Exam, 2008-2009 Solutions

Q1. Power spectrum of synaptic shots.

Consider a model synaptic potential (sometimes called an "alpha-function"), namely

$G(t;G_0,n,\tau_a,\tau_b)=G_0\frac{t^n}{n!}\left(\frac{1}{\tau_a^{\ n+1}}e^{-t/\tau_a}-\frac{1}{\tau_b^{\ n+1}}e^{-t/\tau_b}\right)$*, and assume that these events occur in a Poisson fashion, at rate* λ*. What is the power spectrum of the resulting process?*

The power spectrum of a shot noise process is given by $\lambda\left|\tilde{G}(\omega)\right|^2$, where $\tilde{G}(\omega)$ is the Fourier transform of the shot shape (Noise and Variability Homework 1). To calculate $\tilde{G}(\omega)$ for $G(t;G_0,n,\tau_a,\tau_b)=G_0\frac{t^n}{n!}\left(\frac{1}{\tau_a^{\ n+1}}e^{-t/\tau_a}-\frac{1}{\tau_b^{\ n+1}}e^{-t/\tau_b}\right)$ we treat it term by term. For

$f_n(t)=\frac{t^n}{\tau^{n+1}}e^{-t/\tau}$,

$\tilde{f}_n(\omega)=\int_0^\infty\frac{t^n}{\tau^{n+1}}e^{-t/\tau}e^{-i\omega t}dt$.

For $n\ge 1$, integrate by parts, with $u=t^n$, $du=nt^{n-1}dt$, $dv=e^{-t(1/\tau+i\omega)}dt$, $v=\frac{e^{-t(1/\tau+i\omega)}}{1/\tau+i\omega}$, so

$\tilde{f}_n(\omega)=\int_0^\infty\frac{t^n}{\tau^{n+1}}e^{-t/\tau}e^{-i\omega t}dt=\frac{n}{1+i\omega\tau}\int_0^\infty\frac{t^{n-1}}{\tau^n}e^{-t/\tau}e^{-i\omega t}dt=\frac{n}{1+i\omega\tau}\tilde{f}_{n-1}(\omega)$. Iterating,

$\tilde{f}_n(\omega)=\frac{n!}{\left(1+i\omega\tau\right)^n}$.

Therefore,

$\tilde{G}(\omega)=G_0\left(\frac{1}{\left(1+i\omega\tau_a\right)^n}-\frac{1}{\left(1+i\omega\tau_b\right)^n}\right)$, and the power spectrum is $\lambda\left|\tilde{G}(\omega)\right|^2$.

Q2. How does the Wiener representation depend on input mean?
Recall that for a linear-nonlinear-linear sandwich L_1NL_2, where N is characterized by an input-output relationship f, the nth Wiener kernel is given by

$K_n(t_1,t_2,...,t_n) = c_n \int L_1(t_1-\tau)L_1(t_2-\tau)...L_1(t_n-\tau)L_2(\tau)d\tau$ *, where c_n is the nth coefficient in the orthogonal expansion of f with respect to Hermite polynomials based on the Gaussian that emerges from from L_1 . Determine how c_n depends on the mean of the input signal.*

(This proceeds along the lines of Nonlinear Systems Analysis, Homework 2). Say the mean of the signal emerging from L_1 is μ. Then

$$c_n = \frac{1}{n!P^n} \int_{-\infty}^{\infty} f(x)h_n(x;\mu,P)Gau(x,\mu,P)dx, \tag{1}$$

where $Gau(x,\mu,P) = \frac{1}{\sqrt{2\pi P}} e^{-(x-\mu)^2/2P}$ and the Hermite polynomials $h_n(x;\mu,P)$ are translations of the standard Hermites (because the Gaussian has also been translated): $h_n(x;\mu,P) = h_n(x-\mu;P)$. Here $h_n(x;P)$ are the Hermite polynomials with respect to a Gaussian of power P, centered at 0 (e.g., nonlinear systems theory notes, page 23).

