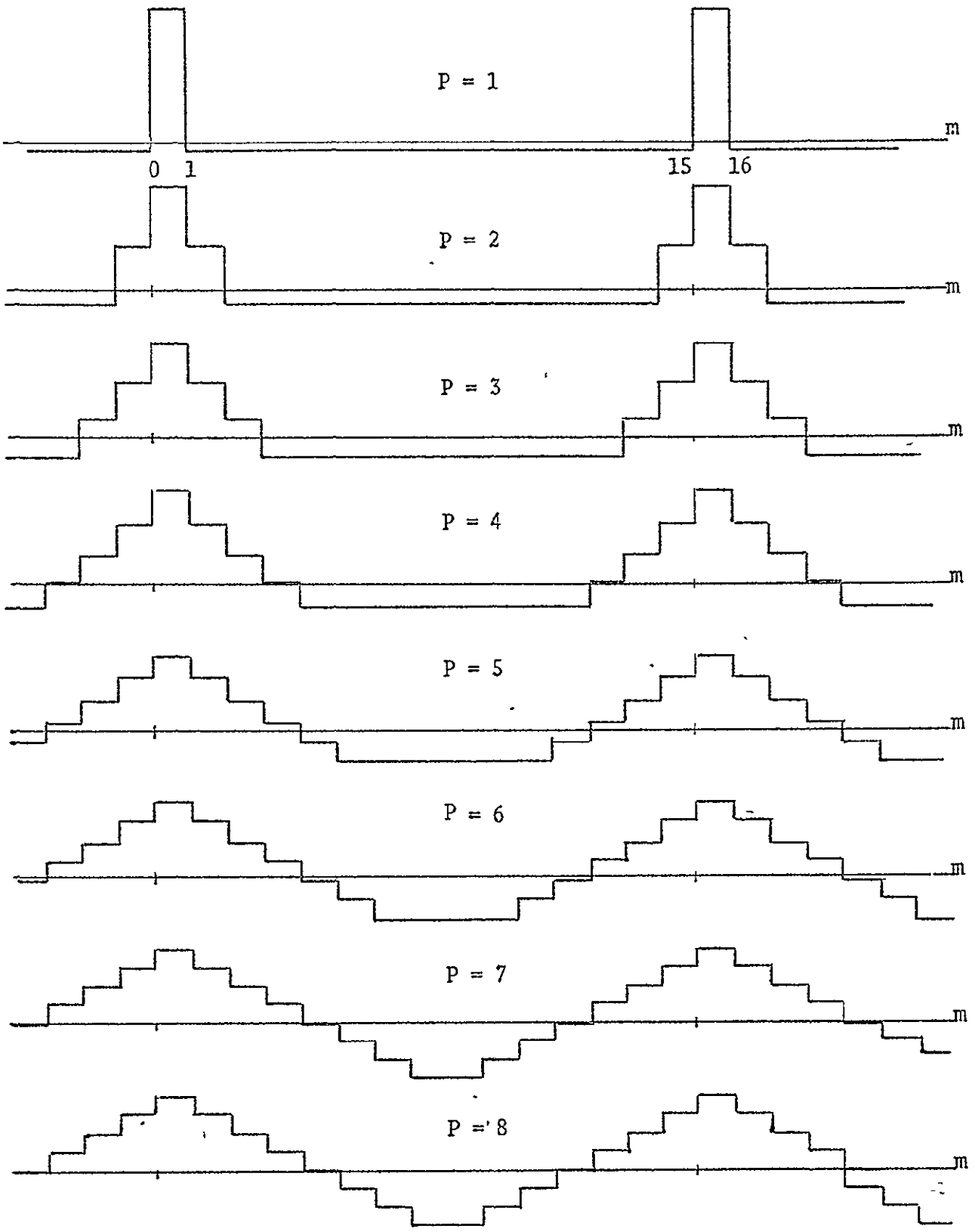


Figure 5. Autocorrelation Function - $R_p(m)$
(scale not consistent for different values of p)

figure 6.

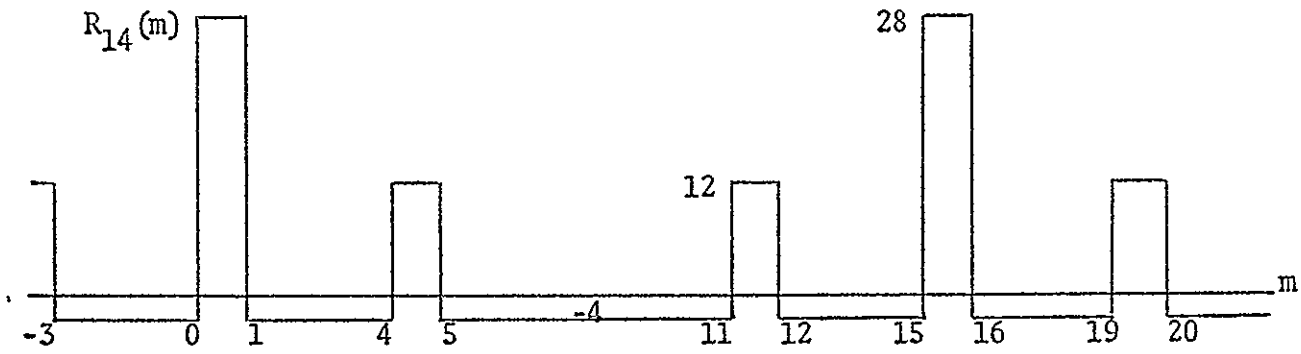


Figure 6

Other interesting autocorrelation functions can be obtained by adding sequences with different delays.

We turn now to the effect of non-unit weights. Let $a_1 = 1$ and $a_2 = e$. The sequence out of the generator is:

$$[b_1 + eb_2,\ b_2 + eb_3,\ \cdots,\ b_k + eb_{k+1},\ \cdots,\ b_{2^n-1} \pm eb_{k+m+1})$$

and its autocorrelation function is:

$$R_2(m)\Big|_{\substack{a_1=1\\a_2=e}} = \sum_{k=1}^{2^n-1} (b_k + eb_{k+1})(b_{k+m} + eb_{k+m+1})$$

$$= (1+e^2)R_1(m) + eR_1(m-1) + eR_1(m+1) \qquad (14)$$

This function is plotted in figure 7 for $e = 1,2,3,$ and 4 . The important difference between these results and those of figure 5 is that in figure 5 the autocorrelation decreases in equal steps (triangular) whereas in figure 7 the function drops off more rapidly for m near 0 than for

