This entire Appendix D was deleted and replaced with a totally different App D.

I will keep this file in case it is needed. This is before I started talking about sources being stationary, so here I had to lean on my "useful ensemble" to get that property. One main idea is the notion of replacing X,Y with Y_m,Y_n and I just added a piece from below to the very end of App G. Reviewing is good!

Appendix D: How α and β are related to μ and σ

We first review some of the basics of random variables X and Y: expected values, means, covariance, variance and standard deviation. We do this in a context in which each variable takes on a set of discrete values rather than continuous values. We then translate our various facts into the language of X,Y → Y_m Y_n where the latter are random variables associated with the amplitudes at positions m and n of a pulse train. The labels m and n add a notational complexity that is a little difficult to comprehend unless things are written out explicitly, which is done below.

Appendix G contains "the long version" of this appendix. Here, we quietly avoid the question of "What *is* a random variable ?", and "Why are positions in a pulse train statistically equivalent ?" . These and related questions are addressed in some detail in Appendix G to which the reader is referred. Here we are mainly interested in relating our pulse train parameters α and β to statistical quantities μ and σ in the case of an uncorrelated pulse train.

(a) Random Variables X and Y

If X and Y are random variables which can take discrete values x_k and y_k, one can express the mean value of each variable as

$$\mu_x = E(X) = \Sigma_k\, x_k\, p_x(x_k)$$
$$\mu_y = E(Y) = \Sigma_k\, y_k\, p_y(y_k) \qquad \text{(D.1a)}$$

where E(X) means the expected value of X. The sum Σ_k is over all values in the set of values that X or Y can take. Probability $p_x(x_k)$ is the probability that X takes the value x_k and similarly for $p_y(y_k)$. If we have a very large ensemble of I systems we can measure the above quantities in this manner,

$$\mu_x = E(X) = (1/I)\, \Sigma_{i=1}^{I}\, x^{(i)}$$
$$\mu_y = E(Y) = (1/I)\, \Sigma_{i=1}^{I}\, y^{(i)} \qquad \text{(D.1b)}$$

where i labels a particular system in the ensemble, and $x^{(i)}$ and $y^{(i)}$ are the values X and Y take in ensemble system i. The probability functions like $p_x(x_k)$ do not appear in (D.1b), but they are present in spirit, being embedded in the "experimental data" like $x^{(i)}$.

The expected value E(XY) is defined by

$$E(XY) = \Sigma_k \Sigma_j\, x_k\, y_j\, p_{xy}(x_k, y_j) \qquad \text{(D.2a)}$$

where now $p_{xy}(x_k,y_j)$ is the **joint probability** that X has value x_k when Y has the value y_j. One would measure E(XY) in this manner,

$$E(XY) = (1/I) \Sigma_{i=1}^{I} x^{(i)} y^{(i)} . \qquad (D.2b)$$

If the random variables X and Y are **independent** (**uncorrelated**), then

$$p_{xy}(x_k,y_j) = p_x(x_k)\, p_y(y_j) \qquad \text{// uncorrelated} \qquad (D.3)$$

and then it is obvious from the above definitions of E(XY), E(X) and E(Y) that

$$E(XY) = E(X)E(Y) . \qquad \text{// uncorrelated} \qquad (D.4)$$

The **covariance** of X and Y is defined as,

$$cov(X,Y) = E([X - \mu_x] [Y - \mu_y]) = \Sigma_k \Sigma_j (x_k - \mu_x) (y_j - \mu_y) p_{xy}(x_k,y_j) . \qquad (D.5a)$$

It could be measured for an ensemble in this way

$$cov(X,Y) = (1/I) \Sigma_{i=1}^{I} (x^{(i)} - \mu_x) (y^{(i)} - \mu_y) \qquad (D.5b)$$

where μ_x and μ_y would be the measured values of E(X) and E(Y) as shown in (D.1b). Note that,

$$cov(X,Y) = E([X - \mu_x] [Y - \mu_y]) = E(XY) - \mu_x E(Y) - \mu_y E(X) + \mu_x \mu_y E(1)$$

$$= E(XY) - \mu_x \mu_y - \mu_y \mu_x + \mu_x \mu_y$$

$$= E(XY) - \mu_x \mu_y . \qquad (D.6)$$

If X and Y are uncorrelated random variables then

$$cov(X,Y) = E(XY) - \mu_x \mu_y = E(X)E(Y) - \mu_x \mu_y = \mu_x \mu_y - \mu_x \mu_y = 0 . \qquad \text{// uncorrelated} \qquad (D.7)$$

