Chapter 5

Spreading Waveforms

In the previous chapters we discussed different ways that standard modulation waveforms can be converted to wideband waveforms through spread spectrum techniques. In each scheme bandwidth expansion was accomplished via a *spreading waveform*. In the case of DS/SS the spreading waveform was a high rate BPSK or QPSK waveform which directly modulated the phase of the data waveform. In the case of FH/SS the spreading waveform was a series of pulsed carriers with random frequencies, thus modulating the carrier frequency of the data waveform. Up until this point we have not discussed how these waveforms are designed. Specifically, we have not discussed the properties of these spreading waveforms. It will be shown that the properties (specifically the auto-correlation and cross-correlation) of these waveforms have a tremendous impact on the performance of spread spectrum systems. Thus, in this chapter we will discuss the design of spreading waveforms, particularly the underlying spreading codes associated with the waveforms.

5.1 Role of the Spreading Waveform

The primary goal of the spreading waveform is to increase the occupied bandwidth of the data signal being transmitted. However, this spreading must be done purposefully. There are two general goals to spreading waveform design: (a) waveforms with good auto-correlation properties and (b) waveforms with good cross-correlation properties. The first goal helps make the signal as noise-like as possible and thus improves synchronization and multipath rejection. The second goal allows for multiple access using different spreading waveforms. As stated previously, military communications was the primary driver behind spread spectrum in its early days. In such situations bandwidth spreading was done to mitigate jamming and to make unauthorized detection difficult. To accomplish these purposes, it is best if the spreading waveform appears to be random. That is, we want the signal to have the properties of random noise to prevent eavesdropping and to reduce the ability of an intentional interferer. In

these cases, the spreading waveform should be pseudo-random. A key measure of the "randomness" of the spreading waveform is its autocorrelation. Ideally the waveform should be uncorrelated with shifted versions of itself. This is beneficial for LPI/LPD, aids in synchronization and provides multipath rejection capabilities as we will discuss in upcoming chapters.

A second goal of the spreading waveform design is more recent. With the advent of commercial spread spectrum systems came more emphasis on multiple access techniques that are compatible with spread spectrum waveforms. While TDMA and FDMA are certainly possible with spread spectrum waveforms, Code Division Multiple Access (CDMA) also becomes possible and is perhaps more natural. We will discuss CDMA in more detail in Chapter 12. For the current discussion, we simply need to understand that when CDMA is used, the cross-correlation properties of the set of spreading waveforms and the code set size become important issues.

5.2 Spreading Codes

We should at this point distinguish between spreading waveforms and spreading codes (or sequences). Spreading waveforms refer to the actual waveforms used to spread the bandwidth of the data signal of interest. The spreading code (or spreading sequence) is the underlying sequence of symbols used to determine the spreading waveform. It is the design of the spreading sequences that is the most important, since they will ultimately control the characteristics of the spreading waveform. We will initially discuss the various types of spreading waveforms and then spend more time on the design of good spreading sequences. We should note that in addition to "spreading code" and "spreading sequence", a third term "signature sequence" is often used to describe spreading codes, typically in the context of CDMA systems.

As mentioned, spreading waveforms and spreading sequences must be designed for both DS/SS systems as well as FH/SS systems. It should be noted that the design of the underlying spreading sequences is very different in the two cases. DS/SS sequences are typically designed on a binary alphabet whereas FH/SS sequences are designed on an N-ary alphabet where N is the number of hop frequencies. Secondly, the design criterion for multiple FH/SS codes (e.g., in multiple access environments) is the Hamming distance between sequences. This is because interference only occurs when two users hop to the same frequency. On the other hand, the design criterion for DS/SS sequences is the overall correlation. Thirdly, the mapping between FH/SS sequence and the frequencies used is arbitrary and can thus be changed regularly making it more difficult for an eavesdropper to determine the sequence. DS/SS sequences on the other hand cannot be changed in this manner making them more susceptible.

5.3 Types of Spreading Waveforms for DS/SS

There are two basic categories of spreading waveforms: pseudo-noise (or PN) waveforms and orthogonal waveforms. The first set of waveforms (sometimes called pseudo-random waveforms) are designed primarily for their auto-correlation properties. Auto-correlation properties are important for making the signal difficult to find by an interceptor while relatively easy to find by the desired receiver. PN codes which underlie PN waveforms appear to the outside observer as random sequences, although they are generated using deterministic generators. We should also note that all codes have finite length and are thus periodic. As a result, PN codes can be further divided according to their length into long codes and short codes. Long codes have a period much longer than the symbol period. That is, the repeat rate of the code is much lower than the data rate. On the other hand, short codes have a length that is on the order of the symbol period. A common technique is to use codes which have a length equal to the symbol duration. These types of systems are referred to as 'code-on-pulse' systems. Orthogonal waveforms are used for spreading when multiple access is the primary application of DS/SS and synchronism between signals can be maintained. Such waveforms are also valuable for multiplexing multiple data streams on the same signal.

The two basic properties of interest when discussing spreading codes are auto-correlation and cross-correlation. Specifically, we define the discrete auto-correlation function as

$$C_{kk}(n) = \frac{1}{N} \sum_{i=0}^{N-1} a_{k,i} a_{k,i+n} \tag{5.1}$$

where $a_{k,i} \in \{+1, -1\}$ is the ith code value of the kth user's spreading code which is related to the binary digits by $a_{k,i} = (-1)^{b_{k,i}}$ and N is the length of the spreading code. Ideally, we would like

$$C_{kk}(n) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \tag{5.2}$$

Ideal auto-correlation properties are useful for synchronization, multipath rejection capability, and low probability of detection. The discrete cross-correlation function can be similarly defined as

$$C_{km}(n) = \frac{1}{N} \sum_{i=0}^{N-1} a_{k,i} a_{m,i+n} \tag{5.3}$$

Ideally we would like to guarantee that

$$C_{km}(n) = \quad 0 \quad \forall n, k \neq m \tag{5.4}$$

Note that in synchronous systems it is possible to guarantee $C_{km}(0) = \quad 0 \quad k \neq m$, but in asynchronous systems the cross-correlation of sequences cannot be held