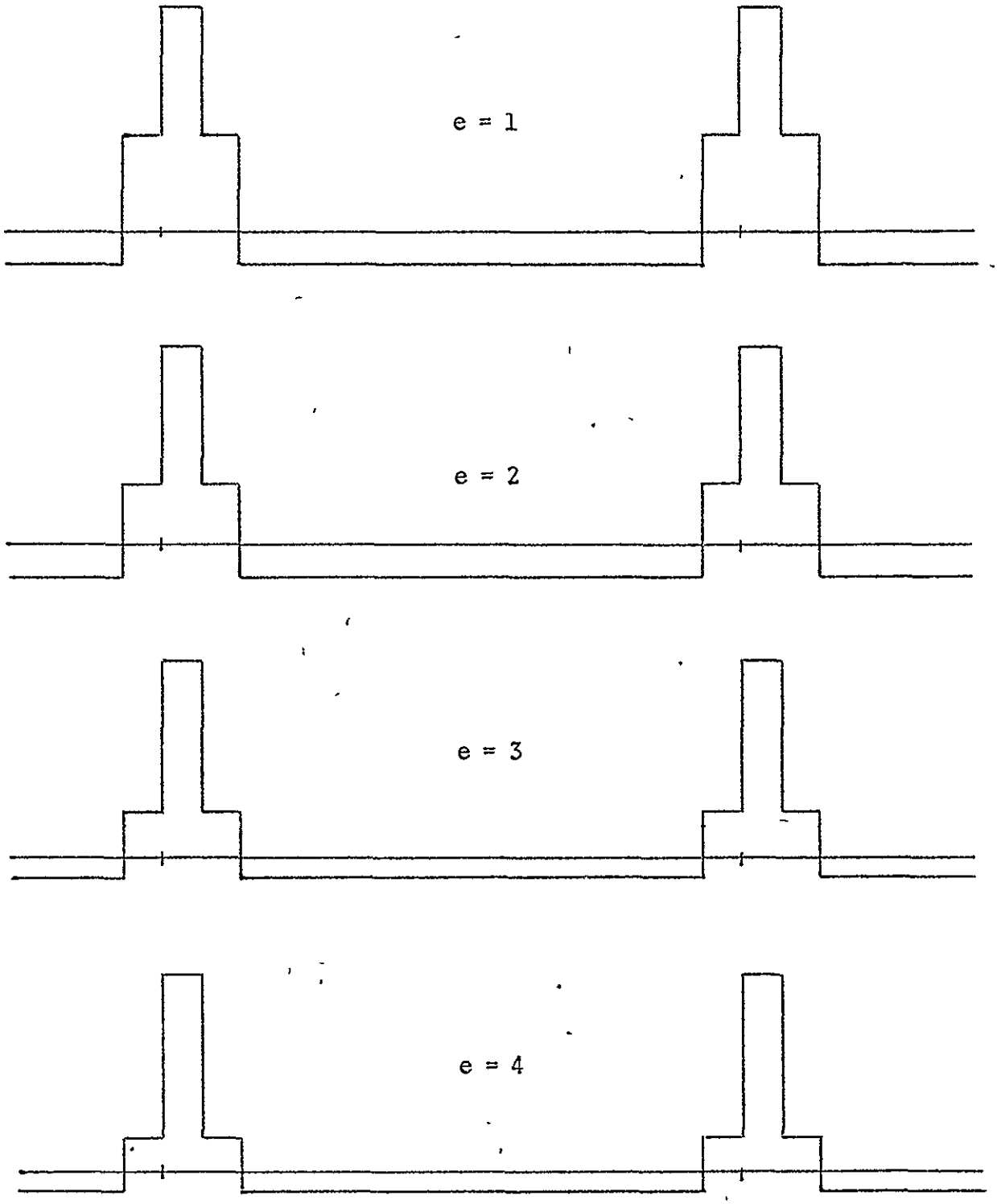


Figure 7. Autocorrelation Functions for Different Weights
(Scale not consistent for different values of e.)

larger m . The latter function resembles an exponentially decreasing function. As e increases the autocorrelation function resembles increasingly the autocorrelation function of the original pseudo-random sequence. This is certainly reasonable because values of e much greater than 1 simply emphasize the heavily weighted sequence. The important point here is that the process of weighting does effect the autocorrelation function and does so in a way which is intuitively reasonable.

Let us consider the effect on the autocorrelation function of binary weights ($a_1 = 1$, $a_2 = 2$, $a_3 = 4$, $\cdots$) . This function is plotted in figure 8 for a number of combinations of binary weights. The functions fall off in an exponential-like way with each successive step being about 1/2 the previous step when all levels are referenced not to zero but to the minimum level.

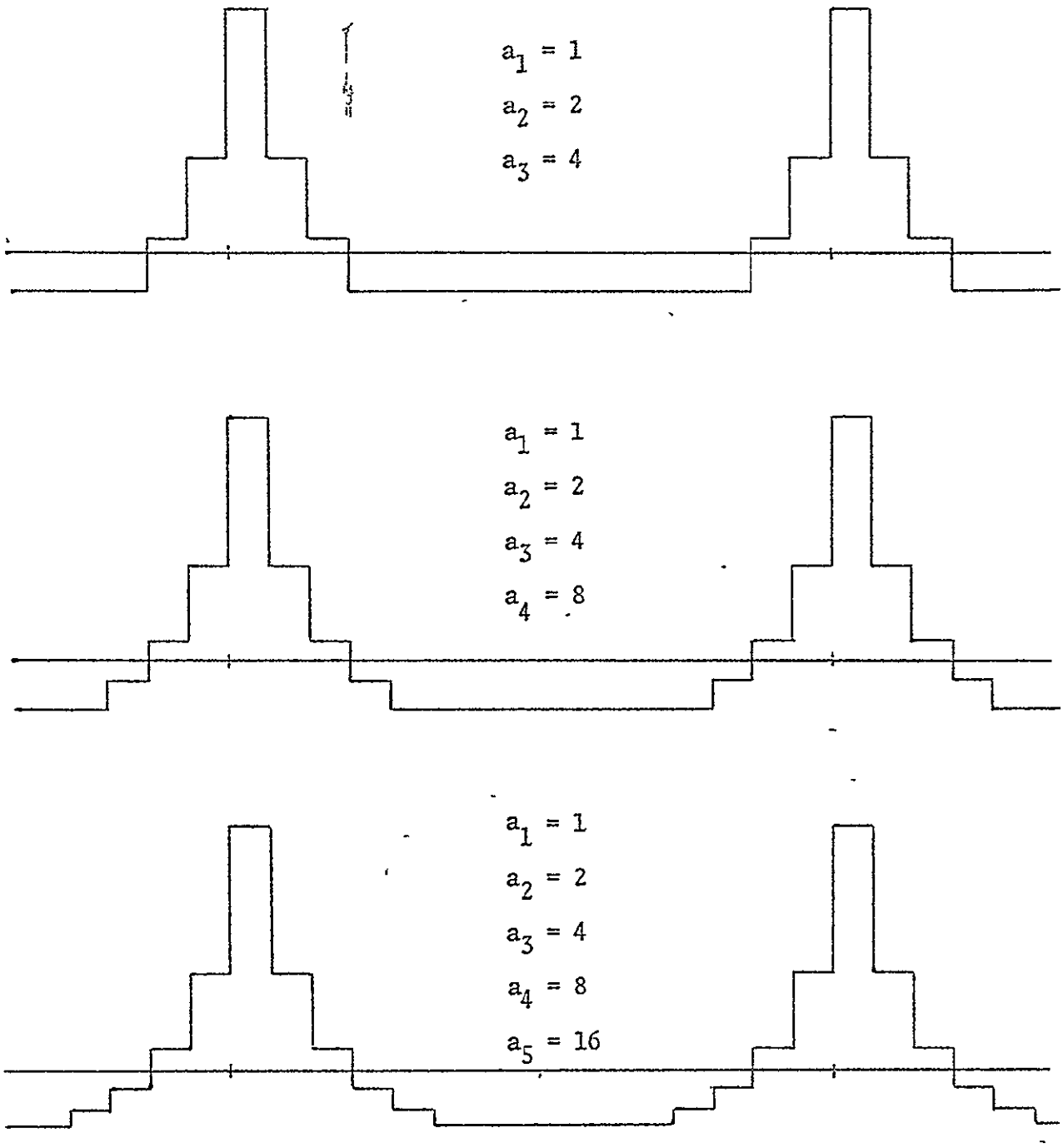


Figure 8. Autocorrelation Functions for Binary Weights
(Scale not consistent for different weights.)