In the special case that X and Y are the same variable, we write cov(X,X) = var(X), the **variance** of X, which is also defined as the square of the **standard deviation** σ_x,

$$var(X) \equiv \sigma_x^2 \equiv cov(X,X) = E(X^2) - \mu_x^2 \qquad (D.8a)$$

with experimental measurement from (D.5b) as

$$var(X) \equiv \sigma_x^2 \equiv cov(X,X) = (1/I) \Sigma_{i=1}^{I} (x^{(i)} - \mu_x)^2 . \qquad (D8.b)$$

Finally, the **correlation** of X and Y is defined as

$$corr(X,Y) = cov(X,Y) / [\sigma(X)\sigma(Y)] \qquad (D.9)$$

and we get the satisfying result from (D.7) that if X and Y are uncorrelated, corr(X,Y) = 0.

(b) Random Variables Y_m and Y_n

We now take a special case of X and Y as follows:

X = Y_m = random variable associated with the amplitudes $y_m^{(i)}$ at pulse location m in pulse train
Y = Y_n = random variable associated with the amplitudes $y_n^{(i)}$ at pulse location n in pulse train

Things are a little complicated due to the extra labels m and n, but we can nevertheless carefully translate all the X and Y equations above. Also, we shall add the following extra notation for an expected value to be consistent with our earlier work,

$$\langle x \rangle \equiv E(X) \qquad \langle xy \rangle \equiv E(XY) \qquad \langle x^2 \rangle \equiv E(X^2) \ .$$

Here then are all the equations of the previous section translated from X,Y to Y_m,Y_n . We have primed the equation numbers of the previous section.

$$\mu_{Ym} = E(Y_m) = \Sigma_k \ y_m^{(k)} \ p_{Ym}(y_m^{(k)}) = \langle y_m \rangle$$
$$\mu_{Yn} = E(Y_n) = \Sigma_k \ y_n^{(k)} \ p_{Yn}(y_n^{(k)}) = \langle y_n \rangle \qquad \text{(D.1a)'}$$

$$\mu_{Ym} = E(Y_m) = (1/I) \ \Sigma_{i=1}^{I} \ y_m^{(i)} = \langle y_m \rangle$$
$$\mu_{Yn} = E(Y_n) = (1/I) \ \Sigma_{i=1}^{I} \ y_n^{(i)} = \langle y_n \rangle \qquad \text{(D.1b)'}$$

$$E(Y_m Y_n) = \Sigma_k \Sigma_j \ y_m^{(k)} \ y_n^{(j)} \ p_{YmYn}(y_m^{(k)}, y_n^{(j)}) = \langle y_m y_n \rangle \qquad \text{(D.2a)'}$$

$$E(Y_m Y_n) = (1/I) \ \Sigma_{i=1}^{I} \ y_m^{(i)} \ y_n^{(i)} = \langle y_m y_n \rangle \qquad \text{(D.2b)'}$$

$$p_{YmYn}(y_m^{(k)}, y_n^{(j)}) = p_{Ym}(y_m^{(k)}) \ p_{Yn}(y_n^{(j)}) \qquad \text{// uncorrelated} \qquad \text{(D.3)'}$$

$$E(Y_m Y_n) = E(Y_m) \ E(Y_n) \ \text{ or } \ \langle y_m y_n \rangle = \langle y_m \rangle \langle y_n \rangle \qquad \text{// uncorrelated} \qquad \text{(D.4)'}$$

$$\text{cov}(Y_m, Y_n) = E(\ [Y_m - \mu_{Ym}] \ [Y_n - \mu_{Yn}] \ ) = \Sigma_k \Sigma_j \ (y_m^{(k)} - \mu_{Ym}) \ (y_n^{(j)} - \mu_{Yn}) \ p_{YmYn}(y_m^{(k)}, y_n^{(j)}) \qquad \text{(D.5a)'}$$

$$\text{cov}(Y_m, Y_n) = (1/I) \ \Sigma_{i=1}^{I} \ (y_m^{(i)} - \mu_{Ym}) \ (y_n^{(i)} - \mu_{Yn}) \qquad \text{(D.5b)'}$$

