Signals & Systems (3F1)

– Examples Paper 3F1/2 Random Processes –

January 5, 2016

Standard notation

$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
a^*	Complex conjugate of $a \in \mathbb{C}$
$\mathcal{F}[\cdot]$	Fourier transform
$\mathbb{E}[\cdot]$	Expectation operator
$\circledast$	Convolution operation

Symbol legend

☞	Important fact
✍	Computations needed
🖹	Use data book
⚠	Pay attention
≷	Clarification

Question 10. Show the following results for a wide-sense stationary (WSS), real-valued random process $\{X(t)\}$ with autocorrelation function $r_{XX}(\tau)$ and power spectrum $\mathcal{S}_X(\omega)$:

(a) $r_{XX}(-\tau) = r_{XX}(+\tau)$;

(b) If $\{X(t)\}$ represents a random voltage across a $1\,\Omega$ resistance, the average power dissipated is $P_{av} = r_{XX}(0)$;

(c) $\mathcal{S}_X(\omega) = \mathcal{S}_X(-\omega)$;

(d) $\mathcal{S}_X(\omega)$ is real-valued;

(e) $P_{av} = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{+\infty} \mathcal{S}_X(\omega)\mathrm{d}\omega$.

(a) By definition of autocorrelation function we have

$$r_{XX}(\tau) = \mathbb{E}[X(t)X(t+\tau)] = \mathbb{E}[X(t+\tau)X(t)], \quad \forall t \in \mathbb{R}.$$

Now if we substitute in the previous expression $t' := t + \tau$, we get

$$r_{XX}(\tau) = \mathbb{E}[X(t')X(t'-\tau)] = r_{XX}(-\tau).$$

(b) Let $V = X^2(t)$ denote the voltage across the resistance $R = 1\,\Omega$, then, by a well-known formula, the instantaneous power is given by

$$P(t) = \frac{V^2}{R} = \frac{X^2(t)}{1} = X^2(t) \quad \text{[Watt]},$$

and the average power has the form

$$P_{av} = \mathbb{E}[P(t)] = \mathbb{E}[X^2(t)] = r_{XX}(0) \quad \text{[Watt]},$$

where the last equality follows from the definition of autocorrelation function and the fact that $X(t)$ is WSS.

(c) By definition of Power Spectral Density,

$$\mathcal{S}_X(\omega) = \mathcal{F}\{r_{XX}(\tau)\} = \int_{-\infty}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau = \lim_{T\to\infty}\int_{-T}^{T} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau.$$

Now, we apply the change of variable $\tau \to -\tau'$, so that we obtain

⚠ Any change of variable affects also the limits of integration! In this case you have to swap the sign of the limits. This can be seen more clearly if you use the definition of improper integral.

$$\begin{aligned}\mathcal{S}_X(\omega) &= \lim_{T\to\infty}\int_{-T}^{T} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &= \lim_{T\to\infty}\int_{T}^{-T} r_{XX}(-\tau')e^{j\omega\tau'}(-\mathrm{d}\tau') \\ &= \lim_{T\to\infty}(-1)\cdot\int_{-T}^{T} -r_{XX}(-\tau')e^{j\omega\tau'}\mathrm{d}\tau' \\ &= \lim_{T\to\infty}\int_{-T}^{T} r_{XX}(-\tau')e^{j\omega\tau'}\mathrm{d}\tau'.\end{aligned}$$

Finally, by exploiting point (a), we have $r_{XX}(-\tau') = r_{XX}(\tau')$ and therefore

$$\mathcal{S}_X(\omega) = \lim_{T\to\infty}\int_{-T}^{T} r_{XX}(\tau')e^{j\omega\tau'}\mathrm{d}\tau' = \int_{-\infty}^{\infty} r_{XX}(\tau')e^{j\omega\tau'}\mathrm{d}\tau' = \mathcal{S}_X(-\omega).$$

(d) By splitting the integral in the definition of Power Spectral Density we obtain

$$\begin{aligned}\mathcal{S}_X(\omega) &= \int_{-\infty}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &= \int_{-\infty}^{0} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau.\end{aligned}$$