to zero. Ideal cross-correlation properties are useful in multiple access situations as well as for multiplexing different data streams on the same signal. Specifically, the cross-correlation between two sequences in the set S is under-bounded according to the Welch bound. That is, defining the maximum cross-correlation

$$C_{\max} = \max_{\substack{k,m \in S \\ k \neq m \text{ or } n \neq 0}} \{|C_{km}(n)|\} \tag{5.5}$$

$$C_{\max} \geq \sqrt{\frac{M-1}{MN-1}} \tag{5.6}$$

where M is the size of set S and N is the length of the codes. This bound represents a fundamental limit to cross-correlation between any two sequences in a set and is termed the Welch bound. Finally, it should also be clear that since the sequences are periodic, the discrete correlation functions are also periodic. That is

$$C_{km}(n) = C_{km}(n+N) \tag{5.7}$$

5.3.1 Pseudo-Noise Waveforms

One of the main attributes of spread spectrum systems is that they appear to be similar to white noise to unintended receivers. As such, they must have low power spectral density levels, and ideally their power spectral densities are flat. In other words they have an ideal autocorrelation function

$$R_x(\tau) = \delta(t) \tag{5.8}$$

Clearly, a realistic spreading waveform cannot have infinite bandwidth, and will thus not have the autocorrelation function described in Equation (5.8). However, consider a truly random binary waveform $x(t)$ which consists of square pulses of duration T_c and is modulated by independent binary random variables which take on the values +1 and -1 with equal probability. Further, let us assume that $x(t)$ is wide-sense stationary. Thus, $E\{x(t)\} = 0$ and $\Pr\{x(t) = i\} = \frac{1}{2}, \;\; i = +1, -1$. Defining $\Pr\{A|B\}$ as the conditional probability of event A given the occurrence of event B the autocorrelation function for this waveform is

$$\begin{aligned} R_x(\tau) &= E\{x(t)\,x(t+\tau)\} \\ &= \frac{1}{2}\Pr\{x(t+\tau)=1|x(t)=1\} - \frac{1}{2}\Pr\{x(t+\tau)=-1|x(t)=1\}... \\ &\quad +\frac{1}{2}\Pr\{x(t+\tau)=-1|x(t)=-1\}... \\ &\quad -\frac{1}{2}\Pr\{x(t+\tau)=1|x(t)=-1\} \qquad (5.9) \\ &= \frac{1}{2}(1-\Pr\{x(t+\tau)=-1|x(t)=1\})... \\ &\quad -\frac{1}{2}\Pr\{x(t+\tau)=-1|x(t)=1\}... \\ &\quad +\frac{1}{2}(1-\Pr\{x(t+\tau)=1|x(t)=-1\})... \\ &\quad -\frac{1}{2}\Pr\{x(t+\tau)=1|x(t)=-1\} \\ &= 1-2\Pr\{x(t+\tau)=1|x(t)=-1\} \qquad (5.10) \end{aligned}$$

where we have used the fact that $\Pr\{x(t+\tau)=1|x(t)=-1\} = \Pr\{x(t+\tau)=-1|x(t)=1\}$ and $\Pr\{x(t+\tau)=1|x(t)=-1\} + \Pr\{x(t+\tau)=-1|x(t)=-1\} = 1$. Now, for $\tau > T_c$

$$\begin{aligned} \Pr\{x(t+\tau)=1|x(t)=-1\} &= \Pr\{x(t+\tau)=1\} \\ &= \frac{1}{2} \qquad (5.11) \end{aligned}$$

since we are guaranteed to have at least one symbol transition and the chip values are independent. For $\tau < T_c$, if a transition occurs on $(t, t+\tau]$ then $x(t)$ and $x(t+\tau)$ are independent, otherwise $x(t) = x(t+\tau)$. Since transitions are uniformly distributed on (0,T_c], the probability of a transition occurring on $(t, t+\tau]$ is $|\tau|/T_c$. Thus, we have the autocorrelation of square pulses modulated by a random binary sequence

$$R_x(\tau) = \begin{cases} 1-\frac{|\tau|}{T_c} & |\tau| \le T_c \\ 0 & |\tau| > T_c \end{cases} \qquad (5.12)$$

In addition to making the signals appear as noise to unintended receivers, such auto-correlation properties also aid in synchronization. Practically speaking, PN codes are periodic since they cannot be infinitely long. This is obviously true due to the practical limitations of generating sequences, but also due to the inherent limitations in synchronization. Periodicities will result in spectral lines. However, provided the periodicities are sufficiently long, the spectral lines will be sufficiently close together as to eliminate discernible spikes in the spectrum. Specifically, let us represent the autocorrelation function with the triangular function $\Lambda(t)$ defined as

$$\Lambda(t) = \left\{ \begin{array}{cc} 1-|t| & |t|<1 \\ 0 & else \end{array} \right\}$$

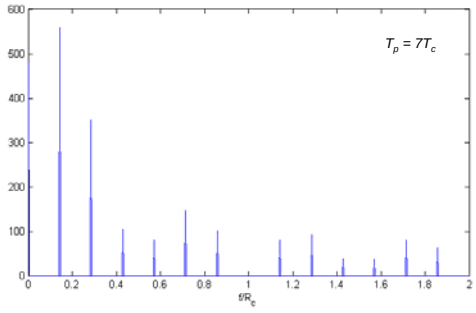

Figure 5.1: Spectrum for Short Code (N=7)

and using the Fourier transform of the triangular function, we can show that the power spectral density is[1]

$$S_x(f) = \frac{1}{N} \sum_{i=-\infty}^{\infty} \text{sinc}^2\left(\frac{i}{N}\right) \delta\left(f - \frac{i}{NT_c}\right)$$

Examples are given in Figure 5.1 through Figure 5.3. Figure 5.1 plots the spectrum of a spreading waveform with square pulses and a short (N=7) spreading code. The spectrum is purely discrete since there is no data modulation. As we increase the period of the code to N=31, we see that the spectral lines get close together (Figure 5.2) until they are indiscernible for N=310 (Figure 5.3). Long codes also have the desirable property that they randomize the interference seen by other in-band users.

