Meaning of averages PhL 7.23.13

Review of this doc. We are in the context of Appendix F on repeated sequences.

1. I look first at the usual $\mathcal{P}(\omega)$ formula for a *single pulse train* with its $y_m y_n$ double sum DS. I try to make the horizontal $<...>_1$ average appear in it. I do not succeed, there is no connection. The vertical (not used here) is the ensemble average <..>, the horizontal is over one pulse train $<...>_1$.

2. I try to make use of the periodicity of the y_m in evaluating the DS. I draw some diagonal lines pictures and think I am able to show that

$$DS/P^2 = (1/P) \sum_{s=0}^{P-1} e^{-i\omega s T_1} < a_m a_{m+s} >_1$$

But then I realize a math error and this is not valid, the diagonal lines method does nothing useful.

3. Summary to this point: I cannot relate DS to $<...>_1$ average, so I cannot state $\mathcal{P}(\omega)$ for single MLS sequence in terms of $<...>_1$. But I know the WK method with R"(z) allows this to somehow happen.

4. I ask: is $<y_n y_{n+s}>$ independent of n? and same question for $<y_n>$? Here I ponder this question and conclude that in general, these things DO depend on n. But is there some "statistical" ensemble for which they do not depend on n? These <...> are the vertical averages over the ensemble. For a (0,1) symbol pulse train, saying $<y_n> = p$ says all $<y_n>$ are the same, not $<y_n> = p_n$. The function p(x) here is so simple because only two values, so p(x) is this one constant.

I continue to play and conclude that the vertical and horizontal averages are really unrelated. And here I find a wrong statement made in App F and I correct it. I realize that the "autocorrelation method" *can* find $\mathcal{P}(\omega)$ for a single sequence, and there is no other way to do that calculation. My respect for the autocorrelation function increases.

The time has come to clear things up.

Starting point from App F is this

$$\mathcal{P}(\omega) = \mathcal{P}_{pulse}(\omega)\ \omega_1 \sum_{m=-\infty}^{\infty} \delta(\omega - \omega_1 m/P)\ \frac{1}{P^2} \sum_{n=0}^{P-1} \sum_{m=0}^{P-1} y_m y_n\ e^{i\omega(m-n)T_1} \qquad (F.24)$$

1. Let's try to cause <..>$_1$ to appear here

$$DS = \sum_{n=0}^{P-1} \sum_{m=0}^{P-1} y_m y_n\ e^{i\omega(m-n)T_1} = S = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} y_m y_n\ e^{i\omega(m-n)T_1}\ \theta(m)\theta(n)\theta(P-1-m)\ \theta(P-1-n)$$

Let s = n-m so n = m+s,

$$DS = \sum_{m=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} y_m y_{m+s}\ e^{-i\omega s T_1}\ \theta(m)\theta(m+s)\theta(P-1-m)\ \theta(P-1-[m+s])$$

$$DS = \sum_{m=-\infty}^{\infty} \theta(m)\ \theta(P-1-m) \sum_{s=-\infty}^{\infty} y_m y_{m+s}\ e^{-i\omega s T_1}\theta(m+s)\ \theta(P-1-[m+s])$$

$$DS = \sum_{m=0}^{P-1} \sum_{s=-\infty}^{\infty} y_m y_{m+s}\ e^{-i\omega s T_1}\theta(m+s)\ \theta(P-1-[m+s])$$

$$DS = \sum_{m=0}^{P-1} \sum_{s=-m}^{(P-1)-m} y_m y_{m+s}\ e^{-i\omega s T_1}$$

Try different path

$$DS = \sum_{s=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} y_m y_{m+s}\ e^{-i\omega s T_1}\ \theta(m)\theta(m+s)\theta(P-1-m)\ \theta(P-1-[m+s])$$

$$DS = \sum_{s=-\infty}^{\infty} e^{-i\omega s T_1} \sum_{m=\max(0,-s)}^{\min(P-1,P-1-s)} y_m y_{m+s}$$

