Research Report

No. 78-4

FINAL REPORT ON
RESEARCH INTO NONLINEAR PSEUDORANDOM
SEQUENCE GENERATION AND CHARACTERIZATION

PREPARED FOR
THE DEPARTMENT OF COMMUNICATIONS
UNDER DSS CONTRACT No. OSU77-00146

BY

S.E. TAVARES P.H. WITTKE

Queen's University at Kingston
Department of Electrical Engineering

FINAL REPORT

RESEARCH INTO NONLINEAR PSEUDORANDOM SEQUENCE GENERATION AND CHARACTERIZATION

Prepared for

The Department of Communications
Under DSS Contract No. OSU77-00146

by

S.E. Tavares and P.H. Wittke
Principal Investigators
Department of Electrical Engineering
Queen's University
Kingston, Ontario, Canada

Scientific Authority:

Dr. J.L. Pearce, Communications Research Centre

Research Assistant:
Mr. D. Caswell

March 1978

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The application of shift register sequences to communications is well documented. For a recent (1976) reference, see the IEEE Press Book edited by R.C. Dixon entitled, "Spread Spectrum Techniques" [2]. Not surprisingly, most of the attention has been focused on maximal length, linear feedback shift register sequences [7] [8] which will be called linear PN sequences here. Two reasons for their popularity are: (i) they have a number of desirable noise-like properties and (ii) their linearity allows for ease of implementation. Golomb's book [8] is a standard reference, and a recent review of the linear theory has been published by MacWilliams and Sloane [12]. See also Chapter 14 in their book [13].

The non-linear sequences were also extensively discussed by Golomb [8] and represents the state of knowledge at that time. Since then, a number of papers have been published [3-6], [9], [10-12], [15-19], but generally the literature is rather sparse. In fact, the paper by Key [9] is really about linear sequences. The nonlinear shift registers in his paper all have linear equivalents which generate the same sequences. A motivation in his work is that the nonlinearity (in the feedforward only) allows fewer stages (delay elements or flip-flops) in the shift register than the linear equivalent.

There are a few explanations which may explain the apparent lack of activity in nonlinear shift register sequences. One of these is the feeling that linear sequences have already achieved all the desirable properties for the anticipated applications. A second reason is that the word "nonlinear" may have scared off many potential investigators.

A third reason is that the view that nonlinear shift-registers would be hard or expensive to build.

Perhaps the easiest to dispose of is the third reason. The revolutionary developments in digital hardware with the accompanying reduction in costs allow much more flexibility in hardware configurations. The second reason mentioned was the "nonlinear" barrier. For the reason stated above, the practical concerns have been removed but there remains the theoretical anxiety. This is quite natural, since the assumption of linearity (mod 2 addition or X-OR) allows a considerable body of mathematical analysis to be employed. However, we still have some analytical tools if we allow nonlinearities (mod 2 multiply or logical AND, and inversions).

To justify the effort in analyzing nonlinear sequences, we will review briefly, in the next section, some of the desirable properties of a pseudo-noise (PN) sequence. However, since these sequences have several applications, some of these requirements may conflict.

RANDOMNESS PROPERTIES OF SEQUENCES

Golomb ([8], Chap. III) gives a discussion of what a "typical" random binary sequence should look like. A familiar example would be a typical sequence generated by tossing an unbiased coin. More formally, consider a binary sequence of 0's and 1's. We will call the sequence random if 0's and 1's occur with equal probability and successive symbols occur independently of all previous symbols. Let

$$\{a_i\} = \{a_0, a_1, a_2, \ldots\} \tag{1}$$

be such a binary sequence, then

$$p(a_{i+1} | a_i, a_{i-1}, \ldots, a_0) = p(a_{i+1}) = 1/2 \tag{2}$$

for any a_{i+1}, $i = 0,1,\ldots$. The notation $p(x|y)$ means the probability of x given y.

Now, although all random binary n-tuples

$$a_0 a_1 a_2 \cdots a_{n-1}$$

are equally likely, the n-tuples with $n/2$ ones (assuming for the moment that n is even) are the most numerous. In fact, there are exactly $\binom{n}{\frac{n}{2}}$ of them. Generally we would expect a random sequence to have roughly an equal number of zeros and ones. Hence, we state the first randomness requirement as follows:

(i) the sequence should have an equal number of zeros and ones.

Even if the above condition is satisfied, the sequence may still not look random. For example, the sequence may consist of a block of zeros followed by a block of ones. We observe that short

runs (of zeros or ones) are more probable than long runs. In particular, a run of length r is twice as probable as a run of length $r+1$. For example, ...XX1001XX... is twice as probable as ...XX10001XX..., where X may be zero or one. The first is a run of zeros of length two and the second is a run of zeros of length three.

Another observation is that all patterns of a given length are equiprobable and we would expect them to occur an equal (or nearly equal) number of times. Although the above observations are related, we state the second and third randomness properties:

(ii) there should be twice as many runs of length r as there are of length $r+1$;

(iii) all patterns of equal length should occur an equal number of times.

Finally there is the autocorrelation property. To avoid ambiguity, we note here that we are discussing the periodic autocorrelation as opposed to the aperiodic autocorrelation. Golomb ([8], p. 25) says that "...random sequences possess a special kind of autocorrelation function, peaked in the middle and tapering off rapidly at the ends." We recall that all binary n-tuples are equiprobable and hence what we are interested in is the nature of the autocorrelation of a "typical" n-tuple. Golomb does not formally justify his observation about the autocorrelation, and it is subject to various interpretations. For example, would the autocorrelation of one of the typical sequences meet his criterion, or should we average over the set of "typical" sequences or perhaps average over all binary n-tuples. One may even average the

autocorrelation for all shifts for some typical sequence. We raise this matter because we feel that the notion of the autocorrelation of a random sequence deserves further investigation and clarification.

To develop our notation, let

$$c(\tau) = \sum_{i=0}^{n-1} a_i a_{i+\tau}$$

be the cyclic autocorrelation of the binary n-tuple

$$(a_0, a_1, a_2, \ldots, a_{n-1})$$

where the subscripts are reduced modulo n. [Note that in our definition of autocorrelation we do not divide by the length of the sequence in contrast to Golomb.] Golomb ([8], p. 26) defines his third random postulate as: "the autocorrelation function is two-valued." Mathematically, this can be stated as

$$c(\tau) = \begin{cases} n & \text{when } \tau = 0 \\ K & \text{when } \tau \neq 0 \end{cases} ,$$

where K is some constant. We will not adopt this definition as one of our random criteria since we feel it is somewhat arbitrary. Our view will be that Golomb's criterion is intrinsically interesting and highly desirable for some applications, but is not necessarily a test of randomness. Indeed, Golomb gives examples of sequences which satisfy his third randomness postulate but do not look like a typical random sequence. For instance, he points out that the sequence ([8], p.88)

0000010001101

of length 13 has

$$c(\tau) = \begin{cases} 13 & \text{if } \tau = 0 \\ 1 & \text{if } \tau \neq 0 \end{cases}$$

although it has 9 zeros and only 4 ones. Further, the pattern distribution in this sequence is very poor. During this discussion we will return to the autocorrelation problem from time to time and hopefully will shed some light on the subject.