$$\text{cov}(Y_m, Y_n) = E(Y_m Y_n) - \mu_{Ym} \ \mu_{Yn} = \langle y_m y_n \rangle - \mu_{Ym} \ \mu_{Yn} \qquad \text{(D.6)'}$$

$$\text{cov}(Y_m, Y_n) = 0 \ . \qquad \text{// uncorrelated} \qquad \text{(D.7)'}$$

$$\text{var}(Y_m) \equiv \sigma_{Ym}^2 \equiv \text{cov}(Y_m, Y_m) = E(Y_m^2) - \mu_{Ym}^2 = \langle y_m^2 \rangle - \mu_{Ym}^2 \qquad \text{(D.8a)'}$$

$$\text{var}(Y_m) \equiv \sigma_{Ym}^2 \equiv \text{cov}(Y_m, Y_m) = (1/I) \ \Sigma_{i=1}^{I} \ (y_m^{(i)} - \mu_{Ym})^2 \qquad \text{(D8.b)'}$$

$$\text{corr}(Y_m, Y_n) = \text{cov}(Y_m, Y_n) / [\sigma(Y_m)\sigma(Y_n)] \qquad \text{(D.9)'}$$

(c) Statistical Pulse Trains

The arm-waving argument is now made that for a large set of random pulse trains, there should be no distinction between different positions (like m and n) in the pulse train. This argument seems particularly compelling if the pulse trains of the ensemble are each infinitely long. This is really just an intuitive argument, and the notion is supported quantitatively by the "useful ensemble" discussion in Appendix G (g), and we shall refer to such an ensemble of pulse trains as being a "statistical ensemble". With no further comment here, we shall accept this fact that no position is different from any other, and this leads to a simplification of the equations above. In particular,

$$\begin{aligned} &\mu_{Y_m} = \mu_{Y_n} \equiv \mu, \\ &\sigma_{Y_m} = \sigma_{Y_n} \equiv \sigma \\ &p_{Y_m}(z) = p_{Y_n}(z) = p(z) \ . \end{aligned} \qquad \text{(D.10)}$$

One other result (as shown in Appendix G), is this fact, written two ways

$$\begin{aligned} &\langle y_m y_n \rangle = \text{depends only on m-n} \\ &\langle y_m y_{m+k} \rangle = \text{depends only on k} \ . \end{aligned} \qquad \text{(D.11)}$$

With these simplifications, we write our equation set a third time, now with double-primed numbers,

$$\begin{aligned} &\mu = E(Y_m) = \Sigma_k \ y_m^{(k)} \ p(y_m^{(k)}) = \langle y_m \rangle \\ &\mu = E(Y_n) = \Sigma_k \ y_n^{(k)} \ p(y_n^{(k)}) = \langle y_n \rangle \end{aligned} \qquad \text{(D.1a)"}$$

$$\begin{aligned} &\mu = E(Y_m) = (1/I) \Sigma_{i=1}^{I} y_m^{(i)} = \langle y_m \rangle \\ &\mu = E(Y_n) = (1/I) \Sigma_{i=1}^{I} y_n^{(i)} = \langle y_n \rangle \end{aligned} \qquad \text{(D.1b)"}$$

$$E(Y_m Y_n) = \Sigma_k \Sigma_j \ y_m^{(k)} \ y_n^{(j)} \ p_{Y_m Y_n}(y_m^{(k)}, y_n^{(j)}) = \langle y_m y_n \rangle \qquad \text{(D.2a)"}$$

$$E(Y_m Y_n) = (1/I) \Sigma_{i=1}^{I} y_m^{(i)} \ y_n^{(i)} = \langle y_m y_n \rangle \qquad \text{(D.2b)"}$$

$$p_{Y_m Y_n}(y_m^{(k)}, y_n^{(j)}) = p(y_m^{(k)}) \ p(y_n^{(j)}) \qquad \text{// uncorrelated} \qquad \text{(D.3)"}$$

$$E(Y_m Y_n) = E(Y_m) \ E(Y_n) \ \text{ or } \ \langle y_m y_n \rangle = \langle y_m \rangle \langle y_n \rangle \qquad \text{// uncorrelated} \qquad \text{(D.4)"}$$

$$\text{cov}(Y_m, Y_n) = E([Y_m - \mu] [Y_n - \mu]) = \Sigma_k \Sigma_j \ (y_m^{(k)} - \mu) \ (y_n^{(j)} - \mu) \ p_{Y_m Y_n}(y_m^{(k)}, y_n^{(j)}) \qquad \text{(D.5a)"}$$

$$\text{cov}(Y_m, Y_n) = (1/I) \Sigma_{i=1}^{I} \ (y_m^{(i)} - \mu) \ (y_n^{(i)} - \mu) \qquad \text{(D.5b)"}$$

$$\text{cov}(Y_m, Y_n) = E(Y_m Y_n) - \mu^2 = \langle y_m y_n \rangle - \mu^2 \qquad \text{(D.6)"}$$