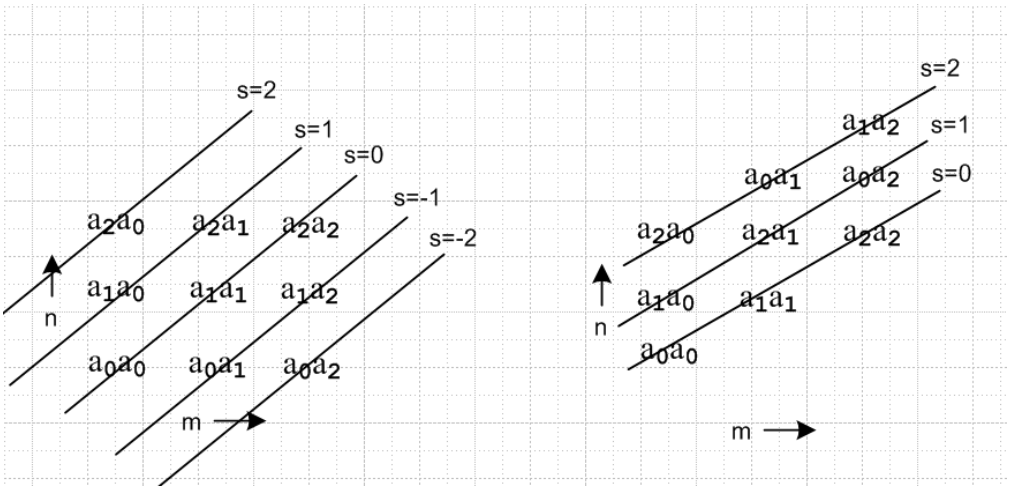
None of these forms helps me much. I cannot write anything as a horizontal sum over the sequence.

2. BUT, let's try to use the periodicity idea

Here is our sum of interest:

$$\sum_{n=0}^{P-1} \sum_{m=0}^{P-1} y_m y_n\ e^{i\omega(m-n)T_1}$$

The picture below left represents adding 9 terms, but I don't show the phases which depend on s = n-m. Since the a_i are periodic with period 3, I can redraw the 9 terms this way



The diagonal lines are those of constant s. On the left, for each value of s, there were a variable number of terms to add, ranging 1,2,3. But on the right each s line has exactly 3 = P terms.

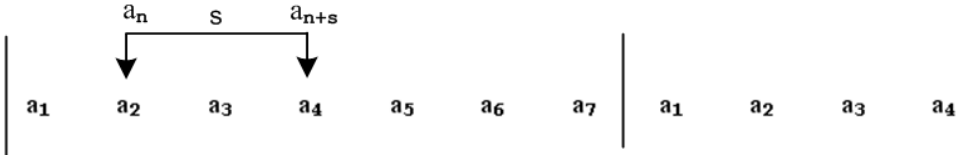
s = n-m so n = m+s

$$DS = \sum_{n=0}^{P-1} \sum_{m=0}^{P-1} a_n a_m e^{i\omega(m-n)T_1} = \sum_{s=0}^{P-1} \sum_{m=0}^{P-1} a_{m+s} a_m e^{-i\omega s T_1}$$

In the sum on the right, there will be a term a_4a_2 when m = 2 and s = 2. This is really a_1a_2. So as long as you know about the periodicity, you CAN write the sum as I show it. Once again

$$\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} y_m y_n e^{i\omega(m-n)T_1} = \sum_{s=0}^{P-1} e^{-i\omega s T_1} \left(\sum_{m=0}^{P-1} a_m a_{m+s} \right)$$

I think now we are getting somewhere. The sum on the right is my horizontal pincer sum where you move the left end to all possible positions m = 0,1,2...P-1 and you sum the products.



Now how about this

$$\frac{1}{P^2} \sum_{n=0}^{P-1} \sum_{m=0}^{P-1} y_m y_n e^{i\omega(m-n)T_1} = (1/P) \sum_{s=0}^{P-1} e^{-i\omega s T_1} \left[(1/P) \sum_{m=0}^{P-1} a_m a_{m+s} \right]$$

$$= (1/P) \sum_{s=0}^{P-1} e^{-i\omega s T_1} < a_m a_{m+s} >_1$$

I have finally found a path to $\langle .. \rangle_1$. So now our double sum is a single sum over these pincer sums for individual values of s.