So,

$$\frac{\partial c_n}{\partial \mu} = \frac{1}{n!P^n} \frac{1}{\sqrt{2\pi P}} \frac{\partial}{\partial \mu} \int_{-\infty}^{\infty} f(x)h_n(x-\mu;P)e^{-(x-\mu)^2/2P}dx$$

Carrying out the derivative by using the product rule (and using the relationship $\frac{\partial}{\partial x} h_n(x;P) = nh_{n-1}(x;P)$ -- see nonlinear systems theory notes page 23) we find

$$\frac{\partial c_n}{\partial \mu} = \frac{1}{n!P^n} \frac{1}{\sqrt{2\pi P}} \int_{-\infty}^{\infty} f(x)\left(-nh_{n-1}(x-\mu;P) + \frac{(x-\mu)}{P} h_n(x-\mu;P)\right) e^{-(x-\mu)^2/2P}dx. \tag{2}$$

Now, use the recursion (nonlinear systems theory notes page 23, or, prove it from the generating function, as in NLSA homework 2) relationship: $h_{n+1}(x;P) = xh_n(x;P) - nPh_{n-1}(x;P)$. With this, (2) becomes

$$\frac{\partial c_n}{\partial \mu} = \frac{1}{n!P^n} \frac{1}{\sqrt{2\pi P}} \int_{-\infty}^{\infty} f(x)\frac{1}{P} h_{n+1}(x-\mu;P)e^{-(x-\mu)^2/2P}dx,$$

or

$$\frac{\partial c_n}{\partial \mu} = \frac{1}{n!P^{n+1}} \int_{-\infty}^{\infty} f(x)h_{n+1}(x-\mu;P)Gau(x;\mu,P)dx \tag{3}$$

Using eq. (1) as

$c_{n+1} = \frac{1}{(n+1)!P^{n+1}} \int_{-\infty}^{\infty} f(x) h_{n+1}(x;\mu,P) Gau(x,\mu,P) dx$, eq. (3) becomes

$$\frac{\partial c_n}{\partial \mu} = (n+1) c_{n+1} . \tag{4}$$

Q3. Relationships among m-sequences.

The polynomial $x^6 = x + 1$ *generates a finite (Galois) field of size* 2^6 *and, consequently, an m-sequence of length 63. Recall that an m-sequence can be created by reading down a column of the table of powers of x. The first portion of this table is presented here:*

k	x^k*:*	*coef of* x^5	*coef of* x^4	*coef of* x^3	*coef of* x^2	*coef of* x^1	*coef of* x^0
0		0	0	0	0	0	1
1		0	0	0	0	1	0
2		0	0	0	1	0	0
3		0	0	1	0	0	0
4		0	1	0	0	0	0
5		1	0	0	0	0	0
6		0	0	0	0	1	1
7		0	0	0	1	1	0
8		0	0	1	1	0	0
9		0	1	1	0	0	0
10		1	1	0	0	0	0
11		1	0	0	0	1	1
12		0	0	0	1	0	1
13		0	0	1	0	1	0
14		0	1	0	1	0	0

Consider the transformation $y = x^c$*. This generates a sequence of coefficients* $y^k = x^{ck}$*, consisting of taking every c-th coefficient from the original m-sequence.*

In answering the questions below, it will be useful to recall the following property of finite fields. Since the nonzero elements form a group of size 63 under multiplication, any nonzero element z satisfies $z^{63} = 1$*. (This corresponds to the fact that every nonzero element is some power of x, and* $x^{63} = 1$*.)*

A. Generate the sequence that results for c=2. How does it relate to the original m-sequence? Is it an m-sequence? If so, find the corresponding polynomial that generates it.

For c=2, $y^k = x^{ck}$ generates the same sequence, albeit shifted. There are two (related) ways to show this. The first is to show that $y = x^2$ satisfies the same polynomial as *x*. To see this, square both sides of $x^6 = x + 1$. This results in $x^{12} = x^2 + 2x + 1 = x^2 + 1$, i.e., $y^6 = y + 1$. So powers of *y* must generate the same table as powers of *x*.