Before closing, we mention another test for randomness based on the conditional probability as defined in (2). The reader may recall that a Markov source of order s is one whose current output depends on the most recent past s outputs. By collecting the appropriate data from a given binary sequence, one may estimate the conditional probabilities for all the 2^s possible "states" of the Markov source. If the sequence $\{a_i\}$ under examination is truly random (independent and equally likely symbols) then all these conditional probabilities should be equal to 0.5. If there is a departure from 0.5, it will reveal a bias or some memory in the sequence.

When we apply the test to linear PN sequences of length $n = 2^m-1$ and test for Markov dependence of $m-1$ or less (i.e., $s \leq m-1$) there is a bias. However, if the same test is applied to a deBruijn sequence, no bias is observed.

In the next section, we will introduce some interesting binary sequences of length $n = 2^m$, called deBruijn sequences [1], and compare them with the linear pseudo-random (PN) sequences of length 2^m-1.

SHIFT REGISTER SEQUENCES OF LENGTH 2^m

There is a class of binary sequences, known as deBruijn sequences [1] [8], which have length 2^m and can be generated by certain non-linear feedback shift registers having m stages. Note that these sequences are one digit longer than the well-known linear PN sequences of length 2^m-1. We observe that both types of sequences are generated by m-stage shift registers, but they differ in the nature of the feedback function. Recall that by "linear function" we mean a modulo 2 addition (exclusive-OR) of the appropriate logical variables. A non-linear function will contain in addition modulo 2 (mod 2) multiplication (logical AND), and/or inversions. We will denote mod 2 addition by $+$ and mod 2 multiplication by juxtaposition.

At first it appears that by allowing nonlinearities we have gained merely a sequence one digit longer. However we wish to observe that there are some important advantages as listed below:

(i) In the deBruign sequences, the number of ones is exactly equal to the number of zeros;

(ii) every pattern of length m appears exactly once;

(iii) the deBruijn sequences are a much larger class of sequences than the linear PN sequences for a given m.

We will now discuss these points in turn.

Concerning property (i), the reader may recall that for linear PN sequences, the ones outnumber the zeros by one. This imbalance may sometimes be undesirable, e.g., for some applications it may result in a resultant DC term.

Our comment on property (ii) is related to the above in that as a

result of the missing zero, linear PN sequences do not contain the all-zero m-tuple. For certain pseudo-random applications it would be desirable to have the symmetry displayed by the corresponding deBruijn sequence (since they contain the missing m-tuple).

Property (iii) has important ramifications for security applications. For a given m, the number of linear PN sequences is given by [8],

$$\lambda(m) = \phi(2^m-1)/m$$

where $\phi(\ )$ is the Euler ϕ-function (see [8], p. 38 for its definition). We make the simple observations that if p is a prime, then

$$\phi(p) = p - 1$$

and in general, for any positive integer $q > 1$,

$$\phi(q) \leq q - 1 \quad .$$

From the above, we conclude that

$$\lambda(m) \leq (2^m-2)/m \ .$$

On the other hand, it is known [1],[8] that the number of deBruijn sequences of length 2^m is given by

$$\delta(m) = 2^{(2^{m-1}-1)} = 2^{2^{m-1}}/2^m \quad .$$

We note that $\delta(m)$ grows like a double exponential. For a quick comparison we give a short table of these two functions. Table I makes it quite clear how rapidly $\delta(m)$ grows. We quote Golomb ([8], p. 111), "This astronomical increase in the number of good codes justifies, in itself, the quest for practical nonlinear shift registers."

TABLE I

m	2^m	$\lambda(m)$	$\delta(m)$
1	2	1	1
2	4	1	1
3	8	2	2
4	16	2	16
5	32	6	2048
6	64	6	67,108,864

DECIPHERING LINEAR AND NONLINEAR PN SEQUENCES

An interesting question that can be posed about deterministic periodic sequences is: How may consecutive digits of the sequence must we observe before we can uniquely determine the sequence or equivalently, its generator? If the sequence is known to be a linear PN sequence of period 2^m-1, then it is known that $2m$ consecutive digits are sufficient [20]. An informal argument to justify this is as follows. We know that the feedback function is a linear combination of some subset of the m contents of the shift register. The unknowns are the locations of the feedback taps. The first m digits specify the state of the shift register and then the second m digits specify m linear equations in m unknowns which can be solved for the tap locations.

In comparison, it requires about 2^{m-1} consecutive digits to determine uniquely which deBruijn sequence of length 2^m is being observed. Notice that we must observe half the length of the sequence to decide which deBruijn sequence it is. An intuitive justification of this result is as follows. Borrowing some ideas from information theory, we assume that all the sequences are equally likely, and hence we need $\log_2\delta(m)$ bits of information to uniquely specify the sequence. Now

$$\log_2\delta(m) = 2^{m-1}-m$$

and if we allow m bits to initially determine the state of the shift register we end up with 2^{m-1} as desired.

If we apply the above arguments to the linear PN sequences, we have

$$\log_2\lambda(m) \le \log(2^m-2) - \log m$$
$$< \log 2^m = m$$

i.e., $$\log_2\lambda(m) < m \quad .$$

If m is large and prime (the most difficult case) then

$$\log_2\lambda(m) \lesssim m$$

where $\lesssim$ means less than but approximately equal to. As before, if we allow m digits to acquire the state of the shift register, we see that $2m$ consecutive digits are sufficient to determine a linear PN sequence of length 2^m-1. A passing thought is that on the average, the number of digits required to determine a linear PN sequence is

$$m + \log_2\lambda(m)$$

which is strictly less than $2m$.

Returning to the deBruijn sequences, we give an explicit procedure to determine the feedback function by observing part of the sequence. Golomb ([8], p. 116) has shown that the feedback function of a deBruijn sequence may be decomposed into a nonlinear function of $(m-1)$ of the variables mod 2 added to the mth variable. Hence we can write

$$f_m(x_{m-1},x_{m-2},\ldots,x_1,x_0) = f_{m-1}(x_{m-1},x_{m-2},\ldots,x_1) + x_0$$

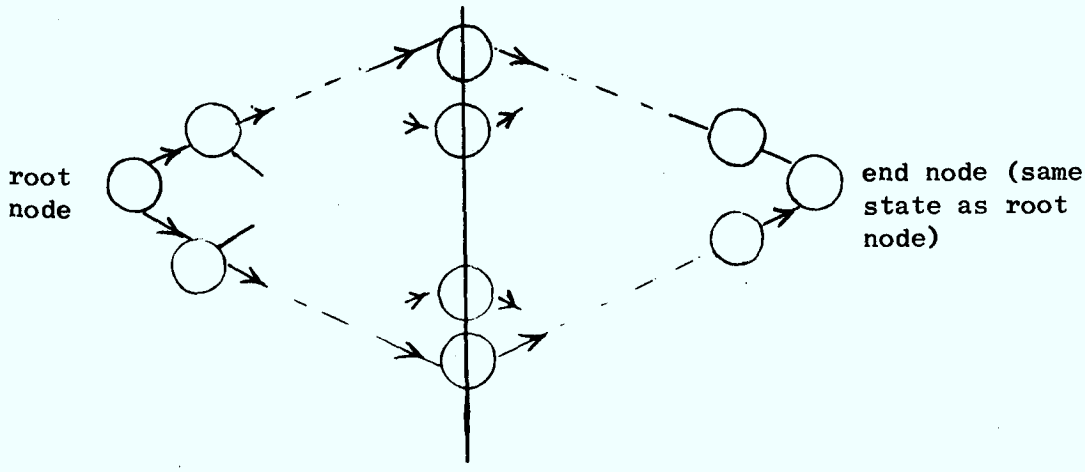
From switching algebra we know that f_{m-1} may be expressed as the mod 2 sum of minterms, each minterm consisting of the mod 2 product (AND) of the $(m-1)$ variables or their complements. Since each variable may be complemented or not, there are 2^{m-1} possible minterms. Once the state of the shift register has been acquired, each succeeding digit will determine whether or not a given minterm is present in f_{m-1}.