Fact: Suppose you are dealing with a periodic sequence which has this horizontal average property

$$\langle a_m a_{m+s} \rangle_1 = \beta \qquad s = NP \qquad = \langle a_n^2 \rangle_1$$

$$\langle a_m a_{m+s} \rangle_1 = \alpha \qquad s \neq NP$$

which I think the MLS sequence does have! Then we have

$$\sum_{s=0}^{P-1} e^{-i\omega s T_1} \langle a_m a_{m+s} \rangle_1 = \beta + \alpha \sum_{s=1}^{P-1} e^{-i\omega s T_1}$$

Now let $x = e^{-i\omega T_1}$ and use the little TK rule

$$\sum_{s=1}^{P-1} x^s = [\, x^P - x \,] / (x-1) = [\, e^{-i\omega P T_1} - e^{-i\omega T_1} \,] / [e^{-i\omega T_1} - 1]$$

I thought this had to be real. Here let $k = \omega T_1$, then the sum is

$$[\, e^{-ikP} - e^{-ik} \,] / [e^{-ik} - 1]$$

I set P = 7 and k = 1.3 Maple shows this thing is not real:

```
s := (exp(-I*k*P) - exp(-I*k))/(exp(-I*k) - 1);
```

$$s := \frac{e^{(-I k P)} - e^{(-I k)}}{e^{(-I k)} - 1}$$

```
P := 7: k := 1.3;
```

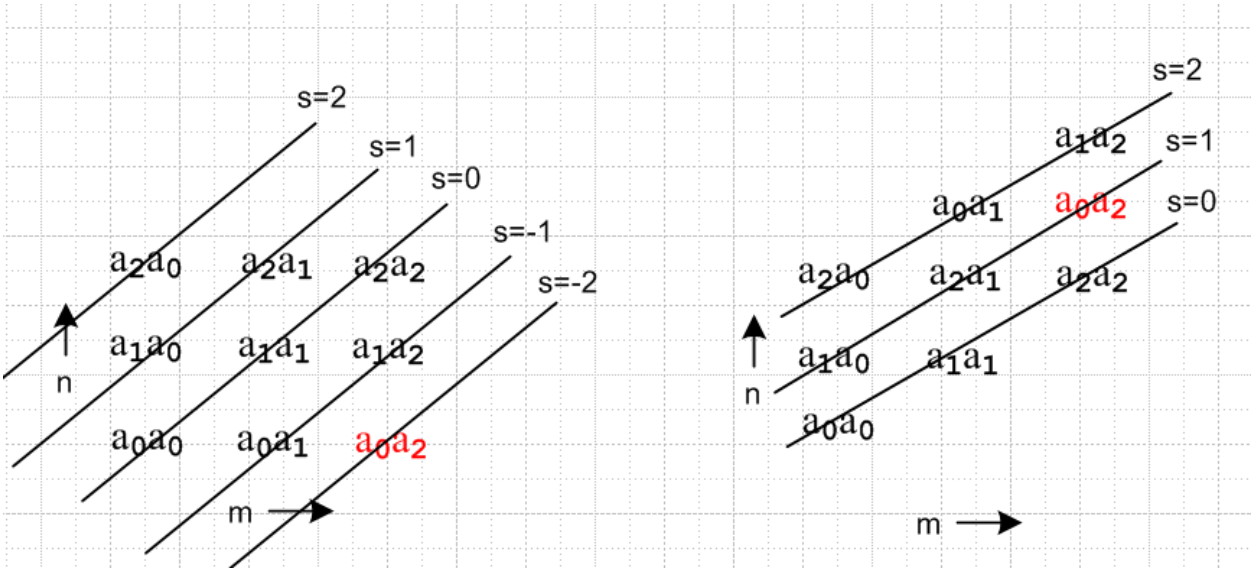
$$k := 1.3$$

```
s;
```

$$.1837374888 - 1.121502076\, I$$

So how did I start with $|\, Y''_P(z)|^2$ which must be real and end up with something not real ???

Go back to



I took the term (and recall that s does $e^{-i\omega s T_1}$)

$a_n a_m \; e^{i\omega(m-n)T_1} = a_0 a_2 \; e^{i\omega(2-0)T_1}$ which has s = -2 (as picture shows)

and I replaced that term with

$a_n a_m \; e^{i\omega(m-n)T_1} = a_0 a_2 \; e^{-i\omega 2T_1}$ which has s = +2.

Thus, because of the phases, the above method does not work! I have altered the sum, and that is why it is no longer real.

3. What I have learned so far:

Consider,

$$\mathcal{P}(\omega) = \mathcal{P}_{pulse}(\omega) \; \omega_1 \sum_{m=-\infty}^{\infty} \delta(\omega - \omega_1 m/P) \; \frac{1}{P^2} \sum_{n=0}^{P-1} \sum_{m=0}^{P-1} y_m y_n \; e^{i\omega(m-n)T_1} \qquad (F.24)$$

I have been unable to write the double sum in terms of the horizontal pincer average $< y_m y_n >_1$. I thought last night this might be possible, and today I did due diligence above trying to find a way.