Short PN codes, on the other hand, will result in discernible spectral lines due to their short period. However, these periodicities may be exploited in multiuser scenarios to build efficient multiuser detection receivers. Adaptive linear multiuser detectors are in particular based on exploiting the cyclostationarity of short code signals. This will be discussed in detail in Chapter 12. Additionally, short codes make fast synchronization more feasible and can provide better autocorrelation properties over short durations. Long codes, in contrast, can provide good auto-correlation properties over the length of the code, but not necessarily over short portions of the code. We will discuss specific types of PN codes shortly.

5.3.2 Orthogonal Waveforms

The most common type of orthogonal waveform is based on what are known as Walsh codes. Walsh codes are based on Haddamard matrices which are formed

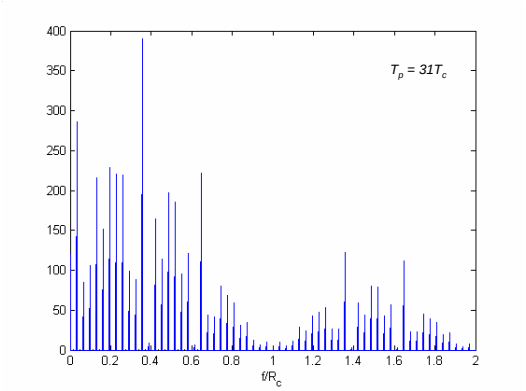

Figure 5.2: Spectrum for Short Code (N=31)

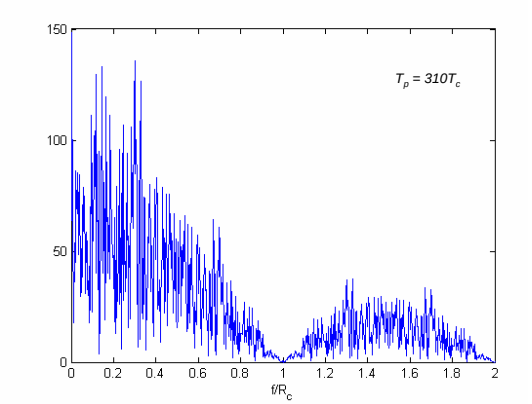

Figure 5.3: Spectrum for Long Code (N=310)

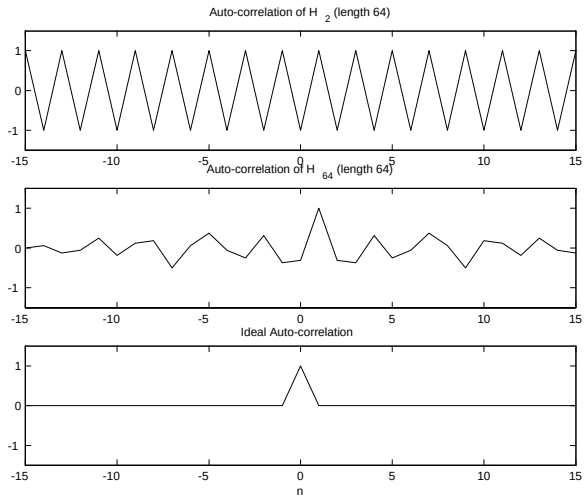

Figure 5.4: Example Autocorrelation Functions for Two Walsh Codes of Length N=64 (code 2 and code 64) Along with Ideal Autocorrelation Function

as

$$H_1 = [1] \tag{5.13}$$

$$H_{2^{i+1}} = \begin{bmatrix} H_{2^i} & H_{2^i} \\ H_{2^i} & -H_{2^i} \end{bmatrix} \tag{5.14}$$

Walsh codes of length 2^i are then formed from the rows of the Haddamard matrix H_{2^i}. Note that the rows of the Haddamard matrix are orthogonal, and thus can be used to form 2^i orthogonal spreading codes. The length of the codes are restricted to be a power of 2. Additionally, orthogonality is obtained only when the codes are aligned properly in time (i.e., synchronous). The cross correlation properties of the codes are poor for non-synchronous alignment. Additionally, the auto-correlation properties are poor as demonstrated in Figure 5.4. Plotted in the figure are example auto-correlation functions for length N=64 Walsh codes. Specifically, the auto-correlation functions for the 2nd and 64th Walsh codes are plotted along with the ideal auto-correlation function. We can see that the auto-correlation properties are not good for Walsh codes, especially code 2. In general, Walsh codes must be augmented with other codes to mitigate this short-coming for synchronization purposes.

The other limitation of the above mentioned Walsh codes is that they require all signals in the system to use the same spreading (i.e., bandwidth expansion) factor. In multimedia scenarios this is not desirable as we may want multiple possible data rates while keeping the same overall chip rate. Thus, we are also interested in codes which are orthogonal for variable spreading factors. Such a scenario can be accommodated through the use of Orthogonal Variable

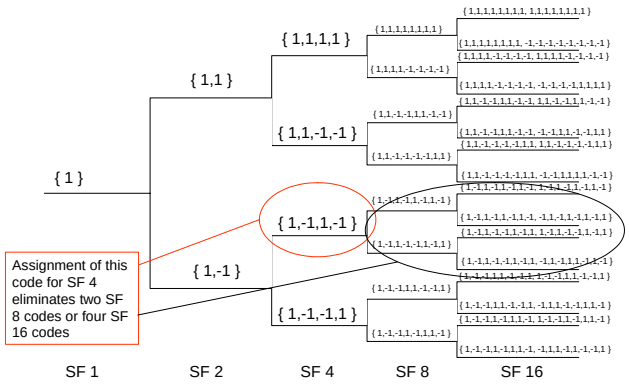

Figure 5.5: Orthogonal Variable Spreading Factor Codes

Spreading Factor (OVSF) codes [2]. The codes are also based on Haddamard matrices, but are assigned using a binary tree as shown in Figure 5.5. If the longest length is $N = 2^k$, codes can be assigned from length 1 to N. However, each code of length 2^{k-i} eliminates 2^i codes of length 2^k from use. For example, consider Figure 5.5 where N=16 and thus there are 16 total codes that can be assigned. Suppose that a code of length $2^{4-2} = 4$ is assigned. We can see from the code tree that 2 codes of length 8 are eliminated or 4 codes of length 16 are eliminated. In this way users with different spreading gains can remain orthogonal, provided that their signals are time synchronous.