If there are no repeats, this requires $2^{m-1}+1$ digits to specify f_{m-1} and hence f_m. If there are repeats, we must observe more digits until all 2^{m-1} distinct (m-1)-tuples appear as arguments of f_{m-1}. This algorithm does not take advantage of all the known properties of deBruijn sequences but it is easy to apply. We now show that there exists an algorithm which will always uniquely identify the sequence after observing $2^{m-1}+m$ consecutive digits.

We prove the following result for deBruijn sequences:

If two deBruijn sequences of length $n = 2^m$ coincide for $2^{m-1}+m$ consecutive digits then they are identical.

To show this, take any m-tuple as a reference state (root node) and grow a tree of successor states. If all the paths in the tree are valid deBruijn sequences, they will all merge at the reference state after 2^m transitions. The state-diagram generated is a branching tree-like structure that grows for a while then starts to remerge until it collapses back to the original node. We now make the simple observation that the mirror image of a deBruijn sequence is also a deBruijn sequence. This means that our state diagram (which is somewhat diamond shaped) is invariant under a left to right flip or a rotation about a vertical axis through its center as shown.



The important conclusion we draw from the above argument is that the state diagram does not branch beyond the midline. It also means there are as many states at the midline as there are distinct deBruijn sequences. Hence any two paths starting at the root node that are not distinct at the midline will not diverge later are thus the same path. This means that if we observe a path from the root node (which is arbitrary) to the midline, we have uniquely determined which deBruijn sequence it is. If we count the number of consecutive digits involved, this comes out to be $2^{m-1}+m$.

It is worth pointing out that the above procedure is not, as it stands, a practical algorithm, since it involves table look-up in a large table (if $n > 32$). Using the left to right symmetry of the state diagram it can be seen that the right half of a deBruijn sequence can be obtained from the left half of the diagram by reading backward from the midline to the root node of the reciprocal sequence. This still requires some easily implementable procedure for pairing up these half sequences, and seems to be worth further investigation.

THE AUTOCORRELATION FUNCTION OF DeBRUIJN SEQUENCES

In several applications, the autocorrelation function of a sequence is of considerable interest, e.g., ranging, timing and synchronization. The autocorrelation function of the linear PN sequences have come to be regarded as ideal. Indeed, sequences which have two-valued autocorrelations, which include the linear PN sequences, are called "perfect sequences" by Golomb ([8], p. 88). It is well-known that a linear PN sequence of length $n = 2^m-1$ has autocorrelation function

$$c(\tau) = \begin{cases} n, & \text{when } \tau = 0 \\ -1, & \text{when } \tau \neq 0 \end{cases}$$

It is clear that if we are trying to synchronize using a periodic sequence with the above autocorrelation with high central peak and no sidelobes, then the chances of false synchronization are relatively small. It is usually assumed that the correlation takes place in the presence of noise.

From the outset, we acknowledge that deBruijn sequences do not have this kind of autocorrelation function. In fact, Golomb ([8], p. 124) has shown that for a deBruijn sequence of length $n = 2^m$,

$$c(\tau) = \begin{cases} n = 2^m, & \text{when } \tau = 0 \\ 0 \quad , & \text{when } |\tau| \leq m-1 \\ \neq 0 \quad , & \text{when } |\tau| = m \end{cases} .$$

Golomb has further shown that if the deBruijn sequence is obtained by inserting the extra zero into a linear PN sequence, then

$$c(\tau) = -4 \quad , \text{ when } |\tau| = m \quad .$$

In general, for $|\tau| > m$ the autocorrelation function is not known.

We will list a few known properties, the proofs of which are given in Appendix I.

(i) $c(\tau) = \pm 4q$, where q is zero or a positive integer.

Let $n = n_0+n_1$, where n is the length of a binary sequence, n_0 is the number of zeros and n_1 is the number of ones, then

$$\bar{c} \triangleq \frac{1}{n} \sum_{\tau=0}^{n-1} c(\tau) = (n_0-n_1)^2/n \quad ,$$

where $\bar{c}$ is the average of $c(\tau)$ over all cyclic shifts. It follows immediately that for deBruijn sequences.

(ii) $\bar{c} = 0$

Note that this is the smallest value possible since $\bar{c}$ is a nonnegative quantity. For a linear PN sequence of length $n = 2^m-1$, $n_1-n_0 = 1$, and hence

$$\bar{c}_{PN} = 1/n \quad .$$

Thus at least in this average sense, deBruijn sequences have a lower (better?) average autocorrelation than linear PN sequences.

We now recall that except for the extended PN sequences (adding the missing zero) the value of $c(m)$ is unknown. Golomb merely states that $c(m) \neq 0$ for a deBruijn sequence of length 2^m. We have been able to specify $c(m)$ in terms of the number of minterms in the feedback function $f_{m-1}(x_{m-1},x_{m-2},\ldots,x_1)$ where $f_m = f_{m-1}+ x_0$. Explicitly, if the number of minterms in f_{m-1} is given by k, then

(iii) $c(m) = 2^m- 4k$.

Golomb ([8], p. 122) has shown that for a deBruijn sequence, the number of minterms k must be odd. Since $c(m)$ is the nearest non-zero

value of $c(\tau)$ to the central peak, it may be desirable to keep it as small as possible in some applications. However, it cannot be zero, so the smallest we can achieve is $c(m) = \pm 4$, in view of property (i).

At this point we would like to suggest that instead of concentrating solely on autocorrelations that look like those of linear PN sequences, we should examine the whole autocorrelation function and thus determine a "signature" for that particular sequence. Depending on their location, some off-center peaks or sidelobes may contribute to a more rapid acquisition of synchronization. As an example, consider the autocorrelation of the deBruijn sequence of length $n = 8$

SEQUENCE: 11101000

AUTOCORRELATION: 8,0,0,-4,0,-4,0,0,8,0,0,-4 ,

where we have specified the autocorrelation beyond one cycle. For comparison we give the linear PN sequence of length seven and its autocorrelation

SEQUENCE: 1110100

AUTOCORRELATION: 7,-1,-1,-1,-1,-1,-1,7,-1,...

Using the autocorrelation of the PN sequence, the receiver has no idea where it is until synchronization is actually obtained, shifting a digit at a time. In contrast, the deBruijn sequence has a more complex autocorrelation pattern which can give clues to the receiver in its search for synchronization. For example, if it observes 0,0,-4, then it knows uniquely where it is relative to true synchronization. It is recognized that such observations are usually made in the presence of noise and hence the computed autocorrelation values are only estimates. However, to ignore these estimates is throwing away useful information.

