Save of Old Section 2.5 of scrambler PhL 7.21.13

Review this for anything important omitted from its replacement: **DONE**
There is no need to ever look at this doc again.

2.5 Autocorrelation, Probability, and Statistics for a sequence in {0,1} and {-1,1}

The Autocorrelation Sequence for a Periodic Sequence

The autocorrelation *function* for a continuous signal a(t) was defined in FT (32.1), where we omitted any kind of normalizing factor,

$$r_x(t) \equiv \int_{-\infty}^{\infty} dt'\, x(t')\, x(t'+t) \qquad // = r_x(-t) \qquad \textit{FT (32.1)}$$

Although we did not use this terminology in FT, it should be clear that one can define an exactly corresponding "digital" autocorrelation *sequence* having the following form:

$$r_k \equiv \sum_{n=-\infty}^{\infty} a_n\, a_{n+k} \qquad // = r_{-k} \qquad (2.5.1)$$

At this point, we still imagine that $a_n = a(t_n)$ is a sample of an analog signal. The sum converges if the sequence tapers off sufficiently in the past and future.

For a sequence of numbers which are elements of some finite set, the only way to make something "taper off" is to have it be zero before some time in the past and after some time in the future. For a periodic sequence (such as an MLS sequence) the above definition of r_k cannot possible converge, so we replace it with the following definition where the sum is over one period of the periodic sequence:

$$r_k \equiv (1/P) \sum_{n=1}^{P} a_n\, a_{n+k} \quad . \qquad (2.5.2)$$

P is the period of the periodic sequence and for an MLS sequence $P = 2^k - 1$. Here it is implied that if a_{n+k} takes us off the end of the sequence, we wrap to the beginning of the sequence, just as if the sequence repeated forever in time.

The sum shown above is obviously the average value of the product $a_n a_{n+k}$ for a periodic sequence which we indicate by $\langle a_n a_{n+k} \rangle$. So we rewrite the above adding this new notation

$$\langle a_n a_m \rangle = r_{m-n} \equiv (1/P) \sum_{n=1}^{P} a_n\, a_m \qquad m \neq n \qquad (2.5.3)$$

$$\langle a_n^2 \rangle \;=\; r_0 \;\equiv\; (1/P) \sum_{n=1}^{P} a_n^2 \;. \qquad\qquad m = n$$

Quantities like $\langle a_n \rangle$ or $\langle a_n^2 \rangle$ don't depend on n since these are averages of some quantity over the sequence, so we can write $\langle a_n^2 \rangle = \beta$. A quantity like $\langle a_n a_{n+k} \rangle$, however, can possibly depend on the distance k between the two symbol locations n and n+k. This is the case, for example, for the AMI code treated in FT ***. But if there is no correlation between the values of symbols at different locations in the sequence, then $\langle a_n a_{n+k} \rangle$ is independent of k (as well as n), so we can write it as $\langle a_n a_{n+k} \rangle = \alpha$. Thus we can write

$$\alpha = \langle a_m a_n \rangle = r_{m-n} \quad \text{for } m \neq n \qquad\qquad \text{// uncorrelated}$$

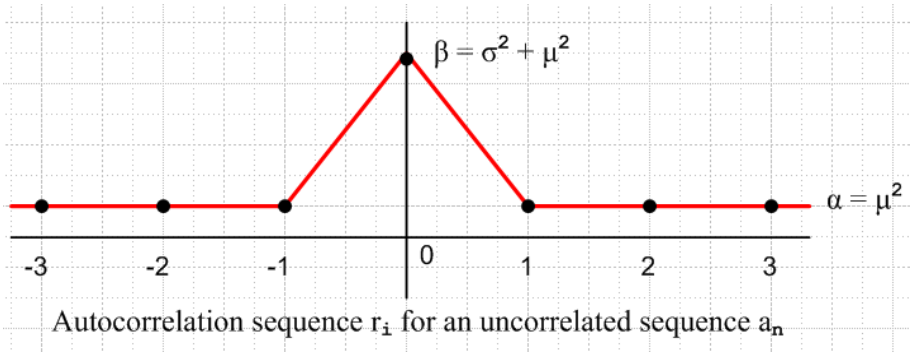
$$\beta = \langle a_n^2 \rangle = r_0 \;.$$

As we show below, in the uncorrelated case, $\langle a_m a_n \rangle = \langle a_m \rangle \langle a_n \rangle = \langle a_m \rangle^2 = \mu^2$ where μ is the mean value of the symbols a_i in the sequence. The variance σ^2 of a sequence is $\sigma^2 = \langle a_n^2 \rangle - \langle a_n \rangle^2 = \beta - \mu^2$, so we can rewrite the above equations to include the mean μ and variance σ^2

$$\alpha = \langle a_m a_n \rangle = \mu^2 = r_{m-n} \qquad \text{for } m \neq n \qquad \text{// uncorrelated}$$

$$\beta = \langle a_n^2 \rangle = (\sigma^2 + \mu^2) = r_0 \qquad \Rightarrow \qquad \sigma^2 = \beta - \mu^2 = (\beta - \alpha) \qquad (2.5.4)$$

For this uncorrelated sequence, we can at once plot the autocorrelation sequence:



Autocorrelation sequence r_i for an uncorrelated sequence a_n

Here the sequence values r_i are at the dots and the red line just interpolates these values.