5.4 Pseudo-random sequences

In this section we study an important aspect of spreading waveforms, the generation of pseudorandom sequences. Before we do so, it is perhaps useful to recall a statement made by John Von Neumann (quoted in Knuth [3]): "Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin." The key point to consider here is that we cannot truly create random sequences but rather deterministic sequences which appear to an outside observer to be random. Thus, we require sequences which have "noise-like" properties, such as auto-correlation.

5.4.1 Galois Fields

In order to study the design of pseudo-random sequences, we must first briefly review finite field (also called a Galois field) arithmetic. Consider the set of M

Table 5.1: Addition

+	0	1
0	0	1
1	1	0

Table 5.2: Multiplication

x	0	1
0	0	0
1	0	1

elements $S = \{e_0, e_1, \ldots, e_{M-1}\}$. In order to define a finite field, we must define two operations, addition and multiplication, such that

- The set must be closed under both operations
- Both operations must be commutative (i.e., $a + b = b + a$ and $ab = ba$)
- The set must have additive and multiplicative identity elements
- The set must have additive and multiplicative inverse elements
- Multiplication must be distributive over addition
- Addition and multiplication must be associative (i.e., $a + (b + c) = (a + b) + c$)

The set of integers $\{0, 1, 2, \ldots, M-1\}$ where M is prime and multiplication and addition are carried out modulo M is a finite field. We will primarily concern ourselves with the binary set $\{0, 1\}$. Note that the additive inverse element for 0 is 0 and the additive inverse element for 1 is 1. That is

$$\begin{aligned} 0+0 &= 0 \\ 1+1 &= 0 \end{aligned} \tag{5.15}$$

Further, the tables for addition and multiplication are given in Tables 5.1 and 5.2 respectively, verifying the other properties.

We must also define polynomial addition and multiplication. Consider a polynomial $f(D) = f_0 + f_1 D + f_2 D^2 + \ldots f_m D^m$ where $f_i \in \{0, 1, ...M-1\}$and addition is defined as

$$\begin{aligned} h(D) &= f(D) + g(D) \\ &= (f_0 + g_0) + (f_1 + g_1)D + (f_2 + g_2)D^2 + \ldots (f_m + g_m)D^m \end{aligned} \tag{5.16}$$

and for multiplication, if $h(D) = f(D)g(D)$ then

$$\begin{aligned} h_0 &= f_0 g_0 \\ h_1 &= f_0 g_1 + f_1 g_0 \\ h_2 &= f_0 g_2 + f_1 g_1 + f_2 g_0 \\ &\vdots \\ h_{2m} &= f_m g_m \end{aligned} \tag{5.17}$$

We can also factor polynomials. For example, the polynomial $f(D) = D^4 + D^3 + D + 1$ can be factored into $f(D) = (D+1)(D^3+1)$. Now, consider a sequence of binary digits $\{\ldots, b_{-3}, b_{-2}, b_{-1}, b_0, b_1, b_2, b_3, \ldots\}$ taken from $S = \{0, 1\}$. The sequence can be represented by the polynomial

$$\ldots, D^{-3}b_{-3} + D^{-2}b_{-2} + D^{-1}b_{-1} + D^0 b_0 + D^1 b_1 + D^2 b_2 + D^3 b_3, \ldots \tag{5.18}$$

which allows us to represent a sequence of binary integers using polynomial notation. Now, a spreading waveform for DS/SS can be defined as

$$a(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT_c) \tag{5.19}$$

where T_c is a chip pulse period and $a_i = (-1)^{b_i}$. The auto-correlation function over the period of the waveform $T_w = NT_c$ is

$$R_a(\tau) = \frac{1}{T_w} \int_0^{T_w} a(t)a(t+\tau)dt \tag{5.20}$$

The cross-correlation between two waveforms $a(t)$ and $a'(t)$ is

$$R_{aa'}(\tau) = \frac{1}{T_w} \int_0^{T_w} a(t)a'(t+\tau)dt \tag{5.21}$$

Note that $R_a(\tau)$ and $R_{aa'}(\tau)$ are periodic in T_w. We can get a better understanding of the correlation function by substituting for $a(t)$:

$$R_{aa'}(\tau) = \frac{1}{T_w} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} a_n a_m \int_0^{T_w} p(t - nT_c)p(t + \tau - mT_c)dt \tag{5.22}$$

Now, if we define $\tau = kT_c + \tau_e$ we have

$$\begin{aligned} R_{aa'}(k, \tau_e) &= \frac{1}{N} \sum_{m=1}^{N-1} a_m a'_{k+m} \frac{1}{T_c} \int_0^{T_c - \tau_e} p(\lambda)p(\lambda + \tau_e)dt... \\ &\quad + \frac{1}{N} \sum_{m=1}^{N-1} a_m a'_{k+m} \frac{1}{T_c} \int_{T_c - \tau_e}^{T_c} p(\lambda)p(\lambda - T_c + \tau_e)dt \end{aligned} \tag{5.23}$$

For square pulses the integrals are

$$\frac{1}{T_c}\int_0^{T_c-\tau_e} p(\lambda)p(\lambda+\tau_e)dt = 1-\frac{\tau_e}{T_c} \tag{5.24}$$

$$\frac{1}{T_c}\int_{T_c-\tau_e}^{T_c} p(\lambda)p(\lambda-T_c+\tau_e)dt = \frac{\tau_e}{T_c} \tag{5.25}$$

Using the definitions for the discrete auto-correlation and cross correlation in equations (5.1) and (5.3) respectively, the cross-correlation function becomes

$$R_{aa'}(k,\tau_e) = \left(1-\frac{\tau_e}{T_c}\right)C_{aa'}(k) + \left(\frac{\tau_e}{T_c}\right)C_{aa'}(k+1) \tag{5.26}$$

We can see from this formulation that the auto-correlation and cross-correlation functions are dominated by the discrete auto-correlation and cross-correlation functions of the sequences. Thus, we wish to find sequences that are (a) easy to generate, (b) have good auto-correlation properties, (c) have good cross-correlation properties, and (d) have a large number of codes within the set. One very efficient means of generating sequences is the use of the Linear Feedback Shift Register (LFSR). An k-element LFSR can generate a sequence of length 2^k-1. There are several sequences of importance that can be generated from LFSR's including m-sequences, Gold Codes, and Kasami sequences. Two common formats (the modular format and the simple format) for representing the LFSR are shown in Figure 5.6.