In some communication systems (e.g., PSK) the polarity of the received sequence may be reversed (i.e., zeros become ones and vice versa). In this case all the autocorrelation values will be multiplied by minus one. As an illustration, the autocorrelation for the eight digit deBruijn sequence becomes

$$c(\tau): \quad -8,0,0,4,0,4,0,0,-8,\ldots$$

Here also, if the receiver knows the signature of the sequence, it will recognize the inversion.

If we allow sidelobes of significant amplitude, a parameter of concern is their spacing. It is clear that the main lobes (in phase peaks) are spaced n digits apart. Hence, if a large positive sidelobe occurs midway between the main lobes it will present a serious risk of false synchronization. For example, a certain deBruijn sequence of length $n=16$ has the autocorrelation shown below

$$c(\tau): \quad 16,0,0,0,-12,0,0,0,8,0,0,0,-12,0,0,0,16$$

The sidelobe $c(8) = 8$ is a potential hazard for false synchronization, exactly $n/2$ digits away from true synchronization. Fortunately, there are many deBruijn sequences of a given length and thus we can delete those with autocorrelations of the above type. Notice that the large sidelobes with amplitude -12 are spaced eight digits apart and are of opposite polarity to the central lobes and thus are a lesser threat even in the presence of polarity inversion of the received sequence.

For the whole class of deBruijn sequences, we have not been able to put a strong upper bound on the height of the sidelobes. However,

they form such a large class of sequences that this may not be surprising. The computer data obtained so far indicates that for any given n, there are some deBruijn sequences with small sidelobes and some with large sidelobes. Perhaps there is some way of defining a subset of the deBruijn sequences that exclude most or all of the sequences with large sidelobes.

Using a simple argument, we have shown that the largest possible positive sidelobe c_M is bounded by

$$c_M \leq n - 2\left\lceil \frac{n}{m} \right\rceil$$

where $\lceil x \rceil$ means if x has a fractional part, we round up to the next highest integer. The argument goes as follows. Consider a deBruijn sequence of length n and consider a shifted version of itself aligned with it. We claim that no opposing pair of m-tuples from the two sequences can be identical since this would imply a repeated m-tuple in the deBruijn sequence. If we partition the whole sequence into m-tuples, the argument repeats. We conclude that any m-tuple or remaining fraction of an m-tuple at the end must have at least one disagreement. Since the autocorrelation is given by the number of agreements minus the number of disagreements, the positive height of a sidelobe cannot exceed

$$n - \left\lceil \frac{n}{m} \right\rceil - \left\lceil \frac{n}{m} \right\rceil = n - 2\left\lceil \frac{n}{m} \right\rceil \quad .$$

We give the following simple examples.

(i) $n = 8 = 2^3$, thus $m = 3$

$$c_M \leq 8 - 2\left\lceil \frac{8}{3} \right\rceil = 8 - 2x3 = 2$$

but $c(\tau)$ must be a multiple of four, thus we conclude

$$c_M \le 0 \quad , \quad \text{for} \quad n = 8 \quad .$$

(ii) $n = 16 = 2^4$, thus $m = 4$

$$c_M \le 16 - 2\left\lceil \frac{16}{4} \right\rceil = 16 - 8 = 8$$

i.e.,

$$c_M \le 8 \quad , \quad \text{for} \quad n = 16 \ .$$

(iii) $n = 32 = 2^5$, thus $m = 5$

$$c_M \le 32 - 2\left\lceil \frac{32}{5} \right\rceil = 32 - 2x7 = 18$$

as before, $c(\tau)$ must be a multiple of four, hence

$$c_M \le 16 \quad .$$

We have determined the autocorrelation for all deBruijn sequences up to length 32, and for each length, there has been a sequence that met this bound. This suggests that the bound may be tight. Ignoring the small end effect for large n, this suggests that the largest positive sidelobe may grow like

$$n(1 - 2/m) \quad .$$

We would aim to exclude such sequences from the ones of interest.

No useful bounds have been found for the negative sidelobes and in general, they tend to be larger than the positive ones. For example, for $n = 32$, the largest positive sidelobe is 16, but the largest negative sidelobe is -24. However, for some applications, negative sidelobes are not necessarily bad.

For reference, we list all the deBruijn sequences of length 16 in Table II.

TABLE II

LIST OF ALL SIXTEEN DeBRUIJN SEQUENCES OF LENGTH 16

A1	0 0 0 0 1 0 0 1 1 0 1 0 1 1 1 1
A2	0 0 0 0 1 1 1 1 0 1 0 1 1 0 0 1
A3	0 0 0 0 1 0 1 0 0 1 1 0 1 1 1 1
A4	0 0 0 0 1 1 1 1 0 1 1 0 0 1 0 1
B1	0 0 0 0 1 0 1 1 0 0 1 1 1 1 0 1
B2	0 0 0 0 1 0 1 1 1 1 0 0 1 1 0 1
B3	0 0 0 0 1 0 1 1 1 1 0 1 0 0 1 1
B4	0 0 0 0 1 1 0 0 1 0 1 1 1 1 0 1
C1	0 0 0 0 1 0 0 1 1 1 1 0 1 0 1 1
C2	0 0 0 0 1 1 0 1 0 1 1 1 1 0 0 1
C3	0 0 0 0 1 0 1 0 0 1 1 1 1 0 1 1
C4	0 0 0 0 1 1 0 1 1 1 1 0 0 1 0 1
D1	0 0 0 0 1 0 1 1 0 1 0 0 1 1 1 1
D2	0 0 0 0 1 1 1 1 0 0 1 0 1 1 0 1
D3	0 0 0 0 1 1 0 1 0 0 1 0 1 1 1 1
D4	0 0 0 0 1 1 1 1 0 1 0 0 1 0 1 1

In Table II, we have grouped the deBruijn sequences according to their autocorrelation functions. At this point we note that two sequences which are reciprocals (mirror images) of each other have the same autocorrelation function. Also, sequences which are complements of each other also have the same autocorrelation. Thus if there is no degeneracy, we obtain four deBruijn sequences with the same autocorrelation. In Table II, the four A sequences (A1, A2, A3, A4) have the same autocorrelation, and similarly for the B, C and D sequences. The autocorrelations are

C_A: 16,0,0,0,-4,0,0,-4,0

C_B: 16,0,0,0,-4,0,-4,0,0

C_C: 16,0,0,0,-4,4,0,-8,0

C_D: 16,0,0,0,-12,0,0,0,8

We also list a few of the more interesting autocorrelations for deBruijn sequences of length $n = 2^5 = 32$. The entire list is given in Appendix II.

(27) 32,0,0,0,0,-4,-4,0,-4,-4,-4,4,0,0,0,0,0

(150) 32,0,0,0,0,-4,-4,0,0,0,0,0,-4,4,-4,-4,0

(191) 32,0,0,0,0,-4,0,0,0,0,0,0,0,0, -8, 0,-8

(333) 32,0,0,0,0,-4,4,-4,0,0,-4,-4,0,0,0,-4,0

The number in brackets on the left is the number assigned to the sequence in the table in Appendix II. Number (191) has only negative sidelobes (and zeros) and the other three obey the rule $|c(\tau)| \leq 4$, $\tau \neq 0$.