The second way to show that $y = x^2$ generates the same table as *x* is to show that the map $\psi(u) = u^2$ is an automorphism of the finite field (that is, it is a one-to-one mapping that preserves the field operations.) To show that ψ preserves the field operations:
$\psi(uv) = (uv)^2 = u^2v^2 = \psi(u)\psi(v)$, and
$\psi(u+v) = (u+v)^2 = u^2 + 2uv + v^2 = u^2 + v^2 = \psi(u) + \psi(v)$.
To show that ψ is one-to-one, we produce its inverse. Observe that
$\psi^2(u) = \psi(\psi(u)) = \psi(u^2) = u^4$, $\psi^3(u) = \psi(\psi^2(u)) = \psi(u^4) = u^8$, so in general

$\psi^r(u) = u^{(2^r)}$. In particular, $\psi^6(u) = u^{64} = u^{63+1} = u$, because $u^{63} = 1$. So if $\psi^6(u) = u$, then $\psi(\psi^5(u)) = u$, so $\psi^{-1} = \psi^5$.

Since the m-sequence is the same (and since $y = x^2$ does not alter the field operations), *y* satisfies the same polynomial as *x*, namely, $y^6 = y + 1$.

B. Generate the sequence that results for c=62. How does it relate to the original m-sequence? Is it an m-sequence? If so, find the corresponding polynomial that generates it.

As in part A, since every field element *z* satisfies $z^{63} = 1$, it follows that $y = x^{62}$ is equivalent to $y = x^{-1}$. So the m-sequence generated by $y = x^{62}$ is the m-sequence generated by *x*, but in reverse order.

To find the generating polynomial:
Multiplying both sides of $x^6 = x + 1$ by x^{-6} yields $1 = x^{-5} + x^{-6}$, which is equivalent to $x^{-6} = x^{-5} + 1$ and, via $y = x^{-1}$, to $y^6 = y^5 + 1$.

C. Generate the sequence that results for c=7. How does it relate to the original m-sequence? Is it an m-sequence? If so, find the corresponding polynomial that generates it.
This sequence takes every seventh element from the original sequence. Since the original sequence has 63 elements, and $63 = 7 \times 9$, the resulting sequence would close (repeat) after 9 elements – and cannot be an m-sequence.

Bonus. D. Generate the sequence that results for c=5. How does it relate to the original m-sequence? Is it an m-sequence? If so, find the corresponding polynomial that generates it.

This sequence takes every fifth element from the original sequence. Since the original sequence has 63 elements and 5 is not a divisor of 63, the new sequence will also close after 63 elements. So it too must have all 6-tuples (except for the all-0 6-tuple).

To see that it is an m-sequence, we need to find a polynomial satisfied by *y*, namely, a relationship $y^6 = \sum_{r=0}^{5} A_r y^r$. This is equivalent to $x^{30} = \sum_{r=0}^{5} A_r x^{5r}$. Using the above table, we find

$x^0 = 1$
$x^5 = x^5$
$x^{10} = x^5 + x^4$
$x^{15} = x^5 x^{10} = x^5(x^5 + x^4) = x^{10} + x^9 = (x^5 + x^4) + (x^4 + x^3) = x^5 + 2x^4 + x^3 = x^5 + x^3$
$x^{20} = (x^5 + x^4)^2 = x^{10} + 2x^9 + x^8 = x^{10} + x^8 = x^5 + x^4 + x^3 + x^2$

$x^{25} = x^{10}x^{15} = (x^5 + x^4)(x^5 + x^3) = x^{10} + x^9 + x^8 + x^7 = (x^5 + x^4) + (x^4 + x^3) + (x^3 + x^2) + (x^2 + x^1) = x^5 + x^1$
$x^{30} = (x^{15})^2 = (x^5 + x^3)^2 = x^{10} + x^6 = x^5 + x^4 + x + 1$