3.0 Augmented Feedback Shift Registers

We turn now to an alternative approach to changing the distribution obtainable from a feedback shift register. Actually there are two ways of implementing this approach. It will be shown that these ways are related to each other and likewise are related to a lesser extent to the approach just discussed in Section 2.2. Before we consider the two ways available under this alternative approach we must briefly consider the defining or characteristic polynomial for the feedback shift register.

3.1 The Characteristic Polynomial

The state of a feedback shift register is completely specified by its initial conditions and a so-called characteristic polynomial which interrelates outputs from various stages and the input to the first stage. One form of the characteristic polynomial for a feedback shift register with n stages connected to generate a maximal length sequences is:

$$D^n \oplus A_{n-1} D^{n-1} \oplus A_{n-2} D^{n-2} + \cdots \oplus A_1 D^1 \oplus D^0 = 0 \qquad (15)$$

where $\oplus$ stands for the modulo 2 sum. This operation is defined in the table below.

$\oplus$	0	1
0	0	1
1	1	0

Table 6

The coefficients A_j are either 0 or 1 depending on the feedback connections in the shift register. D is an operator signifying delay of order equal to its exponent. To clarify this general expression consider

the particular circuit shown in figure 9, and its characteristic polynomial:

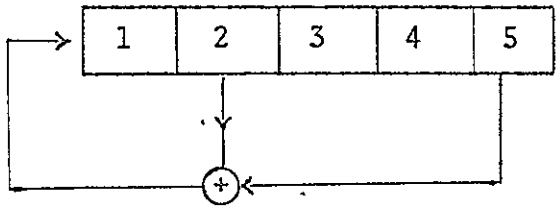


Figure 9. Five Stage Binary Feedback Shift Register

$$D^5 \oplus D^2 \oplus D^0 = 0 \ . \qquad (16)$$

Note first that there is no need or use for a minus sign in modulo 2 arithmetic since -1 and +1 are equivalent. Therefore we can put any term in equation (16) on the right side of the equation without changing signs. Thus, for example, we can write: $D^5 \oplus D^2 = D^0$. (17)

The interpretation of this expression is as follows. The value in a certain stage is equal to the modulo 2 sum of the value two time delay periods earlier and the value five time delay periods earlier. Since the shift register shifts values or outputs one position each delay or clocking period, this is equivalent to saying that the modulo 2 sum of the value in the 5th and 2nd stages is equal to the value in the 0^{th} stage. The latter is a hypothetical stage which feeds the first stage at a clocking instant.

Since D simply stands for delay we can increase the exponent of each term in (17) and still have a correct equation. Thus, for example

$$D^6 \oplus D^3 = D^1 \qquad (18)$$

Similarly, this can be repeated any number of times, a property which will be very useful in the sections to come. The characteristic equation also tells us that values in stages are independent as long as the stages are

no more than 5 delay periods apart. In general, values are independent if stages are no more than n delay periods apart where n is the number of stages in the feedback loop.

3.2 Added Stages

We are ready now to consider the first of two alternative ways to change the distribution available from a feedback shift register. Consider the circuit of figure 10.

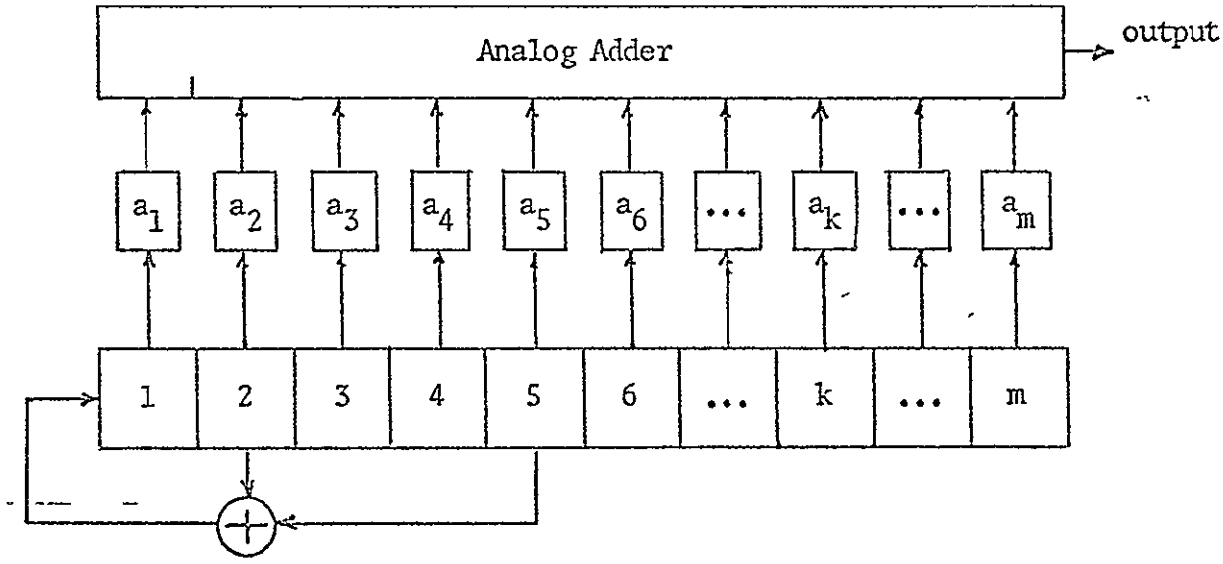


Figure 10. Augmented Feedback Shift Register

The basic feedback shift register is identical to that of Figure 9. To this we have added m-5 delay stages which are part of the shift register but are not in the feedback loop. In much of our discussion to follow, we will assume that the weighting factors related to the stages in the feedback loop are all equal to unity.

Consider as a first example the problem where m = 31 in Figure 10. Let a_1 through a_6 equal one and all other factors equal zero. That is, we are adding the outputs of the five stages plus one additional stage

delayed one clocking period from stage 5. Call this analogsum "S_6".

$$S_6 = D^1 + D^2 + D^3 + D^4 + D^5 + D^6 \tag{19}$$

We know from section 2 that the sum of the first five stages is binomial. But the addition of a sixth stage adds a dependency which makes the resulting sum non-binomial. To see this we substitute D^6 from equation (18) into equation (19) and collect terms to obtain:

$$S_6 = D^1 + D^2 + D^3 + D^4 + D^5 + (D^1 \oplus D^3) = D^2 + D^4 + D^5 + [D^1 + D^3 + (D^1 \oplus D^3)] \tag{20}$$

The terms within the bracket were collected because they are inter-dependent. The total bracketed term and the first three terms are independent and the resulting distribution of the sum can be found by the convolution techniques developed in section 2. The distribution of the terms D^2, D^4 and D^5 is binomial of order 3. That is, it is of the form. [0 3 3 1].

The distribution of the complex term in the brackets cannot be found from convolution techniques. It is possible, however, to find the distribution using a table of possible states as shown below.

D^1	D^3	$D^1 \oplus D^3$	$D^1 + D^3 + D^1 \oplus D^3$
0	0	0	0
0	1	1	2
1	0	1	2
1	1	0	2

Table 7

The resulting distribution of $[D^1 + D^3 + (D^1 + D^3)]$ is: [1 0 3].

When we convolve this with the distribution for D^2, D^4 and D^5 we finally obtain:

[1331] * [103] = [1 3 6 10 9 3]

(see Column 6 of Table 3).

If we now change the problem so that the output of the seventh stage is added to the first five we have:

$$S_7 = D^1 + D^2 + D^3 + D^4 + D^5 + D^7 = D^1 + D^3 + D^5 + [D^2 + D^4 + (D^2 \oplus D^4)] \tag{21}$$

and the resulting distribution is as in the previous example since the structure is the same.