What do I know for an MLS sequence?

y_n^2 = either 0 or 1

I cannot seem to apply F.24 above to a single MLS sequence! But somehow the R"(z) method allows me to do it. That is my Big Mystery.

4. Consider this App F claim about dependence of things on n.

From now on we assume that the y_m are real. Back in (F.25) we made the assumption that

$$\begin{aligned} &\langle y_m y_n \rangle = \alpha \qquad \text{for } m \neq n \qquad s < |m-n| \\ &\langle y_m y_n \rangle = \beta \qquad \text{for } m = n \end{aligned} \qquad \textit{(F.25)}$$

where <...> were ensemble averages. Here we are going to make the much stricter assumption that (F.25) is true for each individual sequence in the ensemble. Then we have

$$\begin{aligned} &\langle y_m y_n \rangle_1 = \alpha \qquad \text{for } m \neq n \qquad s < |m-n| \\ &\langle y_m y_n \rangle_1 = \beta \qquad \text{for } m = n \end{aligned} \qquad \text{(F.42)}$$

If (F.42) is true, then certainly (F.25) is true.

Let's see if this claim is really true:

$$\langle y_n y_{n+s} \rangle = (1/I) \sum_{i=1}^{I} y_n^{(i)} y_{n+s}^{(i)} \qquad (1)$$

$$\langle y_n y_{n+s} \rangle^{(i)}{}_1 = \lim_{N\to\infty} \left[\frac{1}{(2N+1)} \sum_{n=-N}^{N} y_n^{(i)} y_{n+s}^{(i)} \right] \qquad (2)$$

Lots of issues here. First, is (1) independent of n ? I ask this question somewhere, but I am not sure I liked my answer. Similar question is this: is $\langle y_n \rangle$ independent of n ?

$$\langle y_n \rangle = (1/I) \sum_{i=1}^{I} y_n^{(i)} \qquad (3) \qquad \text{for an ensemble of sequences}$$

$$\langle y_n \rangle_1 = \lim_{N\to\infty} \left[\frac{1}{(2N+1)} \sum_{n=-N}^{N} y_n \right] \qquad (4) \qquad \text{for a single sequence}$$

Item (4) shows n as a dummy summation index so $\langle y_n \rangle_1$ is clearly not dependent on n. It is the average of all the symbols in an infinite sequence, cannot possibly depend on n.

The (3) item is less clear. It is a vertical average through the ensemble, all at position n. You must de facto add an assumption that all pulse train positions are "the same" in some sense.

Suppose the sequence y_i were a set of drop times for a ball in a box (pulse), but the boxes have different gravity. Then the pulses in such a pulse train have something that distinguishes them, and then an object like $\langle y_n \rangle$ = the average of ball drop times might depend on n. Somehow we want to claim this:

Assumption: All positions in an infinite pulse train are indistinguishable

Corollary: $\langle y_3 \rangle = \langle y_5 \rangle$

This is WRONG. Here is my new fact

Fact: The ensemble average $\langle y_n \rangle$ might depend on n for certain ensembles.

Proof: By "ensemble" I mean some set of sequences I choose. I could choose a set that has 0's in column m and 1's in column n and then $\langle y_m \rangle = 0$ and $\langle y_n \rangle = 1$ and clearly $\langle y_m \rangle \neq \langle y_n \rangle$. Also:

Fact: The ensemble average $\langle y_n^2 \rangle$ might depend on n for certain ensembles.

Fact: The ensemble average $\langle y_n y_{n+s} \rangle$ for a fixed value of s might depend on n for certain ensembles. Thus, for some arbitrary ensemble, $\langle y_n y_{n+s} \rangle$ is a function of both n and s.

This seems to say my notion of "all positions are indistinguishable" has no meaning at all, it is hot air.

Let's try to find some ensembles for which the claim is true.

"Statistical Ensemble" of pulse trains.

The symbols are called $y_n^{(i)}$ for i^{th} sequence in the ensemble, and position n. Suppose they are all AMI sequences, for example.

I would like to claim that in a random or statistical ensemble of pulse trains the probability of any symbol being a 1 is p and later perhaps p = 1/2. Look at the ensemble, pick a bit in some sequence. That bit is in a vertical column. For that column, we have $\langle y_n \rangle = p$. These could still be AMI sequences, however. But at least in this case, we have $\langle y_n \rangle$ not depending on n, so I have been able to locate one example.