So we need to solve the equation
$x^5 + x^4 + x + 1 = A_5(x^5 + x^1) + A_4(x^5 + x^4 + x^3 + x^2) + A_3(x^5 + x^3) + A_2(x^5 + x^4) + A_1x^5 + A_0$.
Or, equating coefficients of x (beginning with x^5)

$$
\begin{aligned}
A_5 + A_4 + A_3 + A_2 + A_1 &= 1 \\
A_4 + A_2 &= 1 \\
A_4 + A_3 &= 0 \\
A_4 &= 0 \\
A_5 &= 1 \\
A_0 &= 1
\end{aligned}
$$

Back-substituting yields $A_3 = 0$, $A_2 = 1$, and $A_1 = 1$, resulting in $y^6 = y^5 + y^2 + y + 1$.

Bonus. E. Develop a strategy to characterize the values of c that lead to the above 4 kinds of behaviors.
We need to characterize the behavior of the map $\psi_c(u) = u^c$. If c is not relatively prime to 63 (i.e., c and 63 have a nontrivial factor in common), then there will be some power $k < 63k$ for which $ck = 63n$. Since $y^k = x^{ck} = x^{63n} = (x^{63})^n = 1$, the sequence of powers of y will close after at most n steps, so we will not get an m-sequence, as in Part C. Conversely, if c is relatively prime to 63, then there is no such n, and we can use the strategy of Part D to show that $y = x^c$ will generate an m-sequence.

For c equal to a power of 2, the argument of part A applies – so $y = x^c$ generates the same m-sequence as x.

For c equal to 63 minus a power of 2, the argument of part B applies – so $y = x^c$ generates the m-sequence for x but in reverse order.

For all other values of c that are relatively prime to 63, the behavior of part D applies.

All in all, there are 36 integers from 1 to 63 that are relatively prime to 63. 6 of them are powers of 2 (part A), 6 are 63- a power of 2 (part B), and the other 24, together, generate four other m-sequences (part D). These four m-sequences are two entirely distinct sequences, and the same two sequences in reverse order. These observations follow from the fact that the maps $\psi_c(u) = u^c$ with c relatively prime to 63 form a group under composition, and have a subgroup consisting of those values c that are powers of 2.

The above analysis generalizes to m-sequences of any order (here, 6).

For further information, have a look at the “Euler phi-function” and the algebra of Galois fields.

Q4. How does calculation of Wiener kernels behave when datasets are pooled?
Setup: You have two measurements of the response, $R^{[1]}(t)$ *and* $R^{[2]}(t)$*, to the same stimulus* $s(t)$*. Consider estimation of the second-order Wiener kernel* $K_2(\tau_1,\tau_2)$ *from (a)* $R^{[1]}(t)$ *alone, (b)* $R^{[2]}(t)$ *alone, and (c) the average,* $\frac{1}{2}\left(R^{[1]}(t)+R^{[2]}(t)\right)$ *via cross-correlation (e.g., NLST notes page 32). Call these estimates* $K_2^{[1]}(\tau_1,\tau_2)$*,* $K_2^{[2]}(\tau_1,\tau_2)$*, and* $K_2^{average}(\tau_1,\tau_2)$*.*
With this setup,
(A) is there any relationship between $K_2^{[1]}(\tau_1,\tau_2)$*,* $K_2^{[2]}(\tau_1,\tau_2)$*, and* $K_2^{average}(\tau_1,\tau_2)$*?*

(B) Same as (A), but for cross-correlation estimates of the general (nth-order) kernel $K_n(\tau_1,\tau_2,...\tau_n)$*.*

(C) What, if anything, can one learn from $K_2^{[1]}(\tau_1,\tau_2)$ *and* $K_2^{[2]}(\tau_1,\tau_2)$ *that one cannot learn from* $K_2^{average}(\tau_1,\tau_2)$*?*

A. The cross-correlation procedure is linear in the data, i.e.,
$$K_2(\tau_1,\tau_2)=\frac{1}{2P^2}\left\langle R(t)\left(S(t-\tau_1)S(t-\tau_2)-P\delta(\tau_1-\tau_2)\right)\right\rangle,$$
so $K_2^{average}(\tau_1,\tau_2)=\frac{1}{2}\left(K_2^{[1]}(\tau_1,\tau_2)+K_2^{[2]}(\tau_1,\tau_2)\right)$.
B. Same as (A), since the cross-correlation procedure is always linear in the data.
C. One can get a crude estimate of the accuracy of the combined estimate by seeing how its two components agree.