Now, however, consider the case where the added term is in the ninth stage. Then we have:

$$\begin{aligned} S_9 &= D^1 + D^2 + D^3 + D^4 + D^5 + D^9 \\ &= D^1 + D^2 + D^3 + D^4 + D^5 + D^6 \oplus D^4 \\ &= D^1 + D^2 + D^3 + D^4 + D^5 + D^1 \oplus D^3 \oplus D^4 \\ &= D^2 + D^5 + [D^1 + D^3 + D^4 + (D^1 \oplus D^3 \oplus D^4)] \end{aligned} \tag{22}$$

In this case the dependency is in three stages rather than two and the distribution changes. Using the basic approach above, with an 8 column table in this case, we finally obtain the distribution, [1 2 7 12 7 2 1], or, dropping the all-zero term, [0 2 7 12 7 2 1].

After two examples we are led to the general question, what is the distribution for. $S_j = D^1 + D^2 + D^3 + D^4 + D^5 + D^j$ where j can be any

number from 1 to 31? First, we should establish our range of interest on j as 31 (or 2^5-1). We will not exceed 31 since any value of j greater than 31 is equivalent to some j between 1 and 31 . This follows from the fact that the maximal-length pseudo random sequence repeats every 2^n-1 (2^5-1=31 in this case) clocking periods. Furthermore, we note that every S_j involves a different combination of the first five stage terms. For if the D^j term consisted of exactly the same combination of (D^1 , D^2 , D^3 , D^4 and D^5) for different values of j the sequence would have a period equal to the difference of the values of j in question. Thus there are 31 different combinations of terms arising in S_j which should be considered in determining the possible distributions obtainable.

Not all of these combinations produce different distributions. (We saw above that S_6 and S_7 had the same distribution). Since the distribution is dictated by the structure, which reflects the number of interrelated terms, it is fairly obvious that there are 5 possible distributions depending on whether there are 1, 2, 3, 4, or 5 interrelated terms in a particular combination.

The question of how many ways there are to obtain a particular distributior is easily answered. There is just one way in which all five terms can be interrelated. Thus, one of the 31 possible values of j leads to the distribution typical of the case where all five terms are interrelated. There are 5 distinct ways of combining four dependent terms. That is, the combinations of five things taken four at a time is five. And, in fact, the combinaterial law is used to find the number of values of j which lead to a particular distribution. The distributions which are obtained by adding one additional delayed value are summarized in the following table 8.

Included for reference is the binomial case where there are no dependencies. In all cases the all-zero term has been ommitted.

Number of Interdependent Stages	Number of Ways to Obtain Distribution	Distribution
0	1	[0 5 10 10 5 1]
1	5	[0 4 7 8 4 1]
2	10	[0 3 6 10 9 3]
3	10	[0 2 7 12 7 2 1]
4	5	[0 1 10 10 5 5]
5	1	[0 0 15 0 15 0 1]

Table 8

The above example is limited of course to a shift register with five stages in the feedback loop and to a single added stage augmenting the basic five stages. It is of interest to generalize and expand our development.

Consider first the case where we allow more than one additional stage. Thus, for example, we might consider the sum:

$$S_{67} = D^1 + D^2 + D^3 + D^4 + D^5 + D^6 + D^7$$

$$= D^5 + [D^1 + D^3 + (D^1 + D^3)] + [D^2 + D^4 + (D^2 + D^4)] \qquad (23)$$

The bracketed terms, identical in structure to the bracketed term in equation (20), have the distribution [1 0 3] as shown earlier. Thus, the total distribution of S_{67} is:

$$[1\ 1] * [1\ 0\ 3] * [1\ 0\ 3] = [1\ 1\ 6\ 6\ 9\ 9]$$

or, dropping the all-zero term:

[0 1 6 6 9 9] (see column 7 of table 3)

This is a new distribution as we might have expected. Once again, there are a large number of different ways of adding 2 more values. Thus, many different distributions can be obtained depending on which two delayed values are selected. We can, of course, extend the number of added stages to 31 with a very great many different distributions obtainable.

A second way of expanding on the possible distributions is by letting the weighting factors, A_j , take values other than 1. (In section 2, we essentially covered this situation for the case where j was between 1 and n .) Allowing A_j to take <u>any</u> value leads to literally countless different distributions.

A final generalization is to allow n (5 in the above example) to take any value. Again, this opens up a whole new range of distributions.

No attempt has been made to chronicle any significant number of the infinite number of possible distributions. The basic procedure is summarized below.

1) Obtain the characteristic equation corresponding to the feedback shift register of interest (assume its length is n).

2) Write down the sum S in terms of the first n values of D^j $(1 \leq j \leq n)$. Use the characteristic equation to express values of D^j for $j<n$ in terms of values of D^j for $1 \leq j \leq n$.

3) Separate the terms in S into independent groups.

4) Find the distribution of each independent group through the tabular method shown above.

5) Convolve the distributions of the independent groups to find the overall distribution of the sum.

3.3 Added Logic

The approach used in section 3.2 suggests an alternative to adding additional stages. Since we can express any sum in terms of the first n values of D^J we can obtain any sum by simply using logic (modulo 2 adders) prior to analog summing. Assume, for example, that we wished to obtain the sum S_6 defined in equation (20) as:

$$S_6 = D^1 + D^2 + D^3 + D^4 + D^5 + (D^1 + D^3) \tag{24}$$

This sum is easily realized by the circuit shown in figure 6.

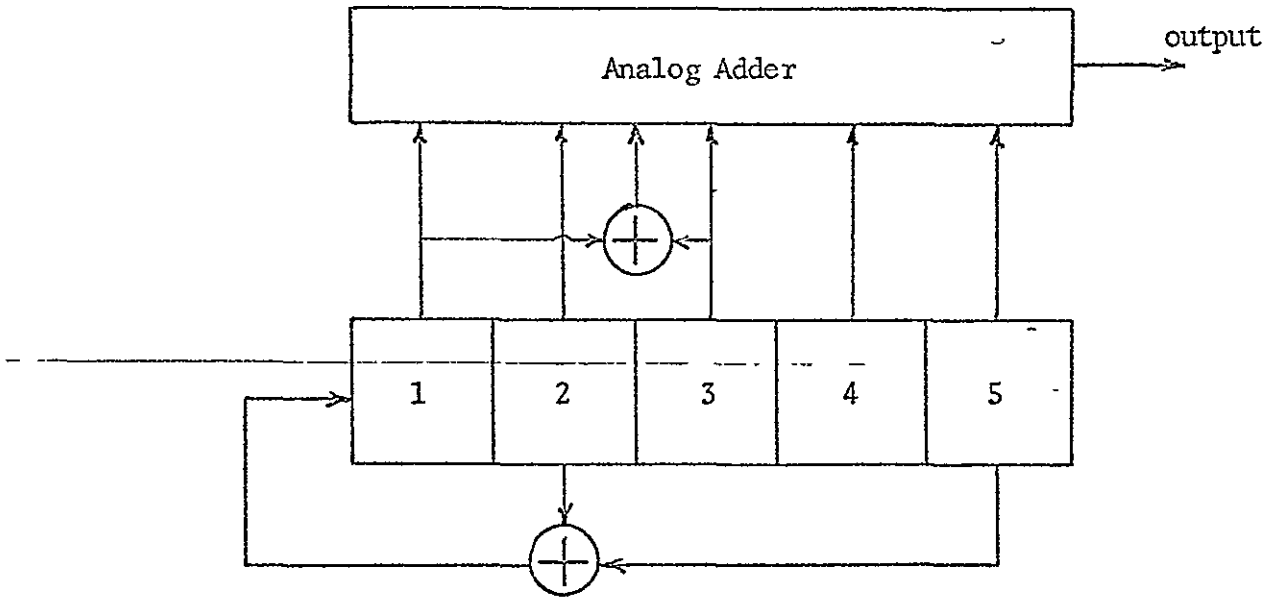


Figure 11. Synthesis of Sum S_6

It is apparent that this basic approach can be used to synthesize any sum discussed in section 3.2. In a particular practical case, it would be necessary to decide whether it would be more desirable to use additional shift register delay stages or additional logic circuits. No general role seems to hold for all situations.