In FT we compute the spectral power density $\langle \mathcal{P}(\omega) \rangle$ for an *uncorrelated* amplitude-modulated pulse train of the form $x(t) = \sum_{n=-\infty}^{\infty} a_n \, x_{pulse}(t - nT_1)$. The result is as follows

$$\langle \mathcal{P}(\omega) \rangle = \mathcal{P}_{pulse}(\omega) \left[(\beta - \alpha) + \alpha \sum_{m=-\infty}^{\infty} 2\pi \, \delta(\omega T_1 - 2\pi m) \right] \qquad \text{FT (35.11)}$$

where α and β are the quantities shown in (2.5.4). Here $\mathscr{P}_{\mathbf{pulse}}(\omega)$ is the spectral power density of the underlying pulse $x_{\mathbf{pulse}}(t)$ used to make the pulse train. If the pulse train sample spacing is time T_1 and if $x_{\mathbf{pulse}}(t)$ is square pulse of amplitude 1 and width T_1, then

$$X_{\mathbf{pulse}}(\omega) = T_1 \, \mathrm{sinc}(\omega T_1/2)$$

$$\mathscr{P}_{\mathbf{pulse}}(\omega) = \frac{|X_{\mathbf{pulse}}(\omega)|^2}{2\pi T_1} = (T_1/2\pi) \, \mathrm{sinc}^2(\omega T_1/2) \qquad \text{FT (36.1)}$$

and then we get, for an uncorrelated sequence $\{a_n\}$,

$$\langle \mathscr{P}(\omega) \rangle = (T_1/2\pi) \, \mathrm{sinc}^2(\omega T_1/2) \, [\, (\beta-\alpha) \; + \; \alpha \sum_{m=-\infty}^{\infty} 2\pi \, \delta(\omega T_1 - 2\pi m) \,]$$

$$= (T_1/2\pi) \, \mathrm{sinc}^2(\omega T_1/2) \, [\, (\beta-\alpha) \; + \; \alpha \; 2\pi \, \delta(\omega T_1) \,]$$

Apart from a DC line of amplitude $\alpha = \mu^2$, the spectrum is entirely continuous. If the amplitudes are selected such that $\mu = 0$, then even the DC line goes away.

It turns out that the MLS sequence is not perfectly uncorrelated, so we cannot apply the above theory to get the autocorrelation sequence and the spectral power density. [this item has been integrated in]

Formula FT (35.11) assumes that the autocorrelation function r_k is constant for all non-zero k. If this is not the case, we have to back up to the earlier formula for $\mathscr{P}(\omega)$, FT (35.2), which is completely general. In fact, this was what we had to do for the AMI line code treated in FT Section 37.

We shall soon compute the digital autocorrelation function r_k for an MLS sequence, and then use the result in the formulas of FT Chapter 6 to obtain the frequency spectrum of the MLS sequence. But first, we insert the following brief review of "probability and statistics" in relation to $\langle a_m a_n \rangle$ and r_k .

A few Basics of Probability and Statistics

The probability of *events* A and B both being true is denoted by $P(A \cap B)$, or P(A,B). This is known as the joint second-order probability density function (pdf). A related concept is P(A|B) meaning the conditional probability that A is true given that B is true. Here are the relations:

$$P(A,B) = P(A \cap B) = P(A|B) \, P(B) = P(B|A) \, P(A) \; . \qquad (2.5.5)$$

If events A and B are *statistically independent*, one gets:

$$P(A,B) = P(A \cap B) = P(A) \, P(B) \qquad \Leftrightarrow \qquad P(A|B) = P(A), \quad P(B|A) = P(B) \; . \qquad (2.5.6)$$

The result P(A|B) = P(A) means that the probability of A being true is not influenced by whether or not B is true.