IMPLEMENTATION OF DeBRUIJN SEQUENCES

It is well-known that there is a direct correspondence between linear PN sequences and primitive polynomials. The coefficients of the polynomial determine the taps for the linear feedback shift register. Finding linear PN sequences is thus equivalent to finding primitive polynomials of the appropriate degree. Primitive polynomials with binary coefficients have been tabulated by various authors and hence for most cases, it is merely necessary to consult a table [8].

There are no corresponding characteristic polynomials for deBruijn sequences. One hopes that there are patterns in the feedback logic which will suggest general structures or "canonical" forms. From the forms of the shift registers obtained for sequences of length 16, some promising patterns have emerged but we do not know at this time how to generalize the results for arbitrary n.

To review our method of attack, we generated all deBruijn sequences of length 8, 16 and 32. There are only two of length 8 and they correspond to extended linear PN sequences. They are reciprocals (mirror images) of each other and have the same autocorrelation. There are 16 of length 16 and 2048 of length 32 (see Table I). Those of length 16 were obtained by hand and those of length 32 by a computer, using a tree search algorithm. The tree search was conducted by employing the fundamental rule that no m-tuple is repeated in a deBruijn sequence of length 2^m. Starting with a root m-tuple, say the all-zero m-tuple, we create two branches by appending a 1 to one branch and a 0 to the other. A branch is terminated if it results in a repeated m-tuple, otherwise it branches again. In the above case, the branch with the 0

is terminated since the all-zero m-tuple is repeated. Branching continues until the tree terminates at depth n. This procedure will generate all deBruijn sequences of length $n = 2^m$, for an assumed m. For $n = 64$, there are over 67 million deBruijn sequences and the need for more analytical tools becomes very pressing.

We now describe a systematic procedure to determine the feedback function from the deBruijn sequence. To do this, recall again the result from switching algebra which says that any function of t switching variables can be expressed as a mod 2 sum of minterms in the t variables. Further, Golomb has shown that the feedback function $f_m(x_{m-1},x_{m-2},\ldots,x_1,x_0)$ for m-stage shift registers which generate deBruijn sequences can be decomposed as

$$f_m(x_{m-1},\ldots,x_1,x_0) = f_{m-1}(x_{m-1},\ldots,x_1) + x_0$$

where the x_i are the contents of the shift register and the next value x_m to be fed back into the register is given by

$$x_m = f_m(x_{m-1},\ldots,x_1,x_0)$$

As an example, consider the deBruijn sequence of length 8 below, time flowing from left to right

0 0 0 1 1 1 0 1

and the corresponding state sequence

$(x_2,x_1,x_0) \rightarrow$ 0 0 0, 1 0 0, 0 1 0, 1 0 1, 1 1 0, 1 1 1, 0 1 1, 0 0 1,...

The corresponding feedback shift register can be drawn as shown in Fig. 1. We can write

$$f_3(x_2,x_1,x_0) = f_2(x_2,x_1) + x_0$$

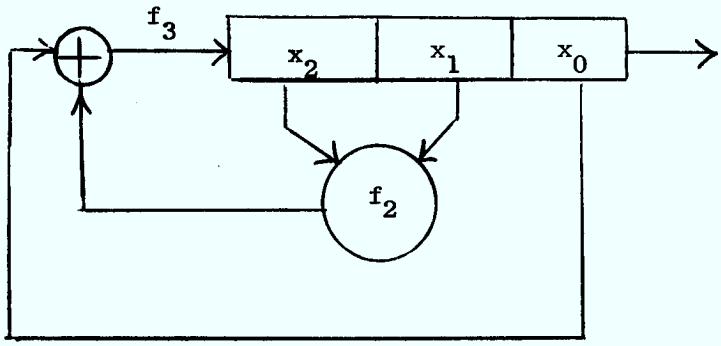


Figure 1.

where

$$f_2(x_2,x_1) = u_1x_2x_1 + u_2x_2\bar{x}_1 + u_3\bar{x}_2x_1 + u_4\bar{x}_2\bar{x}_1$$

where the u_i are unknown binary coefficients to be determined and the bar above a variable means complement. The initial state of the shift register is $(x_2,x_1,x_0) = (0,0,0)$ and the next state is $(1,0,0)$. Hence, when both x_2 and x_1 are zeros, f_2 generates a 1 (note that x_0 is 0 here). This indicates that the minterm $\bar{x}_2\bar{x}_1$ is present and $u_4 = 1$. The next state transition is from (1,0,0) to (0,1,0). By a similar argument, this means that $x_2\bar{x}_1$ is absent and $u_2 = 0$. The next state transition is from (0,1,0) to (1,0,1) and hence $\bar{x}_2,x_1$ is present and $u_3 = 1$. The next state transition provides no new information, but the transition from (1,1,0) to (1,1,1) means that $u_1 = 1$. Collecting these results we have

$$\begin{aligned} f_2(x_2,x_1) &= x_2x_1 + \bar{x}_2x_1 + \bar{x}_2\bar{x}_1 \\ &= x_1 + \bar{x}_2\bar{x}_1 \end{aligned}$$

and for f_3,

$$f_3 = \bar{x}_2\bar{x}_1 + x_1 + x_0 \quad .$$

The shift register corresponding to the above function is shown in Fig. 2. We recall that for n = 8, the two deBruijn sequences are extensions of the two PN sequences of length 7. We also note that the upper mod 2 part of the feedback function corresponds to that of the linear PN sequence. The AND gate can be viewed as inserting the extra zero at the right place and also gets the shift register out of the all-zero state. Otherwise, it is dormant, i.e., puts out zeros. The above argument generalizes for all deBruijn sequences that are extensions of linear PN sequences. This observation is not new and was known to Golomb [8].

The above relationship of the linear PN sequences and some deBruijn sequences is interesting, but for even moderate length sequences, such deBruijn sequences are an insignificant fraction of the deBruijn sequences. We thus set out to see if we could find simplified or canonical forms for all the deBruijn sequences of a given length.

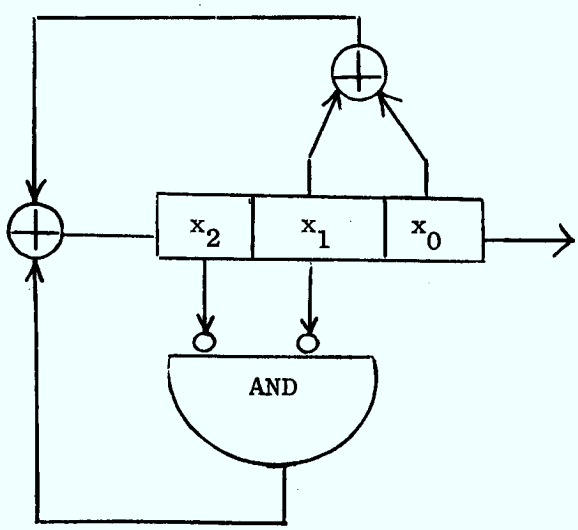


Figure 2.