It is also apparent that the above scheme can be used with different weighting factors to produce a wide range of different distributions.

4.0 Infinite Sums

In this section we are concerned with a problem closely related to that of the first three sections. We seek the probability density function of the random variable:

$$X = \sum_{k=1}^{\infty} a^k \xi_k \tag{25}$$

where "a" is a constant, and ξ_k is a random variable. The relation to our earlier work is evident when we write the sum as obtained in section 2.3 as:

$$S = \sum_{k=1}^{j} 2^k \xi_k \tag{26}$$

where ξ_k is a pseudo-random variable which takes values 0 or 1 with nearly equal probability.

Here we extend the problem to the infinite sum where "a" can take any value and ξ_k can be any random variable. The problem is solved with comparative ease for some cases and is not satisfactorily solved for other cases. We start with the easy case ($a = \frac{1}{2}$, ξ_k binary with equal probability of either state) and consider later the more difficult cases.

4.1 The Uniform Case

We consider first the random variable:

$$X = \sum_{k=1}^{\infty} (\tfrac{1}{2})^k \xi_k \tag{27}$$

where ξ_k is a binary random variable which takes the values c and d with equal probability. That is:

$$P(\xi_k = c) = P(\xi_k = d) = \tfrac{1}{2} \text{ , for all } k \tag{28}$$

We assume further that the ξ_k's are independent. This problem has been solved (for $c = 0$ and $d = 1$) by Blake and Thomas[46]. Hence, we are working with the problem of the infinite sum of independent random variables of the form $a^k\xi_k$. We will use a characteristic function approach. It is well-known[40] that the characteristic function of a sum of independent random variables is the product of the individual characteristic functions.

$$\phi_X(w) = \prod_{k=1}^{\infty} \phi_k(w) \tag{29}$$

$\phi_k(w)$, the characteristic function of $a^k\xi_k$, is:

$$\phi_k(w) = \tfrac{1}{2}(e^{jwc/2^k} + e^{jwd/2^k})$$

$$= e^{jw(c+d)/2^{k+1}} \cos \frac{w(d-c)}{2^{k+1}} \tag{30}$$

Applying (29) we obtain:

$$\phi_X(w) = \prod_{k=1}^{\infty} e^{jw(c+d)/2^{k+1}} \prod_{k=1}^{\infty} \cos \frac{w(d-c)}{2^{k+1}} \tag{31}$$

$$= e^{jw(c+d)/2} \frac{\sin \frac{w(d-c)}{2}}{\frac{w(d-c)}{2}} \tag{32}$$

The two infinite products of equation (31) are found as indicated in equation (32), in the appendix. The characteristic function of equation (32) is that of a uniformly distributed random variable with range from c to d. That is:

$$f_X(x) = \begin{cases} \frac{1}{d-c}, & c \leq x \leq d \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

This is a most interesting result, namely that the infinite sum in equation (27) is uniformly distributed. This problem has been considered, though not originally in a probabilitie framework, for a long time. It was shown by Vieta (see Kac[47]) that any real number t, $0 \leq t \leq 1$, can be expressed uniquely as:

$$t = \frac{G_1}{2} + \frac{G_2}{2^2} + \frac{G_3}{2^3} + \cdots \tag{34}$$

where G_j is either 0 or 1 for each j. The problem is discussed in some detail by Kac.

4.2 General Weights (a = 1/2)

We turn now to the general problem of the distribution of:

$$X = \sum_{k=1}^{\infty} a^k \varepsilon_k \tag{35}$$

where a can be any number from 0 to ∞. This problem is not solved here in any closed form. Rather we suggest some of the results and point out problems which are still under consideration.

We are tempted of course to simply apply the transformation approach

used in section 4.1. This leads us to a transform of X given by:

$$\phi_X(w) = e^{jw(c+d)a^2/(1-a)} \prod_{k=1}^{\infty} \cos w(d-c)a^{k+1} \tag{36}$$

It is seen that the first infinite product, involving exponentials, reduces to an easily expressed "phase" term. The second product, however, is not known to have a simple solution. So the method of section 4.1 does not lead to the closed form solution for which we might hope. In fact, further consideration of the range of X suggests some very real problems.

a) Case 1: $a < \frac{1}{2}$

When $a = \frac{1}{2}$, X is a continuous random variable on the range $[c,d]$. However, if $a < \frac{1}{2}$ it is not possible to obtain all values of X between the minimum and maximum. In analogy to the formula of Vieta (see equation 34) consider for following expression.

$$t = e_1a^1 + e_2a^2 + e_3a^3 + \cdots \tag{37}$$

If e_j can be either 0 or 1 the range on t is $0 \leq t \leq \frac{a}{1-a}$. For example, if $a = \frac{1}{4}$, we have:

$$t = \frac{e_1}{4} + \frac{e_2}{4^2} + \frac{e_3}{4^3} + \cdots \tag{38}$$

and the range on t is $[0,\frac{1}{3}]$. However, there are many values of t which cannot be generated by any combination of e_j's. For example, none of the numbers in the open sub-range $(\frac{1}{8},\frac{1}{4})$ can be obtained. In fact, the closed range $[0,\frac{1}{3}]$ has an infinite number of open sub-ranges in which no number can be generated by equation (38) with any combination of e_j's.

Let us turn now to a consideration of the distribution of F. This problem is discussed by Feller.[48] As we shall see it turns out to be essentially meaningless to speak of the probability density function (PDF), so we shall look at the cumulative distribution function (CDF). The CDF has an infinite number of flat regions separated by jumps which are infinitely small. The largest flat region is $(\frac{1}{8},\frac{1}{4})$ which is centered at $X = \frac{1}{6}$ and had a length of $\frac{1}{6}$. There is also a flat region centered about $X = \frac{1}{24}$ with a length of $\frac{1}{24}$, and another of length $\frac{1}{24}$ centered at $\frac{7}{24}$. The CDF for these regions only is plotted in figure 12. The CDF for these regions plus the next set of regions between these is shown in figure 13. As the number of terms in equation (38) taken into account is increased, the size of the jumps in the CDF, F(X), decreases. (See figure 14.) These jumps go to zero in the limit as all terms are considered. Hence, the CDF becomes continuous, and yet no PDF exists. Feller[48] refers to this as a singular distribution.

The example above is for $a = \frac{1}{4}$ but the same type of distribution arises for all $0 < a < \frac{1}{2}$. Work is presently underway to effect a more meaningful representation of the distribution for this case.

Case 2: $\frac{1}{2} < a < 1$

We consider now the range of t in equation (37) under the assumption $\frac{1}{2} < a < 1$. The maximum and minimum values are 0 and $\frac{a}{1-a}$. The range on t is continuous; that is, every number t between 0 and $\frac{a}{1-a}$ can be expressed as in equation (37) with appropriate e_j's. To see this we consider the way in which we generate appropriate values of e_j's for a given t.