$\text{cov}(Y_m, Y_n) = 0$. // uncorrelated (D.7)"

$\text{var}(Y_m) \equiv \sigma^2 \equiv \text{cov}(Y_m, Y_m) = E(Y_m^2) - \mu^2 = \langle y_m^2 \rangle - \mu^2$ (D.8a)"

$\text{var}(Y_m) \equiv \sigma^2 \equiv \text{cov}(Y_m, Y_m) = (1/I) \Sigma_{i=1}^{I} (y_m^{(i)} - \mu)^2$ (D8.b)"

$\text{corr}(Y_m, Y_n) = \text{cov}(Y_m, Y_n) / [\sigma^2]$ (D.9)"

Of special interest are these equations taken from the above list :

$\langle y_m \rangle = (1/I) \Sigma_{i=1}^{I} y_m^{(i)} = \mu$ no dependence on n *(D.1a)"*

$\langle y_m y_n \rangle = (1/I) \Sigma_{i=1}^{I} y_m^{(i)} y_n^{(i)}$ depends on m-n *(D.2b)"*

$\langle y_m^2 \rangle = (1/I) \Sigma_{i=1}^{I} [y_m^{(i)}]^2 = \sigma^2 + \mu^2$ no dependence on n *(D.8a)"*

Since $\langle y_m^2 \rangle = \sigma^2 + \mu^2$ is independent of m, we can give it the name β so that

$\beta = \langle y_m^2 \rangle = \sigma^2 + \mu^2$. (D.12)

For *certain kinds* of pulse trains, $\langle y_m y_n \rangle$ does not depend on m-n and is in fact then independent of *both* indices m and n. In this case we define the constant $\alpha = \langle y_m y_n \rangle$ for $m \neq n$. So we then have

$\alpha = \langle y_m y_n \rangle$ for $m \neq n$ // for certain kinds of pulse trains of our statistical type

$\beta = \langle y_m^2 \rangle = \sigma^2 + \mu^2$ // for any pulse train of our statistical type (D.13)

One of those special kinds of pulse trains is one in which Y_n and Y_m are uncorrelated for $m \neq n$. In this case we find

$\alpha = \langle y_m y_n \rangle = \langle y_m \rangle \langle y_n \rangle = \mu * \mu = \mu^2$ // uncorrelated (D.14)

and in this case we then have

$\mu \equiv \langle y_m \rangle$
$\alpha = \langle y_m y_n \rangle = \mu^2$ // uncorrelated
$\beta = \sigma^2 + \mu^2$. (D.15)

We can now restate equation (35.6) in this more statistically-oriented manner,

$\langle y_m \rangle = \mu = [p]A + [1-p]B$

$n \neq m \quad \alpha \equiv \langle y_m y_n \rangle = \mu^2 = \{ [p]A + [1-p]B \}^2 = [pp]\,AA + [p(1-p)]\,2AB + [(1-p)(1-p)]BB$

$$n = m \qquad \beta \equiv \langle y_m^2 \rangle = [p]AA + [(1-p)]\,BB = \sigma^2 + \mu^2 \tag{D.16}$$

which is to say

$$\alpha = \mu^2$$
$$\beta = \mu^2 + \sigma^2 = \alpha + \sigma^2$$
$$\sigma^2 = (\beta - \alpha)\,. \tag{D.17}$$

Using the above expressions for α and β we can have Maple compute σ^2:

```
alpha := (p*A + (1-p)*B)^2;
```

$$\alpha := (p\,A + (1-p)\,B)^2$$

```
beta := p*A^2+(1-p)*B^2;
```

$$\beta := p\,A^2 + (1-p)\,B^2$$

```
var := beta-alpha: factor(%);
```

$$-p\,(B-A)^2\,(-1+p)$$

which says

$$\sigma^2 = \text{variance} = [\,p(1-p)]\,(A-B)^2\,. \tag{D.18}$$

If p = 0 or p = 1, this variance vanishes as one would expect.