Both registers in Figure 5.6 are defined by the taps $\mathbf{g}$ and are functionally equivalent. The values of $\mathbf{g}$ are binary from set $S=\{0,1\}$. As a result, the "multiply" functions can be implemented simply by the absence or presence of taps. The modular format is convenient since it is easy to build a delay/adder component. Also, the required clock speed is slower than in the simple form since in the simple form there are potentially $N-1$ adds that must be done serially each shift. Using the polynomial notation introduced earlier, we can determine the output sequence $b(D)$ from the generator taps $g(D)$ and the initial load $a(D)$:

$$b(D) = a(D)/g(D) \tag{5.27}$$

As an example consider a generator with taps defined by $g(D)=1+D+D^3+D^4$ and an initial load of $a(D)=1$.

$$\begin{aligned} b(D) &= \frac{1}{1+D+D^3+D^4} \\ &= 1+D+D^2+D^6+D^7+D^8+\ldots \end{aligned} \tag{5.28}$$

Thus, the output bit stream is 111000111000....

5.4.2 m-sequences

A k-length LFSR has a maximum sequence length of 2^k-1. This comes from the fact that the register can be in 2^k different possible states. However, if the

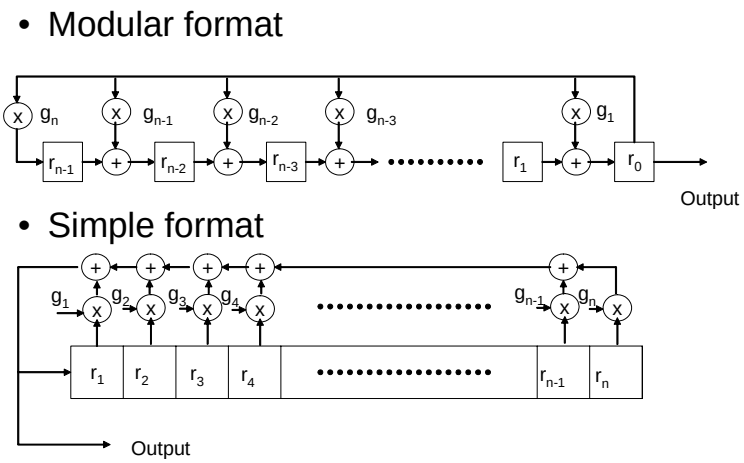

Figure 5.6: Linear Feedback Shift Register

LFSR is in the all-zeros state it will never leave that state. This can be seen by simply examining the LFSR diagram in Figure 5.6. One family of sequences which has length $2^k - 1$ is the maximal length sequence or the m-sequence which is named for the fact that its sequences are of maximal length. m-sequences have other good properties as we will now discuss.

Property 1 *The number of ones in the code minus 1 equals the number of zeros in the code. That is each code is of length* $2^k - 1$ *and has* 2^{k-1} *1's and* $\left(2^{k-1} - 1\right)$ *0's.*

Proof Consider the generator in Figure 5.6. The last value in the modular format a_0 corresponds to the output sequence of chips. If the state cycled through all 2^k values, a_0 would take on the value '1' 2^{k-1} times and the value '0' 2^{k-1} times. However, since the all zeros state cannot occur, we know that '0' will appear one fewer time, i.e., $2^{k-1} - 1$ times.

Property 2 *If a window of length k is slid across the sequence, every k-tuple of values will appear exactly once except the all zeros k-tuple.*

Proof Again, we know that for the length of the 2^{k-1} sequence, we will cycle through all states except the all zeros state. The k-tuple simply corresponds to the states of the sequence generator. Since the states appear only once, the k-tuple will appear only once.

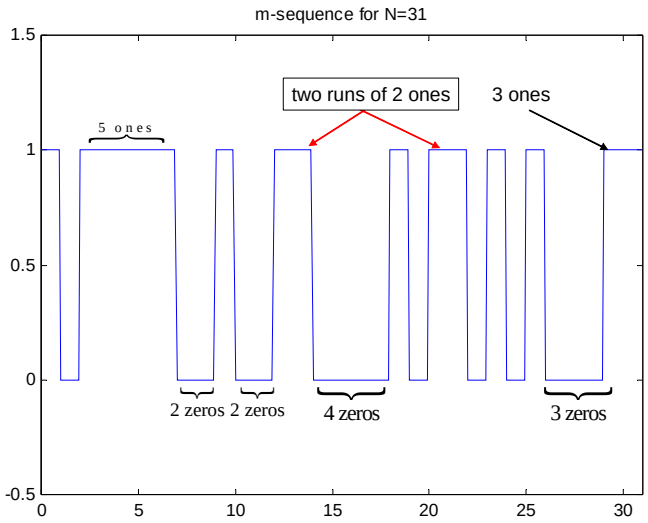

Figure 5.7: Example m-sequence

Property 3 *The modulo-2 sum of an m-sequence with a phase shifted version of itself will result in another shifted version of itself.*

Proof We know that the output of a linear feedback shift register is $b(D) = a(D)/g(D)$ and $b'(D) = a'(D)/g(D)$ where $b(D)$ and $b'(D)$ are distinct phases of the same m-sequence and $a(D)$ and $a'(D)$ are different initial conditions and $g(D)$ are the taps. Now,

$$b(D) + b(D)' = [a(D) + a'(D)]/g(D) \tag{5.29}$$

The modulo-2 sum of any two states will give a third state $a''(D)$. Since the two states are distinct we know that $a(D) + a'(D) \neq 0$ and $a(D) + a'(D) \neq a(D)$. Thus, $a''(D)/g(D)$ is simply a different phase $b''(D)$ of the same m-sequence.