The procedure is to subtract from t the largest value of a^j which

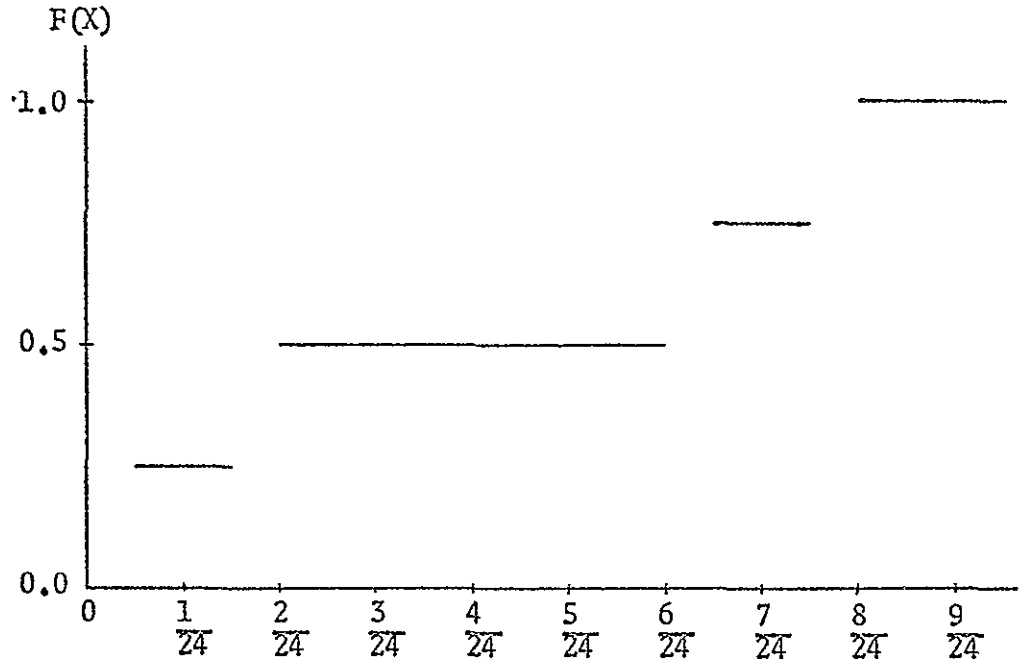


Figure 12. Partial CDF of X (3 open regions shown)

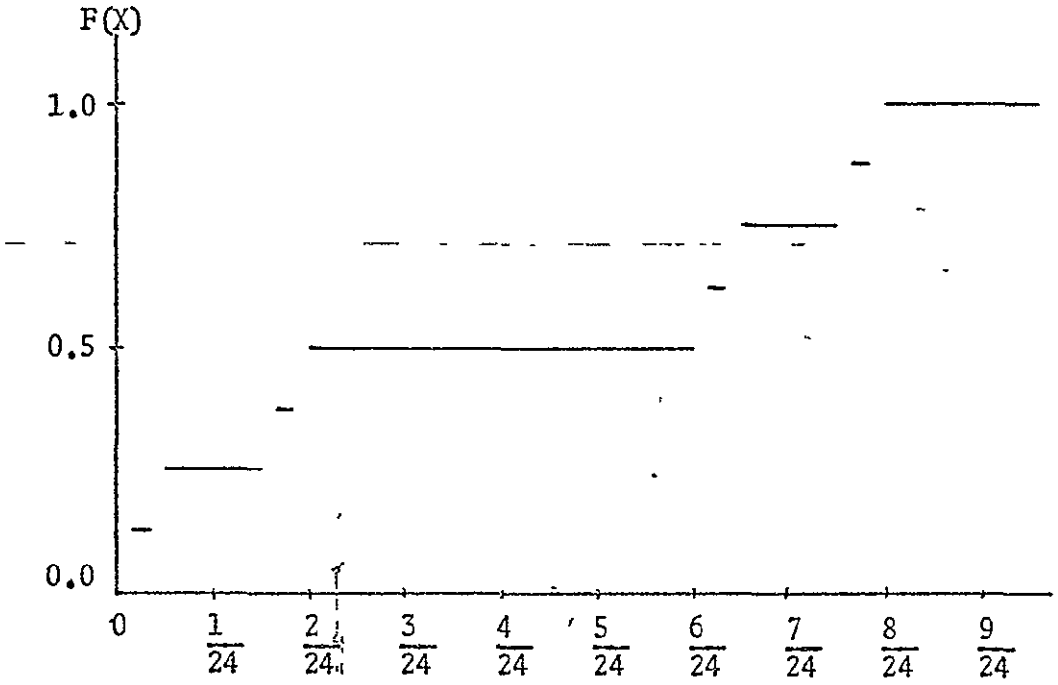


Figure 13. Partial CDF of X (7 open regions shown)

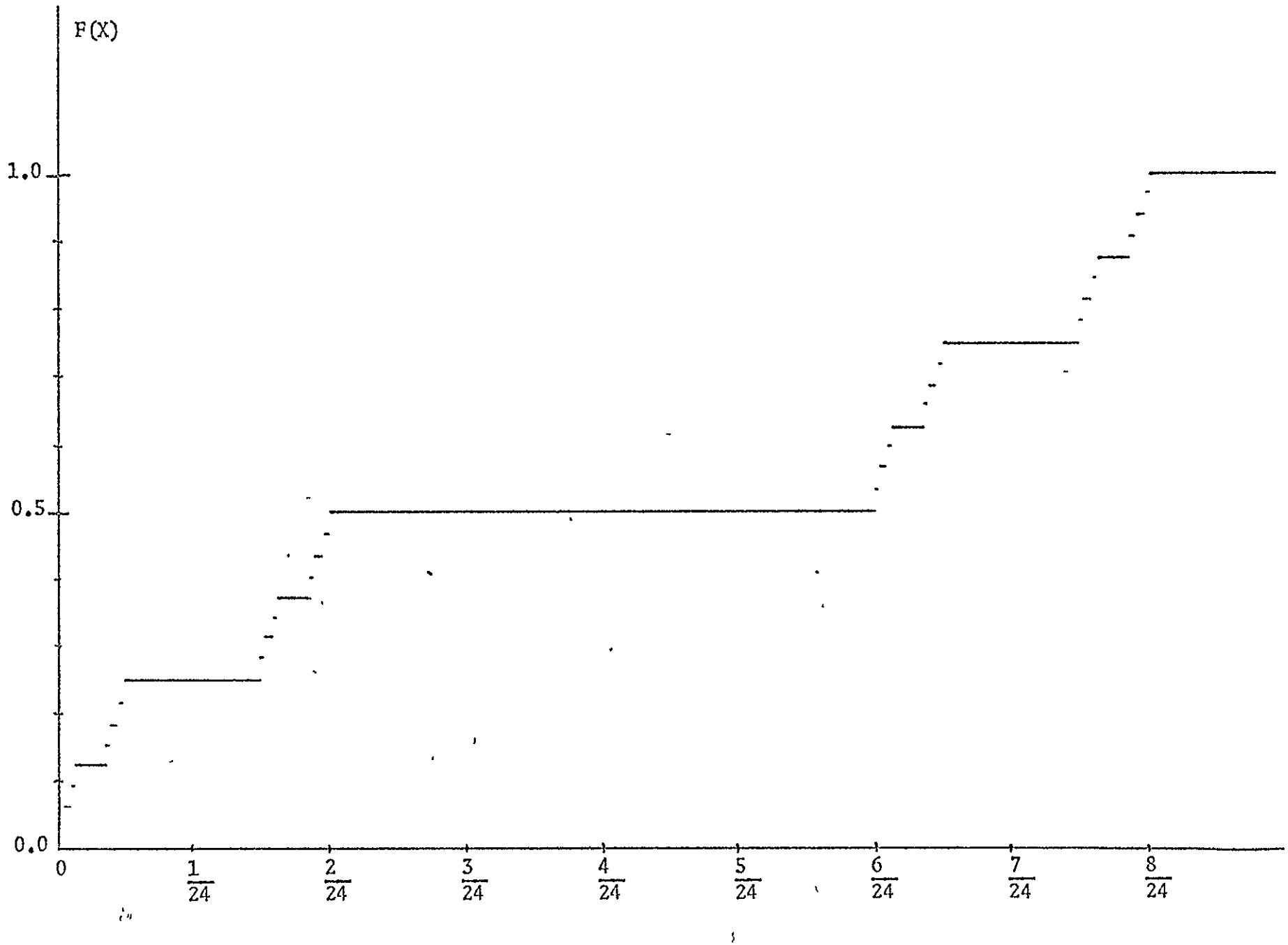


Figure 14. Partial CDF of X. (Many open regions shown.)