What might we say horizontally? Suppose we have a random sequence but every 7th bit is a 1. Then consider $\langle a_n a_{n+7} \rangle_1$. Except on the special bits, probably this will be something like p^2, but for the special bits it will peak at 1, and so $\langle a_n a_{n+7} \rangle_1$ will depend on n for such an ensemble. [It will be p^2 of the two positions are uncorrelated, Y_n and Y_n+7]

If we have a periodic sequence of period 7, then $\langle a_n a_{n+7} \rangle_1 = 1$ for all n.

Lesson: Be very careful what ensemble you have in mind and also what kinds of sequences you have in mind. I think for a random ensemble of AMI sequences, $\langle a_n a_{n+s} \rangle$ depends on s and not on n.

Now go back to the assumptions stated above

$$\begin{aligned} \langle y_m y_n \rangle_1 &= \alpha \quad \text{for } m \neq n \quad s < |m-n| \\ \langle y_m y_n \rangle_1 &= \beta \quad \text{for } m = n \end{aligned} \tag{F.42}$$

Let's write it as

$$\begin{aligned} \langle y_n\, y_{n+s} \rangle_1 &= \alpha \quad s = NP \\ \langle y_n\, y_{n+s} \rangle_1 &= \beta \quad s \neq NP \end{aligned} \tag{F.44}$$

Suppose this really were true for a particular sequence i, so that:

$$\begin{aligned} &\langle y_n^{(i)}\, y_{n+s}^{(i)} \rangle_1 = \alpha \qquad s = NP \\ &\langle y_n^{(i)}\, y_{n+s}^{(i)} \rangle_1 = \beta \qquad s \neq NP \end{aligned} \tag{F.44}$$

It claims no dependence on n. Suppose this were true for *every* sequence in an ensemble of I sequences. Then consider this column average

$$\langle y_n\, y_{n+s} \rangle = (1/I)\ \Sigma_{i=1}^{I}\, y_n^{(i)}\, y_{n+s}^{(i)}$$

I don't think there is any connection !! Suppose in column n of the sequences all bits are contrived to be 0. Then $\langle y_n\, y_{n+0} \rangle = 0$ for particular column n. But that would not rule out (F.44) above. Each sequence might be the same MLS sequence and I have just shifted them to get a column of zeros. So

Fact: $\langle y_n^{(i)}\, y_{n+0}^{(i)} \rangle_1 = \beta$ for each sequence i in an ensemble $\not\Rightarrow$ $\langle y_n\, y_{n+0} \rangle = \beta$.

The reason is that one average is vertical and the other is horizontal and they are just different animals. That means

Fact : (F.44) $\not\Rightarrow$ (F.25)

So this statement I made in App F is wrong. So go try to clean that up. I have made red my strange paragraph about uncorrelated, will return to it later. Nothing until down in (e) with the R"(z) business.

OK, I have repaired Appendix F (I think). Here are my conclusions there:

If

$$\begin{aligned} &\langle y_m y_n \rangle = \alpha \qquad \text{for } m \neq n \qquad s < |\, m-n\, | \\ &\langle y_m y_n \rangle = \beta \qquad \text{for } m = n \end{aligned} \tag{F.25}$$

then // ensemble

$$\langle \mathcal{P}(\omega) \rangle = \mathcal{P}_{pulse}(\omega)\ \omega_1 \sum_{m=-\infty}^{\infty} [\, (\beta-\alpha)\, \frac{1}{P}\, \delta(\omega - \omega_1 m/P) + \alpha\, \delta(\omega - \omega_1 m)\,] \tag{F.52}$$

If

$$\begin{aligned} &\langle y_n\, y_{n+s} \rangle_1 = \alpha \qquad s = NP \\ &\langle y_n\, y_{n+s} \rangle_1 = \beta \qquad s \neq NP \end{aligned} \tag{F.44}$$

then // single sequence

$$\mathcal{P}(\omega) = \mathcal{P}_{pulse}(\omega)\ \omega_1 \sum_{m=-\infty}^{\infty} [\, (\beta-\alpha)\, \frac{1}{P}\, \delta(\omega - \omega_1 m/P) + \alpha\, \delta(\omega - \omega_1 m)\,] \tag{F.52}$$

Question: Why is the "autocorrelation method" able to get this single-sequence result, whereas I cannot seem to achieve this result without use of the autocorrelation function???

This is just the way it is, so I will soften my comments in FT and Scrambler about autocorrelation.