We have done this for length 16, for which there are sixteen sequences, although there are only four distinct autocorrelation functions. We use the term canonical somewhat liberally here in that we manipulate the logical functions until we obtain a form that is "appealing" and seems likely to repeat for larger n. In particular, we have tried to reduce the feedback function to one AND gate in $(m-1)$ variables mod 2 added to a linear sum of some subset of the m variables or their complements.

Using the procedure we described above for the three stage shift register, we obtain the following sixteen feedback functions for the sixteen deBruijn sequences of length $n = 2^4 = 16$. As in Table II, we have grouped them in fours, each group having the same autocorrelation. We will use the same labelling of the functions in Table III as we did for the sequences in Table II.

TABLE III

LIST OF ALL SIXTEEN FEEDBACK FUNCTIONS FOR THE DeBRUIJN SEQUENCES OF LENGTH 16

$$f_{A1} = x_0 + x_1 + \bar{x}_1\bar{x}_2\bar{x}_3$$

$$f_{A2} = x_0 + x_3 + \bar{x}_1\bar{x}_2\bar{x}_3$$

$$f_{A3} = \bar{x}_0 + x_3 + x_1x_2x_3$$

$$f_{A4} = \bar{x}_0 + x_1 + x_1x_2x_3$$

$$f_{B1} = \bar{x}_0 + x_1 + x_3 + \bar{x}_1x_2x_3$$

$$f_{B2} = \bar{x}_0 + x_1 + x_3 + x_1x_2\bar{x}_3$$

$$f_{B3} = \bar{x}_0 + x_1 + x_3 + x_1\bar{x}_2\bar{x}_3$$

$$f_{B4} = \bar{x}_0 + x_1 + x_3 + \bar{x}_1\bar{x}_2x_3$$

$$f_{C1} = \bar{x}_0 + x_2 + x_3 + x_1\bar{x}_2x_3$$

$$f_{C2} = \bar{x}_0 + x_1 + x_2 + x_1\bar{x}_2x_3$$

$$f_{C3} = \bar{x}_0 + x_2 + x_3 + \bar{x}_1x_2\bar{x}_3$$

$$f_{C4} = \bar{x}_0 + x_1 + x_2 + \bar{x}_1x_2\bar{x}_3$$

$$f_{D1} = \bar{x}_0 + \bar{x}_1\bar{x}_2x_3$$

$$f_{D2} = \bar{x}_0 + x_1\bar{x}_2\bar{x}_3$$

$$f_{D3} = \bar{x}_0 + \bar{x}_1x_2x_3$$

$$f_{D4} = \bar{x}_0 + x_1x_2\bar{x}_3$$

We recall that the sequences A1 and A2 are extended linear PN sequences. It is also true that they are reciprocals of each other. Let $\tilde{A}1$ mean the reciprocal of the sequence A1, then $\tilde{A}1 \equiv A2$, where $\equiv$ means the two sequences are cyclic shifts of each other. Equivalently, we could also write $A1 \equiv \tilde{A}2$. Further, let $\bar{A}1$ mean the complement of A1. Then we observe that $\bar{A}1 \equiv A4$, $\tilde{A}3 \equiv A4$ and $\bar{A}2 \equiv A3$. There are some simple relationships between the feedback functions for sequences which are reciprocals or complements of each other. In particular, if two deBruijn sequences of length $n = 2^m$ are reciprocals of each other, then we obtain the feedback function of one from the other by the simple transformation

$$x_0 \rightarrow x_0$$
$$x_i \rightarrow x_{m-i}\ , \quad i = 1,2,\ldots,m-1 \quad .$$

When the sequences are complements of each other, we consider two cases. For case (i), let there be an odd number of variables in the linear sum, then the complementary sequence is obtained by complementing all the variables in the minterms. For example, the minterm for B1 is $\bar{x}_1 x_2 x_3$ and the minterm for $B3 \equiv \bar{B}1$ is $x_1 \bar{x}_2 \bar{x}_3$. For case (ii) the number of variables in the linear sum is even. In this case, we complement all the variables in the minterms (or more generally in $f_{m-1}(x_{m-1},\cdots,x_2,x_1)$) and then complement the entire feedback function. For example, $A3 \equiv \bar{A}2$, where

$$f_{A2} = x_0 + x_3 + \bar{x}_1 \bar{x}_2 \bar{x}_3 \quad ,$$

then according to our rule,

$$f_{A3} = \{x_0 + x_3 + x_1x_2x_3\} + 1$$

$$= \bar{x}_0 + x_3 + x_1x_2x_3$$

where we have chosen to complement x_0 to preserve the apparent symmetry.

It is hoped that the forms of the feedback functions for deBruijn sequences of length sixteen will give good clues about the functions for longer sequences. It is possible, as indicated earlier, to write down the feedback function directly from the deBruijn sequence and then manipulate the expression by various switching algebra reduction techniques. However this could be very tedious and the patterns may not emerge readily.

Obtaining simplified or canonical feedback functions for sequences of length 32 may be worth the effort, by whatever means employed. If one can manage not to become confused by the large number of sequences involved, the effort may be very rewarding in terms of formulating general theorems and algorithms. For example, some prelinimary combinatorial arguments indicate that we may need more than one AND-gate for $m = 5$.

CROSSCORRELATION OF DeBRUIJN SEQUENCES

Not much progress has been made on the crosscorrelation problem. In general it is more difficult and the number of combinations grows much faster than autocorrelation. With the explosive growth of the number of deBruijn sequences, the numbers can quickly become overwhelming. As for autocorrelation, a line of attack is to focus attention on a suitable subset of the deBruijn sequences. The job of defining such subsets remains to be done.

One result that carries over from our work on autocorrelation is the fact that

$$C_{\alpha\beta}(\tau) = \pm 4q$$

where q is zero or a positive integer. This result is proved in Appendix I for a larger class of sequences than deBruijn sequences. Also, using an argument very similar to the one used in Appendix I, it can be shown that

$$\bar{C}_{\alpha\beta} \triangleq \frac{1}{n} \sum_{\tau=0}^{n-1} C_{\alpha\beta}(\tau) = (n_1-n_0)^2/n \quad ,$$

where $C_{\alpha\beta}(\tau)$ is the crosscorrelation of any two binary sequences α and β that both have n_0 zeros and n_1 ones. Hence, for deBruijn sequences,

$$\bar{C}_{\alpha\beta} = 0$$

since $n_0 = n_1$.

This means that the average crosscorrelation for two deBruijn sequences is zero, but it does not put a bound on the value of a particular correlation value.

For example, the two deBruijn sequences of length eight are

1 1 1 0 1 0 0 0 and 1 1 1 0 0 0 1 0 and their crosscorrelation is given by

$$C_{\alpha\beta}(\tau): \quad 4,0,4,0,0,-8,0,0$$

and we see that the values are zero or multiples of four and that their sum is zero.

There are sixteen deBruijn sequences of length 16 and thus 120 crosscorrelations. Since the complement of a deBruijn sequence is also a deBruijn sequence, if we crosscorrelate such a pair we will observe that for some shift,

$$C_{\alpha\beta}(\tau) = -n$$

If such values of crosscorrelation are to be avoided, only one sequence from each complementary pair can be included in the set of interest.