is less than t. For this value of j we let $e_j = 1$. For each smaller value of j we let $e_j = 0$. We now try to subtract a^{j+1} from the new difference $t - a^j$. If we can do so without leaving a negative remainder, we let $e_{j+1} = 1$, if not then $e_{j+1} = 0$. We proceed in like manner with a^{j+2} etc., generating new values of e_{j+k} based on whether or not the corresponding a^{j+k} can be successfully subtracted from the new difference. We continue until the last difference is zero. This may be possible only in the limit as the number of terms approaches infinity.

Suppose that at any point in the procedure the original number t or a new difference $t - a^j - a^{j+k} - \cdots$ falls, as it must, within the range:

$$a^n \leq t \leq a^{n+1} \tag{39}$$

A necessary and sufficient condition that the procedure can be carried out with an eventual difference of zero (if only in the limit as the number of terms is unbounded) is that:

$$t - a^{n+1} \leq \sum_{j=n+2}^{\infty} a^j = \frac{a^{n+2}}{1-a} \tag{40}$$

for every t and every n. Equation (40) simply says that after any a^{n+1} subtraction the terms left to subtract must add to a number at least as large as the new difference. The largest value of t in the range given by (39) is a_1^n so the worst case in (40) is:

$$a^n - a^{n+1} \leq \frac{a^{n+2}}{1-a}$$

which reduces to

$$(1-a)^2 \leq a^2 \tag{41}$$

Note that the equality in (41) is met for $a = \frac{1}{2}$, but the inequality fails for $a < \frac{1}{2}$. This is in agreement with the argument of the preceding section. However, the inequality holds for $\frac{1}{2} < a < 1$. Hence, any $0 \leq t \leq \frac{a}{1-a}$ can be realized by a proper selection of e_j's as long as $\frac{1}{2} \leq a \leq 1$. Thus, we have established that the range on t is continuous for $\frac{1}{2} \leq a < 1$.

We now direct our attention to the random variable defined by equation (35) The range on X is the same as the range on t; following the argument above, $0 \leq X \leq \frac{a}{1-a}$, and X is continuous in this range.

We consider next the distribution of X. For $a = \frac{1}{2}$, X was found to be uniformly distributed. More precisely, the expression,

$$X - \sum_{j=1}^{\infty} a^j \xi_j = 0 \tag{42}$$

can be satisfied only by a unique set of ξ_j's for a particular X. That is, there is one and only one combination of ξ_j's in equation (35) for a given value of X in the range $[0,1]$.

If $\frac{1}{2} < a < 1$ this situation no longer holds. As an example let $a = 0.6$. There is only one way to obtain $X = 0$; it is necessary that $\xi_j = 0$ for all j. Likewise, we obtain $X = 1.5$ only if $\xi_j = 1$ for all j. However, we can obtain numbers in between in many ways. The number 0.96 is obtained from the set $\xi_1 = 1$, $\xi_2 = 1$, $\xi_j = 0$ for $j > 2$. It can also be obtained by letting $\xi_1 = 1$, $\xi_2 = 0$ and ξ_j (for $j > 2$) take on any appropriate values to add to 0.36. That this is possible is demonstrated by the fact that:

$$\sum_{j=2}^{\infty} (0.6)^j = 0.54 , \tag{43}$$

which is greater than 0.36 .

Hence, we have shown that for $\frac{1}{2} < a < 1$ the random variable X is continuous and is not uniformly distributed. Work is continuing on the determination of this distribution.

We close this section with the comment that we believe that as "a" approaches 1, the distribution of X will probably become Gaussian-like.

Case 3: $a = 1$

For this case consider equation (35) with finite range.

$$X = \sum_{k=1}^{N} \xi_k \tag{44}$$

This is simply the problem of N independent trials and the distribution of X is of course binomial of order N . As N increases without bound the distribution of X asyptotically approaches a Gaussian distribution, according to the DeMoivre-Laplace Theorem.

Case 4: $a = 2$

Again, we let the number of terms be finite

$$X = \sum_{k=0}^{N} 2^k \xi_k \tag{45}$$

As was shown in section 2.3, the resulting distribution of X is uniform.

Case 5: $1 < a < \infty$, $a \neq 2$

Little work has been done on the wide range of "a" greater than 1.

There is, however, an intriguing symmetry or mirror effect of the type of distribution with respect to $a = 1$. Note again that for $a = 1$ the distribution is approximately Gaussian. For $a = \frac{1}{2}$ and $a = 2$, the distribution is uniform. There is a strong temptation to suggest an analogy between the ranges $(0,\frac{1}{2})$ and $(2,\infty)$ We are continuing to explore these relationships. Figure 15 summarizes the above 5 cases, including the speculations suggested for case 5.

Figure 15. Types of Distributions

[1] Golomb, S. W., Shift Register Sequences, Holden-Day, San Francisco, 1967.

[2] Zierler, N., "Linear Recurring Sequences", S.I.A.M., Vol. 7 #1, Mar. 1959 pp. 31-48.

[3] Huffman, D. A., "The Generation of Impulse-Equivalent Pulse Trains", IRE Trans. Inf. Th., Sept. 1962, pp. 10-14.

[4] Schmandt, F. D. et al, "Crosscorrelation of a Maximal Length Binary Sequence and its Subsequences", RADC-TR-67-261, August 1967.

[5] Peterson, W. W., Error-Correcting Codes, J. Wiley, N.Y. 1961.

[6] Golomb, S. W., Digital Communications with Space Applications, Prentice-Hall, Englewood Cliffs, N. J. 1964.

[7] Kramer, C., "A Low-Frequency Pseudo-Random Noise Generator", Electronic Engineering, July 1965, pp. 465-467.

[8] Wolf, J. K., McLaughlin, R. G., Schmandt, F. D., "The 'Shift and Add Property' of Maximal Length Binary Sequences Using Cyclotomic Polynomials", RADC-TDR-64-66, March 1961.

[9] David, M., "Synthese et Proprietes D'UnGenerateur de Nombres Aleatoires", C.D.V. 621.373.44 Revue MBLE, 7,249-265, Dec. 1964.

[10] Roberts, P. D., "Generation of Delayed Replicas of Maximal-Length Binary Sequences", Proc. IEE Vol. 112, #4, April 1965, pp. 702-704.

[11] Davies, A. C., "Delayed Versions of Maximal-Length Linear Binary Sequences", Electronics Letters, Vol. 1, #3, May 1965, pp. 61-62.

[12] Tsao, S. H., "Generation of Delayed Replicas of Masimal-Length Linear Binary Sequences", Proc. IEE, Vol. 111, #11, Nov. 1964, pp. 1803-1806.

[13] Douce, J. L., "Delayed Versions of m Sequences", Electronics Letters, 14 June 1968, Vol. 4, #12, p. 254.