Now we use the change of variable $\tau \to -\tau'$ in the first integral of the latter expression and the property in point (a) to get

$$\begin{aligned}\mathcal{S}_X(\omega) &= \int_{-\infty}^{0} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &\overset{\tau'\to-\tau}{=} \int_{\infty}^{0} -r_{XX}(-\tau')e^{j\omega\tau'}\mathrm{d}\tau' + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &= (-1)\cdot\int_{0}^{\infty} -r_{XX}(-\tau')e^{j\omega\tau'}\mathrm{d}\tau' + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &\overset{(a)}{=} \int_{0}^{\infty} r_{XX}(\tau')e^{j\omega\tau'}\mathrm{d}\tau' + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau.\end{aligned}$$

By renaming the variable $\tau' = \tau$ and by using the formula $2\cos(\omega\tau) = e^{j\omega\tau} + e^{-j\omega\tau}$, we arrive at

$$\begin{aligned}\mathcal{S}_X(\omega) &= \int_{0}^{\infty} r_{XX}(\tau)e^{j\omega\tau}\mathrm{d}\tau + \int_{0}^{\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &= \int_{0}^{\infty} r_{XX}(\tau)(e^{j\omega\tau} + e^{-j\omega\tau})\mathrm{d}\tau \\ &= \int_{0}^{\infty} r_{XX}(\tau)2\cos(\omega\tau)\mathrm{d}\tau.\end{aligned}$$

The latter expression must be real since all terms in the integral are real.

(e) From point (b), we have

$$\begin{aligned}P_{av} &= r_{XX}(0) = \mathcal{F}^{-1}\{\mathcal{S}_X(\omega)\}\big|_{\tau=0} \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{S}_X(\omega)e^{j\omega\cdot 0}\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{S}_X(\omega)\mathrm{d}\omega \quad \text{[Watt]}.\end{aligned}$$

♢

Question 11. A white noise process $\{X(t)\}$ is a wide sense stationary, zero mean process with autocorrelation function:

$$r_{XX}(\tau) = \sigma^2\delta(\tau)$$

where $\delta(\tau)$ is the delta-function centred on $\delta = 0$ whose area is unity and width is zero. Sketch the Power Spectrum for this process.

A sample function $X(t)$ from such a white noise process is applied as the input to a linear system whose impulse response is $h(t)$. The output is $Y(t)$.

Derive expressions for the output autocorrelation function $r_{YY}(\tau) = \mathbb{E}[Y(t)Y(t+\tau)]$ and the cross-correlation function between the input and output $r_{XY}(\tau) = \mathbb{E}[X(t)Y(t+$

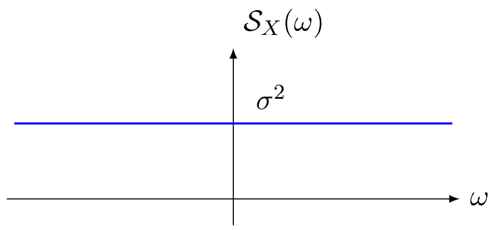
$\tau)]$. Hence obtain an expression for the output power spectrum $\mathcal{S}_Y(\omega)$ in terms of σ^2 and the frequency response $\mathcal{H}(\omega)$ of the linear system.

If the process is correlation ergodic, suggest in block diagram form a scheme for measurement of $r_{XY}(\tau)$. What is a possible application for such a scheme?

The power spectrum of the process $\{X(t)\}$ is given by

$$\begin{aligned}\mathcal{S}_X(\omega) &= \int_{-\infty}^{+\infty} r_{XX}(\tau)e^{-j\omega\tau}\mathrm{d}\tau \\ &= \int_{-\infty}^{+\infty} \sigma^2\delta(\tau)e^{-j\omega\tau}\mathrm{d}\tau = \sigma^2 e^{j\omega 0} = \sigma^2.\end{aligned}$$

Thus the power spectrum is a constant of value σ^2 over all frequencies and its plot is depicted below.