Similarly one can define joint pdf's of third order and beyond. For example, statistical independence for a third-order pdf is indicated by,

$$P(A,B,C) = P(A \cap B \cap C) = P(A)P(B)P(C) \; . \qquad (2.5.7)$$

Case 1: Symbols in {1,0}

The "event" of interest is that $a_n = 1$. The sequence $\{a_n\}$ has some mean value which we denote by μ. The following shorthand notation is useful:

$$p(a_n) \equiv P(a_n=1) = \langle a_n \rangle = \mu$$

$$p(a_n, a_m) \equiv P(a_n= 1, a_m= 1) \quad . \qquad (2.5.8)$$

We can now express various mean values in terms of these pdf's:

$$\langle a_n \rangle = 1 \bullet P(a_n=1) \; + 0 \bullet P(a_n \neq 1) \; = P(a_n=1) = p(a_n) \; = \mu \qquad (2.5.9a)$$

$$\langle a_n^2 \rangle = 1^2 \bullet P(a_n=1) \; + 0^2 \bullet P(a_n \neq 1) \; = P(a_n=1) = p(a_n) = \mu \qquad (2.5.9b)$$

$$\langle a_m a_n \rangle = 1^2 \bullet P(a_m=1,a_n=1) + 3 \;\; \text{zero terms} = \; P(a_m=1,a_n \; =1) = p(a_m, a_n) \; . \qquad (2.5.9c)$$

If $m \neq n$, it is possible that a_m and a_n could be statistically independent. In this case, one would have:

$$\langle a_m a_n \rangle = \; p(a_m, a_n) = \; p(a_m) \, p(a_n) = \langle a_m \rangle \langle a_n \rangle \; = \mu^2 \qquad \text{// uncorrelated} \qquad (2.5.10)$$

For m=n, it is impossible that a_m and a_n be statistically independent, and one has:

$$\langle a_n a_n \rangle = \; p(a_n, a_n) = \langle a_n^2 \rangle = \; p(a_n) = \mu \; \neq \quad p(a_n) \, p(a_n) \qquad \text{// correlated} \qquad (2.5.11)$$

In other words, the "diagonal" second-order statistical quantity $p(a_n, a_n)$ is really the first-order quantity $p(a_n)$.

We can now write down a few related statistical quantities (we are doing GF(2) symbols here):

The moments for N = 0,1,2,3...:

$$m_0 = \langle a_n^0 \rangle = \langle 1 \rangle = 1 \qquad N = 0$$

$$m_N \equiv \langle a_n^N \rangle = 1^N \bullet P(a_n=1) + 0^N \bullet P(a_n=1) = p(a_n) = \mu \qquad N = 1,2,3.. \qquad (2.5.12)$$

The central moments for N = 1,2,3..:

$$\mu_N = \langle (a_n - \langle a_n \rangle)^N \rangle = \langle (a_n - \mu)^N \rangle = \langle \Sigma_{j=0}^N \binom{N}{j} (a_n)^{N-j} [-\mu]^j \rangle$$

$$= \Sigma_{j=0}^N \binom{N}{j} (-1)^j \mu^j \langle (a_n)^{N-j} \rangle = \Sigma_{j=0}^{N-1} \binom{N}{j} (-1)^j \mu^{j+1} + (-1)^N \mu^N \qquad (2.5.13)$$

The N=1,2 central moments

$$\mu_1 = \langle (a_n - \langle a_n \rangle)^1 \rangle = \langle a_n \rangle - \langle a_n \rangle = \mu - \mu = 0$$

$$\mu_2 = \langle (a_n - \langle a_n \rangle)^2 \rangle = \langle (a_n - \mu)^2 \rangle = \langle a_n^2 \rangle - 2\mu \langle a_n \rangle + \mu^2 = \langle a_n^2 \rangle - \mu^2 = \mu - \mu^2 = \mu(1-\mu)$$

$$= \sigma^2 = (\text{standard deviation})^2 = \text{variance} \qquad (2.5.14)$$

Thus, in the special case of symbols in GF(2), all moment and central moments are determined by μ. For some different set of symbol values, this might not be the case.

In general, a full characterization of a statistical sequence requires specification of the probability density functions of all orders. However, much useful information is contained in the first two orders, and this information is sufficient for many applications.

Fact 1: The first and second order statistics are completely determined by the autocorrelation sequence.

Proof: Assume that the autocorrelation sequence r_k is known. Then,

first order:

$$p(a_n) = \langle a_n \rangle = \langle a_n^2 \rangle = r_0 \qquad // \ (2.5.9a), (2.5.9b), (2.54)$$

second order:

$$p(a_n, a_n) = \langle a_n^2 \rangle = p(a_n) = r_0 \qquad // (2.5.9b), (2.54)$$

$$p(a_m, a_n) = \langle a_m a_n \rangle = r_{m-n} \qquad m \neq n \qquad // (2.5.4)$$

Corollary 1: Third and higher order statistics are not determined by the autocorrelation sequence. One needs more information to get these higher statistics.