Q5. Relationship between regression and principal components analysis.
Consider the following two scenarios (detailed below) involving analysis of multiple datasets. In scenario 1, you first regress, and then apply principal components analysis. In scenario 2, you first apply principal components analysis, and then regress. Under what circumstances are the results identical? (Justify your answer)

You are given a matrix Y containing multiple datasets, and the mth element of the rth dataset is the matrix element y_{mr} *. You are also given a matrix X of n regressors (the mth element of the nth regressor is* x_{mn} *).*

Scenario 1: You carry out one regression for each dataset, finding a matrix A for which a_{nr} *is the contribution of the nth regressor to the rth dataset, by minimizing* $R = \sum_{m,r} (\sum_{n} x_{mn}a_{nr} - y_{mr})^2 = tr\left((XA-Y)^T(XA-Y)\right)$. *This yields* $Y^{fit} = XA$. *You then calculate the first h principal components of* Y^{fit}*, and project* Y^{fit} *onto these principal components, leading to* $Y^{fit,PCA}$.

Scenario 2. You calculate the first h principal components of Y, and project Y onto these principal components, leading to Y^{PCA}*. You then regress* Y^{PCA} *on the regressors X, yielding regression coefficients* A^{PCA} *and* $Y^{PCA,fit} = XA^{PCA}$.

Discuss the relationship, if any, between $Y^{fit,PCA}$ *and* $Y^{PCA,fit}$.

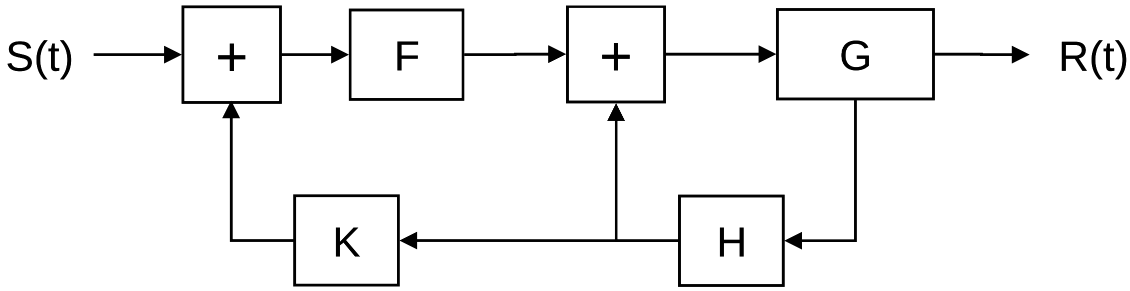
These quantities need not be related to each other. Regression is linear (projection of the data onto a specific subspace), while extraction of principal components is not (it depends on the covariance of the data, i.e., its second moments). There is no guarantee that a nonlinear procedure (here, PCA), will commute with a linear procedure (here, regression).

As a concrete example – consider *h*=1 (extraction of one principal component). Let's take a regression matrix *X* that has only one column, but which happens to be the same as the *second* principal component of Y. Then Y^{fit} will be scalar multiples of *X*, not typically 0. So Y^{fit} has only one principal component, namely *X*, and $Y^{fit,PCA}$ is the same as Y^{fit}. However, with *h*=1, Y^{PCA} (which consists of just its first principal component) is orthogonal to its second principal component (since PC's are orthogonal), and hence, projection of Y^{PCA} onto X, i.e., $Y^{PCA,fit}$, yields 0.