The output of a linear system with impulse response $h(t)$ is given by the convolution of the impulse response and the input, namely

$$Y(t) = h(t) \circledast X(t) = \int_{-\infty}^{+\infty} h(\beta)X(t-\beta)\mathrm{d}\beta.$$

As a consequence, the autocorrelation function of the output takes the form

☞ Recall that for a wide sense stationary process, as $\{X(t)\}$, the autocorrelation function depends only on the difference of the time instants.

$$\begin{aligned}r_{YY}(\tau) &= \mathbb{E}[Y(t)Y(t+\tau)] \\ &= \mathbb{E}\left[\left(\int_{-\infty}^{+\infty} h(\beta_1)X(t-\beta_1)\mathrm{d}\beta_1\right)\left(\int_{-\infty}^{+\infty} h(\beta_2)X(t+\tau-\beta_2)\mathrm{d}\beta_2\right)\right] \\ &= \mathbb{E}\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(\beta_1)h(\beta_2)X(t-\beta_1)X(t+\tau-\beta_2)\mathrm{d}\beta_1\mathrm{d}\beta_2\right] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(\beta_1)h(\beta_2)\mathbb{E}\left[X(t-\beta_1)X(t+\tau-\beta_2)\right]\mathrm{d}\beta_1\mathrm{d}\beta_2 \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(\beta_1)h(\beta_2)r_{XX}(\tau+\beta_1-\beta_2)\mathrm{d}\beta_1\mathrm{d}\beta_2 \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(\beta_1)h(\beta_2)\sigma^2\delta(\tau+\beta_1-\beta_2)\mathrm{d}\beta_1\mathrm{d}\beta_2 \\ &= \sigma^2\int_{-\infty}^{+\infty} h(\beta_2-\tau)h(\beta_2)\mathrm{d}\beta_2 \\ &= \sigma^2 h(-\tau) \circledast h(\tau). \qquad (1)\end{aligned}$$

Similarly, the cross-correlation function between the input and the output takes the form

$$\begin{aligned}
r_{XY}(\tau) &= \mathbb{E}[X(t)Y(t+\tau)] \\
&= \mathbb{E}\left[X(t)\left(\int_{-\infty}^{+\infty} h(\beta)X(t+\tau-\beta)\mathrm{d}\beta\right)\right] \\
&= \mathbb{E}\left[\int_{-\infty}^{+\infty} h(\beta)X(t)X(t+\tau-\beta)\mathrm{d}\beta\right] \\
&= \int_{-\infty}^{+\infty} h(\beta)\mathbb{E}\left[X(t)X(t+\tau-\beta)\right]\mathrm{d}\beta \\
&= \int_{-\infty}^{+\infty} h(\beta)r_{XX}(\tau-\beta)\mathrm{d}\beta \\
&= \int_{-\infty}^{+\infty} h(\beta)\sigma^2\delta(\tau-\beta)\mathrm{d}\beta \\
&= \sigma^2\int_{-\infty}^{+\infty} h(\beta)\delta(\tau-\beta)\mathrm{d}\beta_2 \\
&= \sigma^2 h(\tau).
\end{aligned}$$

Finally, by taking the Fourier transform of (1), we get the power spectrum of $\{Y(t)\}$

☞ Recall that the Fourier transform of a convolution product is the (standard) product of the Fourier transforms.

$$\mathcal{F}\{r_{YY}(\tau)\} = \mathcal{S}_Y(\omega) = \sigma^2\mathcal{H}^*(\omega)\mathcal{H}(\omega) = \sigma^2|\mathcal{H}(\omega)|^2,$$

since $\mathcal{H}^*(\omega)$ is the Fourier transform of $h(-\tau)$.

Now, if the process is correlation ergodic, we can replace the ensemble-average by the time-average in the computation of correlations. Hence, by considering the cross-correlation function between $\{X(t)\}$ and $\{Y(t)\}$, we get

$$r_{XY}(\tau) = \mathbb{E}[X(t)Y(t+\tau)] = \mathbb{E}[X(t-\tau)Y(t)] = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} X(t-\tau)Y(t)\mathrm{d}t,$$

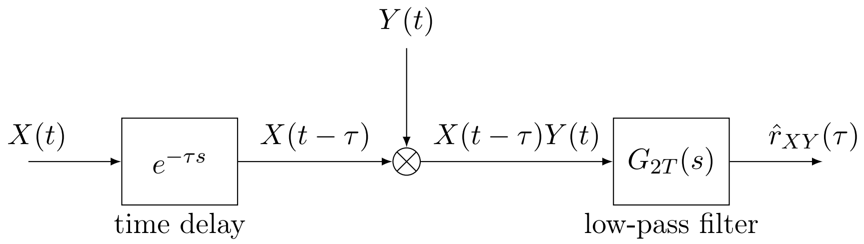
and for some suitably large choice of T, we can approximate the latter quantity by

$$r_{XY}(\tau) \approx \frac{1}{2T}\int_{-T}^{T} X(t-\tau)Y(t)\mathrm{d}t.$$

A scheme for the estimation of $r_{XY}(\tau)$ is given by the following steps:

1. pass $X(t)$ through a delay of τ, in order to obtain $X(t-\tau)$,
2. multiply $X(t-\tau)$ by $Y(t)$,
3. filter $X(t-\tau)Y(t)$ using a low-pass filter $G_{2T}(s)$ with impulse response whose main lobe is of duration roughly $2T$ between its half-amplitude points (this accounts for the integration $\frac{1}{2T}\int_{-T}^{T} X(t-\tau)Y(t)\mathrm{d}t$).

Eventually, by varying the delay τ, we can obtain an estimate $\hat{r}_{XY}(\tau)$ of the cross-correlation function. The figure below shows the block diagram representation of the previously described procedure.



Using the estimated cross-correlation function $\hat{r}_{XY}(\tau)$, we can estimate the impulse response of the system as follows

$$\hat{h}(\tau) = \frac{\hat{r}_{XY}(\tau)}{\sigma^2}.$$

Therefore the proposed scheme can be regarded as a "System Identification" technique. $\diamondsuit$

Question 12. It is desired to predict a future value $X(t+T)$ of a WSS random process $\{X(t)\}$ from the current value $X(t)$ using the formula:

$$\hat{X}(t+T) = cX(t)$$

where c is a constant to be determined.

If the process has autocorrelation function $r_{XX}(\tau)$, show that the value of c which leads to minimum mean squared error between $X(t+T)$ and $\hat{X}(t+T)$ is

$$c = \frac{r_{XX}(T)}{r_{XX}(0)}$$

Hence obtain an expression for the expected mean squared error in this case.

If $\{X(t)\}$ has non-zero mean, suggest an improved formula for $\hat{X}(t+T)$, giving reasons.

Consider the prediction error at time t

$$e(t) := X(t+T) - \hat{X}(t+T) = X(t+T) - cX(t).$$

By taking the expectation of $e^2(t)$ we obtain the mean squared prediction error

$$\begin{aligned} \text{MSE} := \mathbb{E}[e^2(t)] &= \mathbb{E}[(X(t+T) - cX(t))^2] \\ &= \mathbb{E}[X^2(t+T) - 2cX(t+T)X(t) + c^2X^2(t)] \\ &= (1+c^2)\mathbb{E}[X^2(t)] - 2c\mathbb{E}[X(t+T)X(t)] \\ &= (1+c^2)r_{XX}(0) - 2cr_{XX}(T), \end{aligned} \tag{2}$$

where we used the fact that the process $\{X(t)\}$ is WSS which implies that $\mathbb{E}[X^2(t+T)] = \mathbb{E}[X^2(t)] = r_{XX}(0)$. The mean squared prediction error is a convex quadratic function of c, therefore it admits a minimum $c_{\min}$ which can be found by equating to zero the derivative of MSE with respect to c, namely

$$\frac{\text{dMSE}}{\text{d}c} = 2cr_{XX}(0) - 2r_{XX}(T) \overset{!}{=} 0 \quad \Rightarrow \quad c_{\min} = \frac{r_{XX}(T)}{r_{XX}(0)}.$$

The minimum mean squared error is given by

$$\begin{aligned}
\mathrm{MSE}_{\min} &= (1 + c_{\min}^2) r_{XX}(0) - 2c_{\min} r_{XX}(T) \\
&= \left(1 + \frac{r_{XX}^2(T)}{r_{XX}^2(0)}\right) r_{XX}(0) - 2\frac{r_{XX}^2(T)}{r_{XX}(0)} r_{XX}(T) \\
&= \frac{r_{XX}^2(0) + r_{XX}^2(T) - 2r_{XX}^2(T)}{r_{XX}(0)} \\
&= \frac{r_{XX}^2(0) - r_{XX}^2(T)}{r_{XX}(0)}.
\end{aligned}$$