Example: Consider the third-order density:

$$p(a_n, a_{n+k}, a_{n+k'}) = \langle a_n a_{n+k} a_{n+k'} \rangle$$

One could relate this to some kind of fancier 2-variable autocorrelation sequence defined this way,

$$r_{k,k'} \equiv (1/P) \sum_{n=1}^{P} a_n \, a_{n+k} \, a_{n+k'}$$

Fact 2: For a *statistically independent sequence*, all higher order statistics are determined by the first order statistics: ($p(a_m)$ is independent of m)

$$\langle a_m a_n \rangle = p(a_m, a_n) = p(a_m)\,p(a_n) = [p(a_m)]^2 = \mu^2 \qquad m \neq n$$

$$\langle a_m a_n a_k \rangle = p(a_m, a_n, a_k) = p(a_m)\,p(a_n)\,p(a_k) = [p(a_m)]^3 = \mu^3 \qquad m \neq n \neq k$$

etc.

Proof: This follows from the definition of "statistical independence" given above (= uncorrelated).
ok to here
Definition: A *white* sequence over GF(2) is one which is uncorrelated and has $p(a_n) = \mu = 1/2$.

For such a sequence, we have:

$$\langle a_n \rangle = p(a_n) = \mu = 1/2 = r_0$$

$$\langle a_m a_n \rangle = p(a_m, a_n) = p(a_m)\,p(a_n) = \mu^2 = (1/2)^2 = 1/4 = r_{m-n} \qquad m \neq n$$

$$\langle a_m a_n a_k \rangle = p(a_m, a_n, a_k) = p(a_m)\,p(a_n)\,p(a_k) = \mu^3 = (1/2)^3 = 1/8 \qquad m \neq n \neq k$$

etc. (2.5.15)

From (2.5.15) here is a plot of the autocorrelation sequence for a white sequence a_n defined over GF(2):

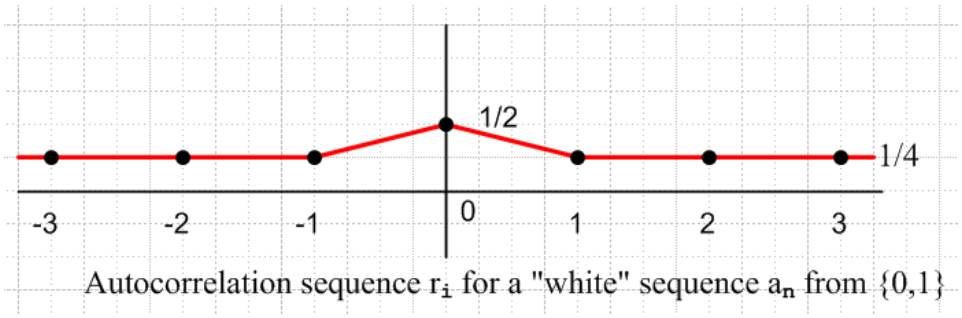


Fig 2.5.1: Autocorrelation function of a white sequence over GF(2) = {0,1}.

This fits the general template shown in Fig *** above where $\alpha = \mu^2 = 1/4$ and $\beta = \sigma^2+\mu^2 = \mu = 1/2$

Case 2: Symbols in {1,-1}

It is sometimes convenient to think of the elements of the sequence a_n as belonging to the set {-1, +1} instead of to the official GF(2) set { 0,1 }. In this case, various statements made above must be modified. For example,

$$P(a_n = 1) = p(a_n) \equiv p$$

$P(a_n \neq 1) = P(a_n = -1) = 1 - p(a_n) = 1-p$

$$\langle a_n \rangle = 1 \bullet p(a_n) + (-1) \bullet [\, 1 - p(a_n)] = P(a_n=1) = [\, 2\, p(a_n) - 1\,] = (2p-1) = \mu \qquad (2.5.9a)'$$

$$\langle a_n^2 \rangle = 1^2 \bullet p(a_n) + (-1)^2 \bullet [\, 1 - p(a_n)] = 1 \qquad (2.5.9b)'$$

$$\langle a_m a_n \rangle = 1^2 \bullet P(a_m=1, a_n=1) + 1 \bullet (-1)\, P(a_m=1, a_n=-1) + (-1) \bullet 1 \bullet P(a_m=-1, a_n=1) + (-1)^2\, P(a_m=1, a_n=-1) \qquad (2.5.9c)'$$

If $m \neq n$, it is possible that a_m and a_n could be statistically independent. In this case, one would have:

$$\langle a_m a_n \rangle = p(a_m, a_n) = p(a_m)\, p(a_n) = \langle a_m \rangle \langle a_n \rangle = (2p-1)^2 = \mu^2 \qquad \text{// uncorrelated} \quad . \qquad (2.5.10)'$$