Property 4 *The periodic auto-correlation function is two-valued taking on the values*

$$C_{k,k}(b) = \begin{cases} 1 & b = lN \\ -\frac{1}{N} & b \neq lN \end{cases} \tag{5.30}$$

Proof The value of the periodic auto-correlation function is $(N_a - N_d)/N$ where N_a is the number of agreements and N_d is the number of disagreements between the sequence and a shifted version of itself. This is also equal to the number of zeros (agreements) and the number of ones (disagreements) in the modulo-2 sum of sequence $\mathbf{b}$ and a shifted version of $\mathbf{b}$. Clearly for any value

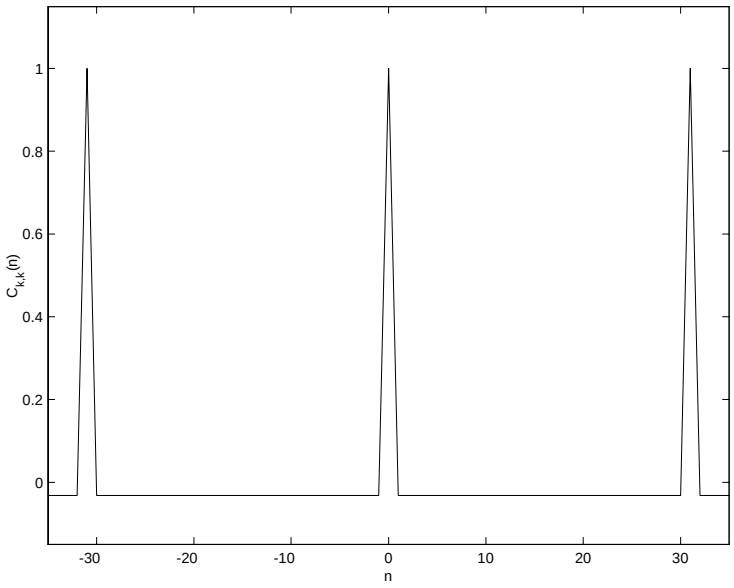

Figure 5.8: Auto-correlation of an Example m-sequence for $N = 31$

lN the modulo-2 sum will be all zeroes ($N_a = N$ and $N_d = 0$) and thus the auto-correlation is 1. For $n \neq lN$ we know that the modulo-2 sum is another phase of the same m-sequence. From the first property we know that it will contain one more '1' than '0'. Thus, the auto-correlation function will be $-\frac{1}{N}$. An example is shown in Figure 5.8.

Property 5 *Define a run as a subsequence of identical symbols within the m-sequence. The length of the subsequence is the length of the run. Then for any m-sequence there is*

- *1 run of ones of length k*
- *1 run of zeros of length* $k-1$
- *1 run of ones and 1 run of zeros of length* $k-2$

  $\vdots$

- 2^{k-3} *runs of ones and runs of zeros of length 1*

Proof Again consider the shift register in Figure 5.6. Clearly there can be no run of k zeros since that would require the all zeros state to appear, which it cannot. Thus, there cannot be a string of k zeros. Similarly, we know that the all ones state occurs only once. This means that a run of ones of length k will

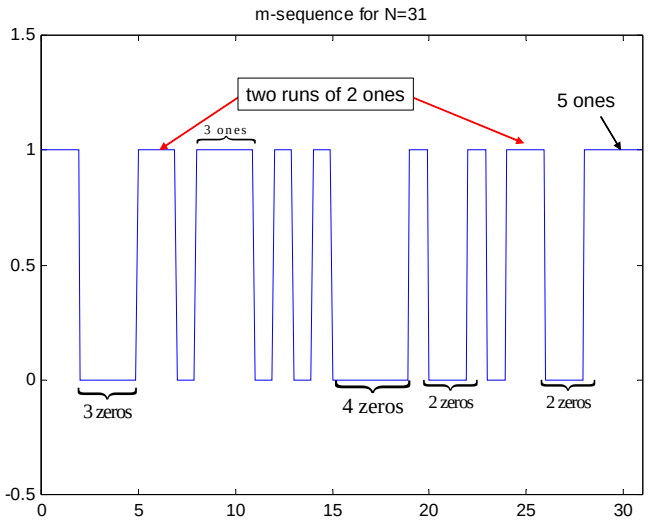

Figure 5.9: Example m-sequence

occur but not more than once. Further, any run of ones longer than k would require consecutive appearances of the all ones state which cannot happen. We also know that a run of $k-1$ ones cannot occur. This is because a run of $k-1$ ones requires a zero on either side of it and that requires the states $0111\ldots1$ and $111\ldots110$ to occur successively. However, we know that both of these states must occur before and after the run of k ones. Since the states cannot occur more than once, a string of $N-1$ ones must not occur. Consider a run of l ones with $0 \leq l \leq k-1$. The shift register must pass through the state containing $0\underbrace{11\ldots1}_{l}0$. Since $l+2$ bits are defined, there are $k-l-2$ remaining values in the shift register that can take on any values. Thus, there are 2^{k-l-2} ways this run can occur. Similarly for l zeros.

As an example let's consider a size $k=5$ shift register with taps $\mathbf{g}=[00101]$, which results in a length $N=2^k-1=31$ sequence and an initial load of 11011. The resulting sequence is plotted in Figure 5.9. We can see the runs of ones and zeros that were predicted. The discrete cross-correlation function between two example sequences is plotted in Figure 5.10

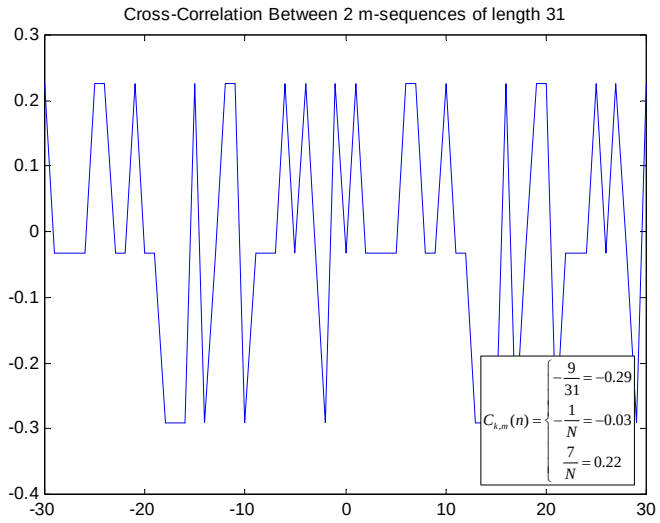

Figure 5.10: Cross-correlation of Two Example m-sequences

n	$N = 2^n - 1$	$N_p(n)$
2	3	1
3	7	2
4	15	2
5	31	6
6	63	6
7	127	18
8	255	16
9	511	48
10	1023	60