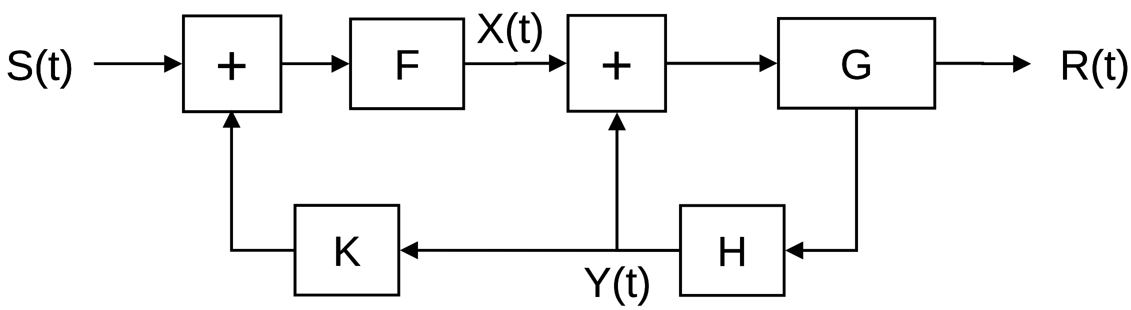
$Y^{fit,PCA}$ is guaranteed to lie in the space spanned by the columns of *X*, but $Y^{PCA,fit}$ is not.

Q6. Composition of linear filters.

Consider the following linear system, with input S(t), output R(t), and linear filters F, G, H, K with transfer functions $F(\omega)$, $G(\omega)$, $H(\omega)$, *and* $K(\omega)$. *Find the transfer function of the combined system.*



Label as follows.



Since *F* acts on the sum of *S* and the output of *K* to produce *X*,

$$\tilde{X}(\omega) = \tilde{F}(\omega)(\tilde{S}(\omega) + \tilde{K}(\omega)\tilde{Y}(\omega)). \qquad (5)$$

Since *G* acts on the sum of *X* and the output of *H* to produce *R*,

$$\tilde{R}(\omega) = \tilde{G}(\omega)(\tilde{X}(\omega) + \tilde{Y}(\omega)). \qquad (6)$$

Since *H* acts on *R* (the output of *G*) to produce *Y*,

$$\tilde{Y}(\omega) = \tilde{H}(\omega)\tilde{R}(\omega).$$

Putting the last two together yields

$$\tilde{Y} = \tilde{H}\tilde{R} = \tilde{H}\tilde{G}(\tilde{X} + \tilde{Y}), \text{ i.e., } \tilde{Y} = \frac{\tilde{H}\tilde{G}}{1 - \tilde{H}\tilde{G}}\tilde{X}, \text{ so } \tilde{X} + \tilde{Y} = \frac{1}{1 - \tilde{H}\tilde{G}}\tilde{X}. \qquad (7)$$

Putting this in eq. (5) yields

$$\tilde{X} = \tilde{F}(\tilde{S} + \tilde{K}\tilde{Y}) = \tilde{F}\tilde{S} + \tilde{F}\tilde{K}\tilde{Y} = \tilde{F}\tilde{S} + \tilde{F}\tilde{K}\frac{\tilde{H}\tilde{G}}{1 - \tilde{H}\tilde{G}}\tilde{X}.$$

That is,

$$\tilde{X} = \frac{\tilde{F}}{\left(1 - \tilde{F}\tilde{K}\dfrac{\tilde{H}\tilde{G}}{1 - \tilde{H}\tilde{G}}\right)}\tilde{S} = \frac{\tilde{F}(1 - \tilde{G}\tilde{H})}{1 - \tilde{G}\tilde{H}(1 + \tilde{F}\tilde{K})}\tilde{S}.$$

Now using (6) and (7),

$$\tilde{R} = \tilde{G}(\tilde{X} + \tilde{Y}) = \frac{\tilde{G}}{1 - \tilde{G}\tilde{H}}\tilde{X} = \frac{\tilde{F}\tilde{G}}{1 - \tilde{G}\tilde{H}(1 + \tilde{F}\tilde{K})}\tilde{S}.$$