For m=n, it is impossible that a_m and a_n be statistically independent, and one has:

$$\langle a_n a_n \rangle = p(a_n, a_n) = \langle a_n^2 \rangle = 1 \neq p(a_n)\, p(a_n) = p^2 \quad . \qquad (2.5.11)'$$

<u>The moments for N = 0,1,2,3...:</u>

$$m_0 = \langle a_n^0 \rangle = \langle 1 \rangle = 1 \qquad N = 0$$

$$m_N = \langle a_n^N \rangle = 1^N \bullet p(a_n) + (-1)^N \bullet [\, 1 - p(a_n)] = p + (-1)^N (1-p) \qquad N = 1,2,3.. \qquad (2.5.12)'$$

$$= \begin{cases} 1 & N \text{ even} \\ (2p-1) & N \text{ odd} \end{cases}$$

<u>The central moments for N = 1,2,3..:</u>

$$\mu_N = \langle (a_n - \langle a_n \rangle)^N \rangle = \langle (a_n - p)^N \rangle = \langle \Sigma_{j=0}^N \binom{N}{j} (a_n)^{N-j} [-p]^j \rangle$$

$$= \Sigma_{j=0}^N \binom{N}{j} (-1)^j p^j \langle (a_n)^{N-j} \rangle = \Sigma_{j=0}^N \binom{N}{j} (-1)^j p^j\, m_{N-j} \qquad (2.5.13)'$$

<u>The N=1,2 central moments:</u>

$$\mu_1 = \langle (a_n - \langle a_n \rangle)^1 \rangle = \langle a_n \rangle - \langle a_n \rangle = 0$$

$$\mu_2 = \langle (a_n - \langle a_n \rangle)^2 \rangle = \langle a_n^2 \rangle - 2\langle a_n \rangle^2 + \langle a_n \rangle^2 = \langle a_n^2 \rangle - \langle a_n \rangle^2 = 1 - (2p-1)^2 = 4p(1-p) = 1-\mu^2$$

$$= \sigma^2 = (\text{standard deviation})^2 = \text{variance} \qquad (2.5.14)'$$

Fact 1': The first and second order statistics are completely determined by the autocorrelation function.

Proof: Assume that the autocorrelation function r_k is known. Then,

first order:

$p(a_n) = \langle a_n \rangle = \langle a_n^2 \rangle = r_0$ // (2.5.9a), (2.5.9b), (2.54)

second order:

$p(a_n, a_n) = \langle a_n^2 \rangle = p(a_n) = r_0$ // (2.5.9b), (2.54)

$p(a_m, a_n) = \langle a_m a_n \rangle = r_{m-n}$ $m \neq n$ // (2.5.4)

For a *white sequence*, we have p = 1/2 so

$\langle a_n \rangle = (2p-1) = 0 = \mu$

$\langle a_n^2 \rangle = 1 = r_0$ // (2.5.4)

$\langle a_m a_n \rangle = p(a_m, a_n) = p(a_m)\, p(a_n) = \langle a_n \rangle \langle a_n \rangle = 0 = r_{m-n}$ $m \neq n$ // (2.5.4)

$\langle a_m a_n a_k \rangle = p(a_m, a_n, a_k) = p(a_m)\, p(a_n)\, p(a_k) = 0$ $m \neq n \neq k$

etc. (2.5.15)'

From (2.5.15)' here is a plot of the autocorrelation sequence for a white sequence a_n defined over {-1, +1} where we set p = 1/2 in the above equations:

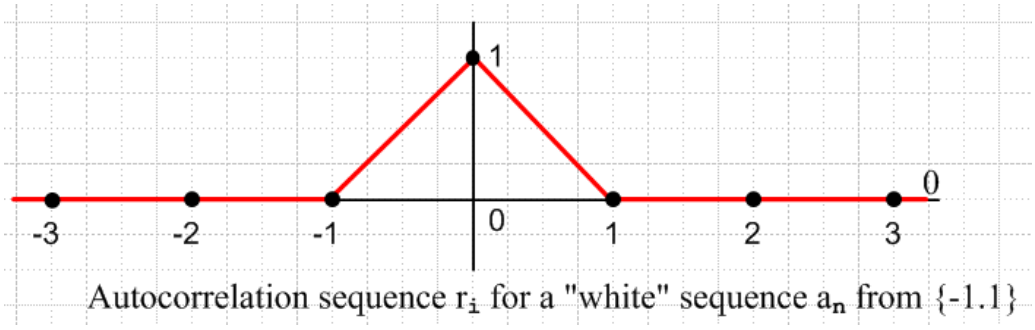


Fig 2.5.2: Autocorrelation function of a white sequence over {-1,+1}.