Table 5.3: Code Length N and the number of m-sequences $N_p(n)$

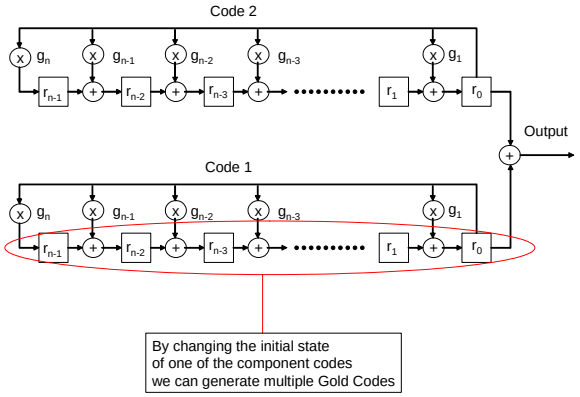

Figure 5.11: Gold Code Generation from m-sequences

5.4.3 Gold Codes

While m-sequences have excellent auto-correlation properties, they are not great candidates for CDMA systems. This is because there are only a small number of m-sequences for any given sequence length $N = 2^k - 1$ as can be seen in Table 5.3 [2]. However, it was shown by Gold in 1967 that certain pairs of m-sequences (called preferred pairs) have well behaved cross-correlation properties (they are 3 valued) and can be combined to form sequences called Gold codes [4]. Specifically, $N+2$ Gold codes can be created from a length N preferred pair of m-sequences as shown in Figure 5.11. These sequences come from modulo-2 summing an m-sequence with N phases of the other half of a preferred pair as well as the original two sequences. The cross-correlation between any two of these Gold codes is three-valued:

$$C_{k,m}(b) \in \left\{ -\frac{1}{N}\left(1 + 2^{0.5(n+2)}\right), -\frac{1}{N}, \frac{1}{N}\left(2^{0.5(n+2)} - 1\right) \right\} \tag{5.31}$$

for n even and

$$C_{k,m}(b) \in \left\{ -\frac{1}{N}\left(1 + 2^{0.5(n+1)}\right), -\frac{1}{N}, \frac{1}{N}\left(2^{0.5(n+1)} - 1\right) \right\} \tag{5.32}$$

for n odd.

The benefit of Gold codes is that there are a large number of them for a given length N while having controlled cross-correlation properties as shown in Figure 5.13. However, the downside is that the auto-correlation properties of Gold codes are not as good as with m-sequences. This can be seen in the example in Figure 5.12. Compared with Figure 5.8, we can see that the auto-correlation value for $n \neq 0$ are substantially larger. Specifically, the out-of-phase portion of the auto-correlation is now three-valued taking on the same three values as the cross-correlation function given in equations (5.31) and (5.32).

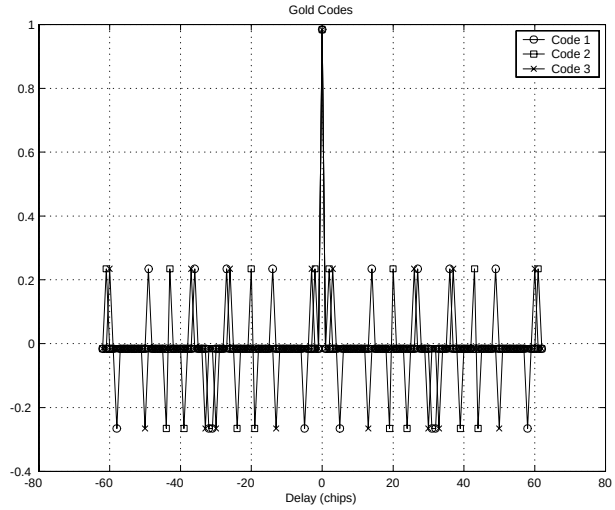

Figure 5.12: Example Autocorrelation Functions for length 63 Gold Sequences

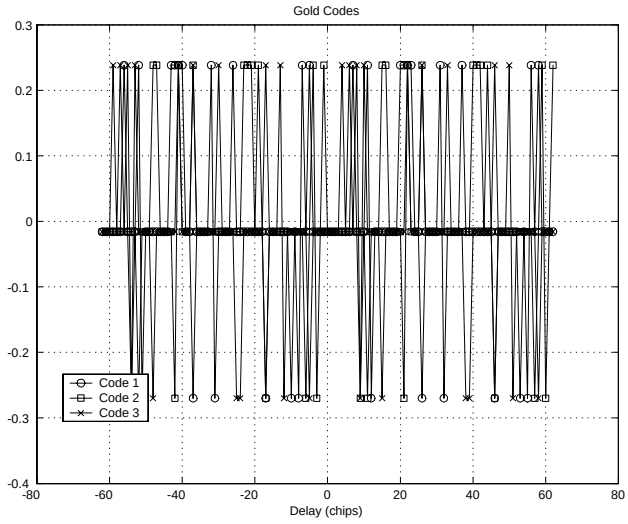

Figure 5.13: Example Cross-correlation Functions for Gold Codes of Length 63

5.5 Kasami Sequences

As discussed previously, Welch showed that the maximum cross-correlation between any two sequences in a set of M sequences is lower bounded [5]. Specifically, he showed that the maximum cross correlation between two sequences was lower bounded by equation (5.6). Thus, for relatively large sets the maximum cross-correlation is greater than $\sqrt{\frac{1}{N}}$. From our discussion on Gold codes we know that the maximum cross-correlation is