[14] McFadden, J. A., "The Probability Density of the Output of an RC Filter When the Input is a Binary Random Process", IRE Trans. on Inf. Theory, Dec. 1959, pp. 174-178.

[15] Cumming, I. C., "Autocorrelation Function and Spectrum of a Filtered, Pseuo-random Binary Sequence", Proc. IEE, Vol. 114, #9, Sept. 1967, pp. 1360-1362.

[16] Roberts, P. D., "Statistic Properties of Smoothed Maximal-Length Linear Binary Sequence", Proc. IEE, Vol. 113, #1, Jan. 1966, pp. 190-196.

[17] Davies, A. C., "Digital Filter of Binary Sequences", Elect. Letters, July 1967, Vol. 3, #7, pp.

[18] Matthews, S. B., "Generation of Pseudorandom Noise Having a Gaussian Spectral Density", IEEE Tran. on Computers, April 1968, pp. 382-385.

[19] Parker, H. A., "Choice of Pseudorandom Binary Signals for System Identification", Elect. Letters, Nov. 1967, Vol. 3, #11, pp. 524-526.

[20] Davies, W.D.T., "Identification of a System in the Presence of Low-Frequency Drift", Elect. Letters, Sept. 1966, Vol. 2, #9, pp. 327-329.

[21] Brunner, R., "What Makes Test Instruments Tick?", Electronic Design 20, Sept. 27, 1967, pp. T6-T11.

[22] "Testing With Pseudo-Random and Random Noise", Hewlett-Packard, 1967.

[23] Anderson, G. C., Finnie, B. W., Roberts, G. T., "Pseudo-Random and Random Test Signals", Hewlett-Packard, 1967.

[24] Shepertycki, T. H., "Telemetry Error Measurements Using Pseudo-Random Signals", IEEE Trans. on Space Elec. and Telemetry, Sept. 1964, pp. 111-115.

[25] Barna, A., "Pseudorandom Frequency Modulation in Range-Doppler Radar", Stanford Electronics Lab., SU-SEL-68-046, May 1968.

[26] Hampton, R. L., "A Hybrid Analog-Digital Pseudo Random Noise Generator", Simulation, March 1965, pp. 179-190.

[27] Balza, C., Fromageot, A., Maniere, M., "Four-Level Pseudorandom Sequences", Elect. Letters, July, 1967, #7, pp. 313-315.

[28] Brown, R. F., Goodwin, C. C., "New Class of Pseudorandom Binary Sequences", Elect. Letters, May 1967, Vol. 3, #5, pp. 198-199.

[29] Hampton, R.L.T., "Experiments Using Pseudo-Random Noise", Simulation, April 1965, pp. 246-254.

[30] Darnell, M., "Synthesis of Pseudorandom Signals Derived from p-level m Sequences", Electronics Letters, Nov. 1966, Vol. 2, #11, pp. 428-430.

[31] Braasch, R. H., "The Distribution of (n-m) Terms for Maximal Length Linear Pseudo-Random Sequences", IEEE Trans. on Inf. Th., July 1968, pp. 607-608.

[32] Swick, D. A., "Wideband Ambiguity Functions of Pseudo-Random Sequences: An Open Problem", IEEE Trans. on Inf. Th., July 1968, pp. 602-603.

[33] Lindholm, J. H., "An Analysis of the Pseudo-Randomness Properties of Subsequences of Long m-Sequences", IEEE Trans. on Inf. Th., July 1968, pp. 569-576.

[34] Davies, W.D.T., "Using the Binary Maximum Length Sequence for the Identification of System Dynamics, Proc. IEE, Vol. 114, #10, October 1967, pp. 309-312.

[35] Davio, M., "Random and Pseudorandom Number Generators", Electronic Engineering, Sept. 1967, pp. 558-559.

[36] Kramer, C., "A Low-Frequency Pseudo-Random Noise Generator", Elect. Engineering, July 1965, pp. 465-467.

Bibliography

[37] Gilson, R. P., "Some Results of Amplitude Distribution Experiments on Shift Register Generated Pseudo-Random Noise", IEEE Tran. on Computers, Dec. 1966, pp. 926-927.

[38] White, R. C., "Experiments with Digital Computer Simulation of Pseudo-Random Noise Generators", IEEE Tran. on Electronic Computers, June 1967, pp. 355-357.

[39] Parzen, E., Modern Probability Theory and Its Applications, John Wiley & Sons, New York 1960.

[40] Bracewell, R., The Fourier Transform and Its Application, McGraw-Hill, New York, 1965.

[41] Healy, T., "Convolution Revisited", IEEE Spectrum, April, 1969, pp. 87-93.

[42] Healy, T., "The Synthesis of Distributions of Pseudo-Random Variables", 2nd Asilomar Conference on Circuits and Systens, Oct. 30 - Nov. 1, 1968.

[43] Davies, A., "Probability Distributions of Waveforms Generated by a Digital Technique", Electronics Letters, Vol. 4, 20 Sept. 1968, pp. 421-423.

[44] Aggarwal, G., "Sampled-Data Multichannel Telemetry Using Pseudorandom Sequences", IEEE Proc., Vol. 57, No. 3, p. 366.

[45] Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

[46] Blake, I. and Thomas, J., "The Linear Random Process", IEEE Proc. Vol. 56, No. 10, pp. 1696-1703.

[47] Kac, M., Statistical Independence in Probability, Analysis and Number Theory, John Wiley and Sons, Inc., 1959.

[48] Feller, W., An Introduction to Probability Theory and Its Applications, Vol. II, John Wiley and Sons, New York, 1966.

6.0 Appendix: Infinite Products

1) Simplify $\prod_{k=1}^{\infty} e^{jw(c+d)/2^{k+1}}$

$$\prod_{k=1}^{\infty} e^{jw(c+d)/2^{k+1}} = e^{ja(c+d)\left(\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots\right)}$$

$$= e^{jw(c+d)\frac{1}{4}\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n}$$

$$= e^{jw(c+d)/2} \qquad \text{(A-1)}$$

Since $\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n = 2$

2) Simplify $\prod_{k=1}^{\infty} \cos\frac{w(d-c)}{2^{k+1}}$

$$\prod_{k=1}^{\infty} \cos\frac{w(d-c)}{2^{k+1}} = \cos\frac{\frac{w(d-c)}{2}}{2} \cos\frac{\frac{w(d-c)}{2}}{4} \cos\frac{\frac{w(d-c)}{2}}{8} \cdots$$

$$\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}} = \cos\frac{\theta}{4} \sin\frac{\theta}{4}\Big/\frac{\theta}{4} \qquad \text{(A-2)}$$

$$= \cos\frac{\theta}{4}\cos\frac{\theta}{8}\sin\frac{\theta}{8}\Big/\frac{\theta}{8}$$

$$= \cos\frac{\theta}{4}\cos\frac{\theta}{8}\cos\frac{\theta}{16}\sin\frac{\theta}{16}\Big/\frac{\theta}{16} \qquad \text{(A-3)}$$

As the trigonometric identity is continually applied to the last $(\sin\gamma)/\gamma$ term, that term approaches unity and the infinite product becomes:

$$\frac{\sin\frac{\theta}{2}}{\frac{\theta}{2}} = \prod_{k=1}^{\infty} \cos\frac{\theta}{2^{k+1}} \qquad \text{(A-4)}$$

which suggests that

$$\prod_{k=1}^{\infty} \cos \frac{w(d-c)}{2^{k-1}} = \frac{\sin \frac{w(d-c)}{2}}{\frac{w(d-c)}{2}}$$