This fits the general template shown in Fig *** above where now $\alpha = \mu^2 = 0$ and $\beta = \sigma^2 + \mu^2 = 1$.

2.6 Autocorrelation and Spectral Power Density of an MLS Sequence

Consider a shift register generator of k stages having a primitive feedback polynomial h(x). What is the frequency spectrum of the MLS sequence output by this piece of hardware? There are really two different answers to this question, depending on what the output driver looks like. We consider the following two cases:

Case 1: (single-ended) On a 1 we make a pulse $x_{pulse}(t)$, and on a 0 we make no pulse. For 1101,

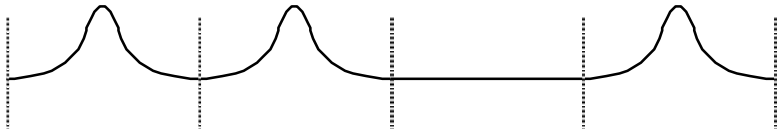


Fig 2.6.1: Encoding of 1101 using arbitrary pulse shape, singled ended drive.

Case 2: (differential) On a 1 we make a pulse $x_{pulse}(t)$, and on a 0 we make a pulse $-x_{pulse}(t)$. For 1101,

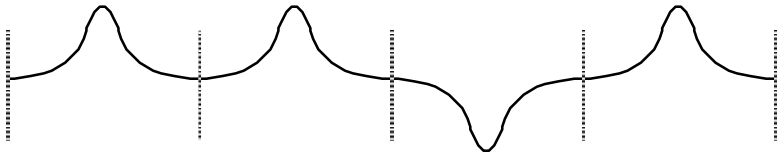


Fig 2.6.2: Encoding of 1101 using arbitrary pulse shape, differential drive.

When $x_{pulse}(t)$ is a box of width T_1 (as it would be for an ideal NRZ line code), Case 1 and Case 2 are related by a factor of 2 scale change and a DC offset. T_1 is the spacing between dotted lines in the above figures. We prefer to deal with an arbitrary pulse shape in the following analysis, so Case 1 and Case 2 are quite different.

In either case, our method of analysis will be the same. We compute the two quantities $\alpha = \langle a_m a_n \rangle$ and $\beta = \langle a_n^2 \rangle$, and we make use of our general formula given in Spectral Chapter 6, boxed equation (6.4.7). We anticipate that these averages will be independent of the subscripts, and so (6.4.7) applies. If this were not the case, we would have to back up to equation (6.3.9) as we did in the analysis of the AMI line code in Spectral Section 6.6(d). In quantity $\alpha = \langle a_m a_n \rangle$ it is implied that $m \neq n$.

Case 1: MLS Sequence Statistics for Symbols in {1,0}

We replicate our table from Section 2.4 above. Recall that this table describes the product of two MLS sequences, and just such a product sequence is what appears summed in the definition of the autocorrelation function, Equation (2.5.2).

The Counts	Column 3	Column 4	Shifted n's	Unshifted n's
n_1 = number of (0,0)'s	sum = 0	product = 0	$(2^k - 1) - 3 \bullet 2^{k-2}$	$2^{k-1} - 1$
n_2 = number of (0,1)'s	sum = 1	product = 0	2^{k-2}	0
n_3= number of (1,0)'s	sum = 1	product = 0	2^{k-2}	0
n_4 = number of (1,1)'s	sum = 0	product = 1	2^{k-2}	2^{k-1}

The product of two MLS sequences consists of 0's and 1's as shown in Column 4. To get $\alpha = \langle a_m a_n \rangle$, we use the Shifted column for the values of the n_i counts, while for $\beta = \langle a_n^2 \rangle$ we use the Unshifted counts column. In these sums, the product=0 combinations make no contribution, so we add up only the product=1 combinations. We state the results for general k, and for k = 9 as an example: ($P = 2^k - 1$)

$$\alpha = \langle a_m a_n \rangle = n_4 / P = \frac{2^{k-2}}{2^k - 1} = \frac{1}{4}\frac{2^k}{2^k - 1} = \frac{1}{4}\frac{P+1}{P} = (1/4)(1+1/P) = r_{m-n} \qquad = 128/511$$

$$\beta = \langle a_n^2 \rangle = n_4 / P = \frac{2^{k-1}}{2^k - 1} = 2\frac{2^{k-2}}{2^k - 1} = (1/2)(1+1/P) = r_0 \qquad = 256/511$$