Now assume that $\{X(t)\}$ has non-zero mean, specifically $\mathbb{E}[X(t)] = \mu$ and define the zero-mean process $\tilde{X}(t) := X(t) - \mu$. The autocorrelation of $\{X(t)\}$ can be written as

$$\begin{aligned}
r_{XX}(\tau) &= \mathbb{E}[X(t+\tau)X(t)] \\
&= \mathbb{E}[(\tilde{X}(t+\tau) + \mu)(\tilde{X}(t) + \mu)] \\
&= r_{\tilde{X}\tilde{X}}(\tau) + \mu^2.
\end{aligned}$$

By applying formula (2) as it is, we get the following mean squared prediction error

$$\begin{aligned}
\mathrm{MSE}_{\min} &= \frac{r_{XX}^2(0) - r_{XX}^2(T)}{r_{XX}(0)} \\
&= \frac{(r_{XX}(0) + r_{XX}(T))(r_{XX}(0) - r_{XX}(T))}{r_{XX}(0)} \\
&= \left(\frac{r_{\tilde{X}\tilde{X}}(0) + \mu^2 + r_{\tilde{X}\tilde{X}}(T) + \mu^2}{r_{\tilde{X}\tilde{X}}(0) + \mu^2}\right)(r_{\tilde{X}\tilde{X}}(0) - r_{\tilde{X}\tilde{X}}(T)) \\
&= \left(1 + \frac{r_{\tilde{X}\tilde{X}}(T) + \mu^2}{r_{\tilde{X}\tilde{X}}(0) + \mu^2}\right)(r_{\tilde{X}\tilde{X}}(0) - r_{\tilde{X}\tilde{X}}(T)).
\end{aligned}$$

Now, since $r_{\tilde{X}\tilde{X}}(0) \geq r_{\tilde{X}\tilde{X}}(T)$ (see the clarification box in the next page), it follows that

1. $\frac{r_{\tilde{X}\tilde{X}}(T)+\mu^2}{r_{\tilde{X}\tilde{X}}(0)+\mu^2} \geq \frac{r_{\tilde{X}\tilde{X}}(T)}{r_{\tilde{X}\tilde{X}}(0)}$,
2. $r_{\tilde{X}\tilde{X}}(0) - r_{\tilde{X}\tilde{X}}(T) \geq 0$,

we conclude that,

$$\mathrm{MSE}_{\min} = \left(1 + \frac{r_{\tilde{X}\tilde{X}}(T) + \mu^2}{r_{\tilde{X}\tilde{X}}(0) + \mu^2}\right)(r_{\tilde{X}\tilde{X}}(0) - r_{\tilde{X}\tilde{X}}(T)) > \frac{r_{\tilde{X}\tilde{X}}^2(0) - r_{\tilde{X}\tilde{X}}^2(T)}{r_{\tilde{X}\tilde{X}}(0)},$$

the addition of a non-zero mean μ increases the mean squared error and, therefore, the smallest possible MSE is attained for a zero-mean process.

To improve the prediction performance, we can subtract the mean of the process and then apply the original predictor to the zero-mean remainder process $\{\tilde{X}(t)\}$,

$$\hat{\tilde{X}}(t+T) = c_{\min}\tilde{X}(t), \quad c_{\min} = \frac{r_{\tilde{X}\tilde{X}}(T)}{r_{\tilde{X}\tilde{X}}(0)} = \frac{r_{XX}(T) - \mu^2}{r_{XX}(0) - \mu^2}.$$

In this way, the "optimal" predictor of the original process $\{X(t)\}$ can be recovered as

$$\hat{X}(t+T) = \hat{\tilde{X}}(t+T) + \mu.$$

Here we prove that for a WSS process $\{X(t)\}$ and for all T

$$r_{XX}(0) \geq r_{XX}(T),$$

i.e. the autocorrelation function of a WSS process achieves its maximum value at 0. Given any two random variables X and Y, the Cauchy-Schwarz inequality reads as

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]},$$

or equivalently $\mathbb{E}^2[XY] \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$. Using this result, we have that for all $T \neq 0$,

$$r_{XX}^2(T) = \mathbb{E}^2[X(t)X(t+T)] \leq \mathbb{E}[X^2(t)]\mathbb{E}[X^2(t+T)] = r_{XX}^2(0),$$

where we used the fact that $\{X(t)\}$ is WSS.

$\diamondsuit$