$$\max_{n,k,m} C_{k,m}(n)^{Gold} \approx \sqrt{\frac{2}{N}} \tag{5.33}$$

for n odd, and

$$\max_{n,k,m} C_{k,m}(n)^{Gold} \approx \sqrt{\frac{4}{N}} \tag{5.34}$$

for n even. Thus, the maximum cross-correlation of Gold codes is higher by at least a factor of $\sqrt{2}$ than optimal codes. Another set of codes called Kasami sequences [2][6][7] can be constructed from m-sequences. These sequences have a cardinality of $2^{n/2}$ for a length of $N = 2^n - 1$ and have a three-valued cross-correlation function. Specifically, the cross correlation function takes on values from the set $\left\{-\frac{1}{N}, -\frac{1}{N}\left(2^{n/2}+1\right), \frac{1}{N}\left(2^{n/2}-1\right)\right\}$ which satisfies the Welch lower bound[1]. Kasami sequences are formed in a manner similar to Gold codes. We start with an m-sequence **a** of length $N = 2^n - 1$ where n is even. By decimating the sequence by $2^{n/2}+1$, we obtain a second m-sequence **a'** of length $2^{n/2}-1$. By adding (modulo two) **a** and $2^{n/2}-1$ shifted versions of **a'** we obtain a set of $2^{n/2}-1$ sequences. By also inlcuding the original sequence **a** we can ultimately obtain M=$2^{n/2}$ total sequences. Unfortunately, while this set satisfies the Welch bound, this is not a very large set and is often called the *small set* of Kasami sequences. A larger set of Kasami sequences can be obtained which inlcudes Gold sequences and the small set of Kasami sequences provided that $\mathrm{mod}\,(n,4) = 2$. Again let **a** be an m-sequence of length $N = 2^n - 1$. Now, let sequences **a'** and **a"** be formed by decimating the original sequnece **a** by $2^{n/2}+1$ and $2^{(n+2)/2}+1$. The first is a length $2^{n/2}-1$ m-sequence, but the second is another length $2^n - 1$ m-sequence. We can form the small set of Kasami sequences by modulo-2 summing **a** with shifted versions of **a'**. If we further take all sequences by summing **a**, and a" we obtain a set of $2^n - 1$ Gold codes. We can also obtain $2^{n/2}-1$ codes by modulo-2 summing **a'** and **a"**. Finally, we can obtain $\left(2^{n/2}-1\right)\left(2^n-1\right)$ by summing all phases of **a**, **a'** and **a"**. Including a and a" we thus have $M = 2^{3n/2} + 2^{n/2}$ total codes. All auto-correlation and cross-correlation values from members of this set are limited to the set $\left\{-1, -1 \pm 2^{n/2}, -1 \pm 2^{n/2+1}\right\}$.

[1] This can be readily seen by recalling that $N = 2^n - 1$.

5.6 Layered Spreading Waveforms

As we have seen, the two classes of spreading waveforms have very different attributes. Pseudo-random codes are useful for their auto-correlation properties while orthogonoal waveforms are useful for their synchronous cross-correlation properties. In cellular CDMA systems, both properties are useful. Specifically, in cellular CDMA downlinks, the system supports several serparate users on a single waveform, necessitating orthogonal channelization codes. Further, because the system supports many cells in a single band, we desire codes that appear to be noise to other base stations.

This situation motivates the use of *layered* spreading codes. Within a specific cell (i.e., signals coming from the same base station, are separated using orthogonal codes, while signals from different base stations are randomized using different pseudo-random sequences. As an example consider the downlink of the IS-95 cellular standard outlined in Figure 5.14. The different user signals originating from the same base station are separated through the use of orthogonal channelization codes using 64-ary Walsh-Haddamard codes. The coded data stream has a bit rate of 19.2kbps and spreading via length 64 Walsh-Haddamard codes results in 1.2288Mcps transmission. After all user signals are summed, they are multiplied by a common pseudo-random code to improve the autocorrelation properties. The pseudo-random code, often called a covering code, is unique to a particular base station and thus allows mobile stations to distinguish between different base stations[2]. Similar uses of layered spreading codes are also used in the UMTS cellular 3G standard.

[2] To be precise, the IS-95 standard does not call for base stations to use distinct covering codes. Instead, base stations use different phases of a common PN code.

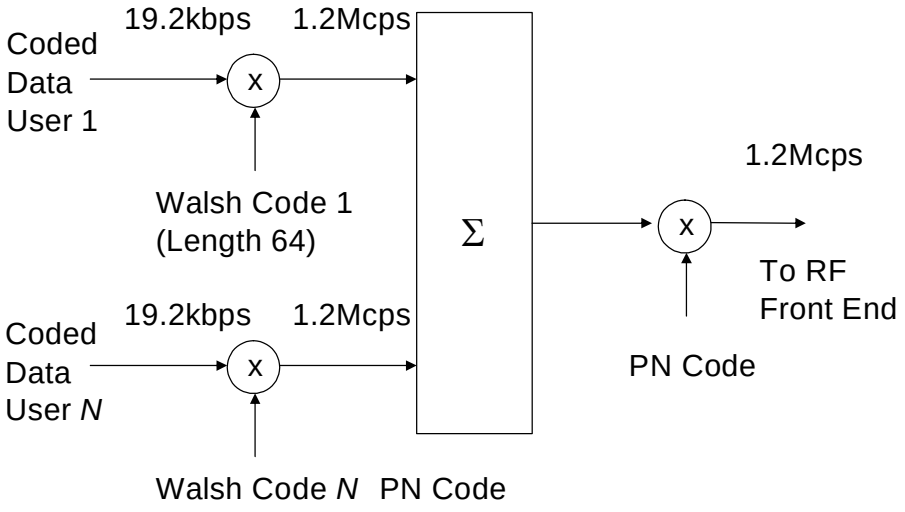

Figure 5.14: Example of Layered Spreading Codes from IS-95 Cellular Standard (downlink)

Bibliography

[1] D. Torrieri, *Principles of Spread Spectrum Systems*. New York, NY: Springer, 2005.

[2] E. Dinan and B. Jabbari, "Spreading codes for direct sequence cdma and wideband cdma cellular networks," *IEEE Communications Magazine*, vol. 36, pp. 48–54, September 1998.

[3] D. Knuth, *The Art of Computer Programming*, vol. 2. New York, NY: Addison-Wesley, third ed., 1998.

[4] R. Gold, "Optimum binary sequences for spread spectrum multiplexing," *IEEE Transactions on Information Theory*, vol. IT-13, pp. 619–621, October 1967.

[5] L. Welch, "Lower bounds on the maximum cross-correlation of signals," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 397–399, May 1974.

[6] T. Kasami, "Weight distribution formula for some class of cyclic codes," *University of Illinois-Urbana Technical Report R-285*, April 1966.

[7] M. Simon, J. Omura, R. Scholtz, and B.K.Levitt, *Spread Spectrum Communications Handbook*. New York, NY: McGraw-Hill, electronic ed., 2002.