$$(\beta - \alpha) = 2\frac{2^{k-2}}{2^k - 1} - \frac{2^{k-2}}{2^k - 1} = \frac{2^{k-2}}{2^k - 1} = (1/4)(1+1/P) \qquad = 128/511 \qquad (2.6.1)$$

$$\mu = \langle a_n \rangle = \text{prob of } 1 = \frac{2^{k-1}}{2^k - 1} = (1/2)(1+1/P) = r_0 \qquad \text{// see } *****$$

$$\sigma^2 = \langle a_n^2 \rangle - \langle a_n \rangle^2 = \beta - \mu^2 = r_0 - r_0^2 = r_0(1-r_0) = (1/2)(1+1/P)[\,1 - (1/2)(1+1/P)]$$
$$= (1/2)(1+1/P)\,[\,(1/2)(1-1/P)] = (1/4)(\,1 - [\,\frac{1}{P}\,]^2)$$

Is the MLS sequence correlated or uncorrelated? If it were uncorrelated, we would have $\langle a_m a_n \rangle = \langle a_n \rangle^2$, but from the above we see that

$\langle a_m a_n \rangle = (1/4)(1+1/P)$
$\langle a_n \rangle = (1/2)(1+1/P)$
$\langle a_n \rangle^2 = (1/4)(1+1/P)^2$

so

$$\langle a_m a_n \rangle - \langle a_n \rangle^2 = (1/4)[\,(1+1/P) - (1+1/P)^2] = -\frac{1}{4}(\frac{1}{P} + \frac{1}{P^2})$$

No, the MLS sequence is *not* uncorrelated, but it becomes so as $P \to \infty$. And in this limit $\mu = 1/2$.

From (2.6.1) here is a plot of the autocorrelation function:

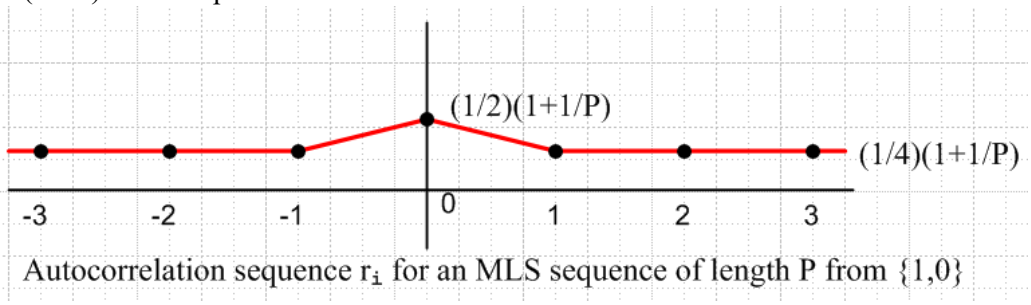


Fig 2.6.3: Autocorrelation function for Case 1 MLS sequence on {1,0}

One can see that as P→∞ this replicates Fig *** for a white sequence in {1,0}

Fact 1: Comparison of Fig 2.6.3 with Fig 2.5.1 shows that the MLS sequence has an autocorrelation function that differs only slightly from that of a white sequence. For k = 9, P = 511 so (1 + 1/P) ≈ 1.002. Thus, the first and second order statistics of an MLS sequence are very close to those of a white sequence.

Proof: See Fact 1 of Section 2.5 above.

The corresponding spectral power density from Spectral (6.4.7) is, [add FT D.32 reference!]

$$\frac{\langle |X(\omega)|^2 \rangle}{T} = \frac{|X_{pulse}(\omega)|^2}{T_1} \{ (1/4)(1 + 1/P) + (1/4)(1+ 1/P) \sum_{m=-\infty}^{\infty} 2\pi\delta(\omega T_1 - 2\pi m) \}$$

update (2.6.2)

Notice that the continuous and discrete spectra have exactly equal strengths, and these strengths are just slightly greater than 1/4. For k=9, we get coefficients 128/511 = .0250489.

The above result is in terms of a general pulse shape. If we set P = ∞ in the above formula, we recover the spectral density of a white sequence. If we also specialize to the case of a square wave pulse, we replicate our earlier result for an NRZ line code having p = 1/2. See Spectral equation (6.5.1).

Case 2: MLS Sequence Statistics for Symbols in {1,-1}

We modify the above table by replacing (0,1) with (+1,-1). We are basically scaling our coefficients by 2 and giving them an offset of -1 unit compared to Case 1. It was already noted that the resulting analog signals are considerably different between Case 2 and Case 1.