For positive values of crosscorrelation, $+n$ is not allowed since this would imply the sequences were identical. The next largest value is $n-4$ and this has been achieved for sequences of length 16. In particular, for the sequences (see C1 and A1 in Table II)

$$0,0,0,0,1,0,0,1,1,1,1,0,1,0,1,1$$

and

$$0,0,0,0,1,0,0,1,1,0,1,0,1,1,1,1$$

there are 14 agreements and two disagreements and thus

$$\begin{aligned} C_{\alpha\beta}(\tau) &= 14-2 = 12 \\ &= n-4 \quad . \end{aligned}$$

The above expression is not offered as a useful bound on positive crosscorrelation sidelobes. We need an algorithm for selecting suitable

subsets of dissimilar sequences which would have low crosscorrelation sidelobes. The number of minterms in common in the feedback functions could be used as a measure of the similarity of two sequences. Some crosscorrelation values for a few sequences of length 32 are given in Appendix III.

REFERENCES

[1] N.G. deBruijn, "A Combinatorial Problem," Nederl. Akad. Wetensch., Proc., vol. 49, 1946, pp. 758-764.

[2] R.C. Dixon, ed., Spread Spectrum Techniques, IEEE Press, New York, N.Y., 1976.

[3] H. Fredricksen, "The Lexicographically Least deBruijn Cycle," J. Combinatorial Theory, vol. 9, 1970, pp. 1-5.

[4] H. Fredricksen," Generation of the Ford Sequence of Length 2^n, n Large," J. Combinatorial Theory, vol. 12(A), 1972, pp. 153-154.

[5] H. Fredricsson, "A Class of Nonlinear deBruijn Cycles," J. Combinatorial Theory, vol. 19A, 1975, pp. 192-199.

[6] S. Fredricsson, "Pseudo-Randomness Properties of Binary Shift Register Sequences," IEEE Trans. Inform. Theory, vol. IT-21, 1975, pp. 115-120.

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

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

[9] E.L. Key, "An Analysis of the Structure and Complexity of Nonlinear Binary Sequence Generators," IEEE Trans. Inform. Theory, vol. IT-22, Nov. 1976, pp. 732-736.

[10] D.E. Knuth, The Art of Computer Programming: Fundamental Algorithms, vol. 1, Addison-Wesley, Reading, Mass., 1968, p. 379.

[11] A. Lempel, "On a Homomorphism of the deBruijn Graph and its Applications to the Design of Feedback Shift Registers," IEEE Trans. Comput., vol. C-19, 1970, pp. 1204-1209.

[12] F.J. MacWilliams and N.J.A. Sloane, "Pseudo-Random Sequences and Arrays," Proc. IEEE, vol. 64, Dec. 1976, pp. 1715-1729.

[13] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1977, Chap. 14, Part II.

[14] P.S. Moharir, " Generalized PN Sequences," IEEE Trans. Inform. Theory, vol. IT-23, 1977, pp. 782-784.

[15] S. Mossige, "Constructive Theorems for the Truth Table of the Ford Sequence," J. Combinatorial Theory, vol. 11, 1971, pp. 106-110.

[16] F.J. Mowle, "Relation Between P_n Cycles and Stable Feedback Shift Registers," IEEE Trans. Electr. Comput., vol. EC-15, 1966, pp. 375-378.

[17] J. Mykkeltveit, "Generating and Counting the Double Adjacencies in a Pure Cycling Shift Register," IEEE Trans. Comput., vol. C-24, 1975, pp. 299-304.

[18] J. Mykkeltveit, "Nonlinear Recurrences and Arithmetic Codes," Inform. & Contr., vol. 33, 1977, pp. 193-209.

[19] E.J. van Lantschoot, "Double Adjacencies Between Cycles of a Circulating Shift Register," IEEE Trans. Comput., vol. C-22, 1973, pp. 944-954.

[20] K.-P. Yiu and R.B. Ward, "A Method for Deciphering a Maximal-Length Sequence," Proc. IEEE, vol. 65, July 1977, pp. 1075-1076.

APPENDIX I

(i) To prove that for a deBruijn sequence the autocorrelation function is given by

$$C(\tau) = \pm 4q$$

where q is zero or a positive integer.

We will prove this result for a larger class of sequences and in fact will prove it for crosscorrelation.

Consider the crosscorrelation of two binary n-tuples α and β where α has n_0 zeros and n_1 ones and β has n_0' zeros and n_1' ones. Let $C_{\alpha\beta}(\tau)$ be the crosscorrelation function and write

$$C_{\alpha\beta}(\tau) = A - D$$

where A is the number of agreements and D is the number of disagreements. There are four possible (α_i, β_i) pairs, where α_i and β_i are the ith element of α and β respectively. Let A_0 be the number of $(0,0)$ pairs and A_1 be the number of $(1,1)$ pairs. Similarly, let D_0 be the number of $(0,1)$ pairs and D_1 be the number of $(1,0)$ pairs. Then the following relationships are true:

$$n = n_0 + n_1 = n_0' + n_1' \qquad \text{(I-1)}$$

$$A + D = n \qquad \text{(I-2)}$$

$$A = A_0 + A_1 \qquad \text{(I-3)}$$

$$D = D_0 + D_1 \qquad \text{(I-4)}$$

$$C_{\alpha\beta}(\tau) = A - D = 2A - n \qquad \text{(I-5a)}$$

$$= n - 2D \qquad \text{(I-5b)}$$

Lemma I-1: If n is even, $C_{\alpha\beta}(\tau)$ must be even for all τ.

This follows immediately from (I-5).

It is not difficult to determine that

$$D_0 = n_0 - A_0 = n_1' - A_1 \qquad \text{(I-6)}$$

$$D_1 = n_1 - A_1 = n_0' - A_0 \quad . \qquad \text{(I-7)}$$

Hence,

$$\begin{aligned} D &= D_0 + D_1 \\ &= (n_0 - A_0) + (n_1 - A_1) \qquad \text{(I-8)} \\ &= (n_0' - A_0) + (n_1' - A_1) \quad . \qquad \text{(I-9)} \end{aligned}$$

Lemma I-2: If $n_0 = n_1'$ and n is a multiple of four, then $C_{\alpha\beta}(\tau)$ is zero or a multiple of four.

Proof.

$n_0 = n_1'$ implies $A_0 = A_1$ from (I-6). It follows that

$$A = 2A_0 = 2A_1 = \text{even} \qquad \text{(I-10)}$$

from (I-3). If n is divisible by four, we conclude from (I-5a) and (I-10) that $C_{\alpha\beta}(\tau)$ is zero or divisible by four. QED .

In words, Lemma I-2 says that if the number of zeros in one n-tuple is equal to the number of ones in the other n-tuple and n is divisible by four, then the crosscorrelation is zero or divisible by four.

Lemma I-3: If $n_0 = n_0'$ and n is a multiple of four, then $C_{\alpha\beta}(\tau)$ is zero or a multiple of four.

Proof.

$n_0 = n_0'$ implies $D_0 = D_1$ from (I-6) and (I-7) and hence

$$D = 2D_0 = 2D_1 = \text{even} . \qquad \text{(I-11)}$$

If n is divisible by four, we conclude from (I-5b) and (I-11) that $C_{\alpha\beta}(\tau)$ is zero or divisible by four. QED.

In view of either Lemma I-2 or Lemma I-3 the property we set out to prove for the autocorrelation $C(\tau)$ of deBruijn sequences is immediately true. In fact the result is also true for the crosscorrelation of deBruijn sequences since Lemmas I-2 and I-3 were proved for crosscorrelation.

(ii) To prove that

$$\bar{C} \triangleq \frac{1}{n} \sum_{\tau=0}^{n-1} C(\tau) = (n_0-n_1)^2/n$$

where $C(\tau)$ is the autocorrelation of a binary sequence of length n, having n_0 zeros and n_1 ones.

Proof.

Let us make the transformation

$$b_i = 2a_i - 1$$

where the a_i are digits from the binary sequence with elements from $\{0,1\}$ and the b_i are elements from $\{-1,+1\}$. Then we can write

$$C(\tau) = \sum_{i=0}^{n-1} b_i b_{i+\tau}$$

and

$$\overline{C} = \frac{1}{n}\sum_{\tau=0}^{n-1} C(\tau) = \frac{1}{n}\sum_{\tau=0}^{n-1}\sum_{i=0}^{n-1} b_i b_{i+\tau}$$

$$= \frac{1}{n}\sum_{i=0}^{n-1} b_i \sum_{\tau=0}^{n-1} b_{i+\tau}$$

$$= \frac{1}{n}(n_1-n_0)(n_1-n_0)$$

$$= (n_1-n_0)^2/n = (n_0-n_1)^2/n \quad . \qquad \underline{\text{QED}}.$$

(iii) To prove that

$$C(m) = 2^m - 4k$$

for a deBruijn sequence of length $n = 2^m$ where k is the number of minterms in f_{m-1}.

<u>Proof</u>.

We recall Golomb's decomposition of the feedback function $f_m(x_{m-1},\ldots,x_1,x_0)$ into

$$f_m(x_{m-1},\ldots,x_1,x_0) = f_{m-1}(x_{m-1},\ldots,x_1) + x_0 \quad .$$

This is illustrated in the figures below.

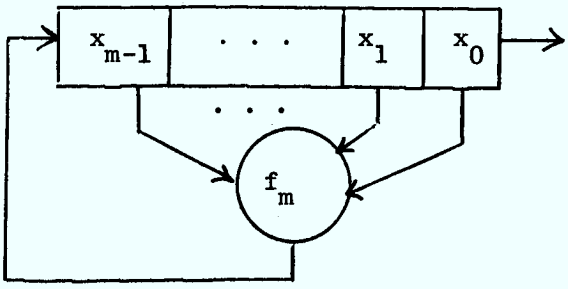


Figure A-1

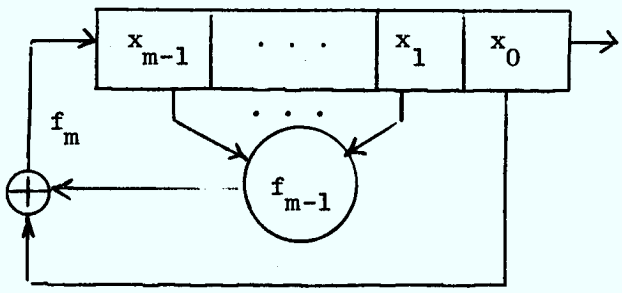


Figure A-2

Now,

$$x_m = f_m(x_{m-1},\ldots,x_1,x_0)$$
$$= f_{m-1}(x_{m-1},\ldots,x_1) + x_0 \quad .$$

Next we observe that $C(m)$ is given by the number of times x_m and x_0 are the same minus the number of times that they differ and write

$$C(m) = A(m) - D(m)$$

We can express $A(m)$ as the sum of two parts

$$A(m) = A_0(m) + A_1(m)$$

where $A_D(m)$ is the number of times $x_m = x_0 = 0$ and $A_1(m)$ is the number of times $x_m = x_0 = 1$. Also, let $D_0(m)$ be the number of times $x_m = 0$ when $x_0 = 1$ and let $D_1(m)$ be the number of times $x_m = 1$ when $x_0 = 0$. Now, f_{m-1} is the mod 2 sum of k minterms in the variables $x_{m-1},\ldots,x_1$ and $f_{m-1} = 1$ whenever one of its minterms assumes the value 1. This will happen k times with $x_0 = 0$ and k times with $x_0 = 1$ since each m-tuple appears exactly once in a complete cycle. We now claim that

$$A_0(m) = 2^{m-1}-k \quad , \quad \{x_0 = 0,\ f_{m-1} = 0,\ f_m = 0\}$$
$$A_1(m) = 2^{m-1}-k \quad , \quad \{x_0 = 1,\ f_{m-1} = 0,\ f_m = 1\}$$
$$D_0(m) = k \quad , \quad \{x_0 = 1,\ f_{m-1} = 1,\ f_m = 0\}$$
$$D_1(m) = k \quad , \quad \{x_0 = 0,\ f_{m-1} = 1,\ f_m = 1\} \quad .$$

Then from the definitions

$$C(m) = (2^m - 2k) - 2k$$
$$= 2^m - 4k = n - 4k \quad .$$

QED.

Remark. Equivalently and somewhat more directly, we could argue that $x_m \neq x_0$ whenever $f_{m-1} = 1$, since $x_m = f_{m-1} + x_0$. The two cases are $x_0 = 1$ when $f_{m-1} = 1$ and $x_0 = 0$ when $f_{m-1} = 1$. Combining these, we have

$$\begin{aligned} D(m) &= D_0(m) + D_1(m) = k + k \\ &= 2k \quad . \end{aligned}$$

Using the simple fact that

$$A + D = n$$

we have

$$A = n - D = n - 2k \quad .$$

Finally,

$$\begin{aligned} C(m) &= A - D = (n-2k) - 2k \\ &= n - 4k \\ &= 2^m - 4k \quad . \end{aligned}$$

Note on Appendix II

Copies of Appendix II can be obtained from Dr. J.L. Pearce, the project officer for the contract, or from the Principal Investigators, namely Drs. S.E. Tavares and P.H. Wittke. It is a computer printout of all the deBruijn sequences of length 32 (there are 2048 of them) and their autocorrelation functions. Sequences with the same autocorrelation are grouped together with their corresponding autocorrelation function. As discussed in the report, sequences which are mirror images (reciprocals) of each other or complements of each other have the same autocorrelation. Hence for the large majority of cases, four sequences have the same autocorrelation function. In some degenerate cases there are only two sequences with the same autocorrelation. This happens when the complement and reciprocal of a sequence are cyclic shifts of each other. In one instance, eight sequences share the same autocorrelation.

Note on Appendix III

Copies of Appendix III can be obtained from the same sources as Appendix II. It is a listing of the crosscorrelation of those sequences which were considered to have good autocorrelation functions. There were the sequences numbered (27), (150) and (333) in Appendix II. As for Appendix II, sequences with the same crosscorrelation are grouped together.

TAVARES, S.E.
Research into nonlinear pseudorandom sequence generation and

DATE DUE
DATE DE RETOUR