The Counts	Column 3	Column 4	Shifted n's	Unshifted n's
n_1 = number of (-1,-1)'s	sum = -2	product = 1	$(2^k - 1) - 3 \cdot 2^{k-2}$	$2^{k-1} - 1$
n_2 = number of (-1,1)'s	sum = 0	product = -1	2^{k-2}	0
n_3 = number of (1,-1)'s	sum = 0	product = -1	2^{k-2}	0
n_4 = number of (1,1)'s	sum = 2	product = 1	2^{k-2}	2^{k-1}

The product of two MLS sequences consists of -1's and 1's as shown in Column 4. To get $\alpha = \langle a_m a_n \rangle$, we use the Shifted column for the values of the n_i counts, while for $\beta = \langle a_n^2 \rangle$ we use the Unshifted counts column. This time, we get contributions from all four possible combinations. We state the results for general k, and for k = 9 as an example: (recall P = length of sequence = $2^k - 1$)

$$\alpha = \langle a_m a_n \rangle = (n_1 - n_2 - n_3 + n_4)/P = -1/P = r_{m-n} = -1/511 \quad \text{explain}$$

$$\beta = \langle a_n^2 \rangle = (n_1 - n_2 - n_3 + n_4)/P = 1 = r_0 = 1 \quad \text{explain}$$

$$(\beta - \alpha) = (1 + 1/P) = 1 + 1/511 \qquad (2.6.3)$$

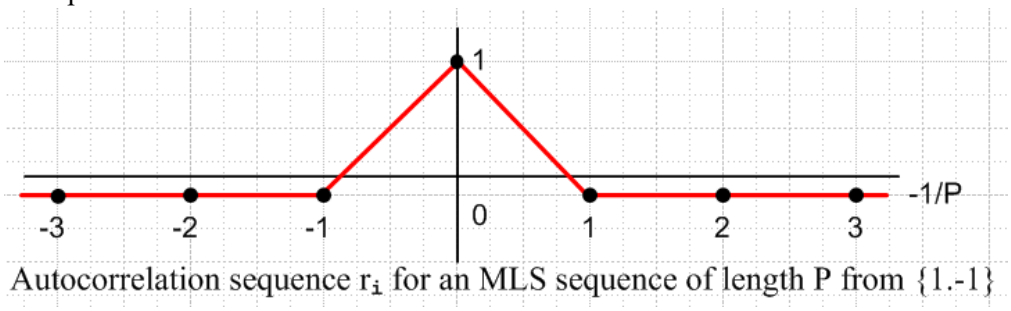
$$\mu = \langle a_n \rangle = 1 \bullet \frac{2^{k-1}}{2^k - 1} + (-1) \bullet \frac{2^{k-1} - 1}{2^k - 1} \quad \text{// see *****}$$

$$= \frac{2^{k-1}}{2^k - 1} - \frac{(1/2)2^k - 1}{2^k - 1} = \frac{P+1}{P} - \frac{(1/2)(P+1) - 1}{P} = \frac{1}{2P}[\,2P + 2 - P - 1 + 2\,] = \frac{P+3}{2P}$$

$$= \frac{1}{2}\left(1 + \frac{3}{P}\right)$$

$$\sigma^2 = \langle a_n^2 \rangle - \langle a_n \rangle^2 = 1 - \mu^2 = 1 - \left(\frac{P+3}{2P}\right)^2 = \frac{3}{4}\left[\,1 - \frac{2}{P} - \frac{3}{P^2}\,\right]$$

Here is a plot:



Autocorrelation sequence r_i for an MLS sequence of length P from {1.-1}

Fig 2.6.4: Autocorrelation function for Case 2 MLS sequence

Upon comparing Fig 2.6.4 with Fig 2.5.2, one sees again that the autocorrelation function for the MLS sequence (now defined over +1, -1) is just slightly different from that of a white sequence.

The corresponding Case 2 spectral power density from (6.4.7) is,

$$\frac{\langle |X(\omega)|^2 \rangle}{T} = \frac{|X_{pulse}(\omega)|^2}{T_1} \left\{\, 1 + 1/P + (-1/P) \sum_{m=-\infty}^{\infty} 2\pi\delta(\omega T_1 - 2\pi m) \,\right\} \qquad (2.6.4)$$

where $P = 2^k - 1$. Now almost all the power is in the continuous spectrum. For a white sequence, we set $P = \infty$ and find that all power is in the continuous spectrum .

For an NRZ signal, Case 1 and Case 2 are essentially the same. They differ only by a factor of 4 due to the factor-of-2 scale change on the height of the pulse. The apparent discrete spectrum in (2.6.2) vanishes due to zeros in the quantity $|X_{pulse}(\omega)|^2$. Therefore:

Fact 2: Apart from a possible DC line, the spectrum of the MLS sequence output by a shift register generator is entirely continuous, and it is the envelope of the pulse shape used.