2.160 IDENTIFICATION, ESTIMATION, AND LEARNING

LECTURE NOTES NO. 11

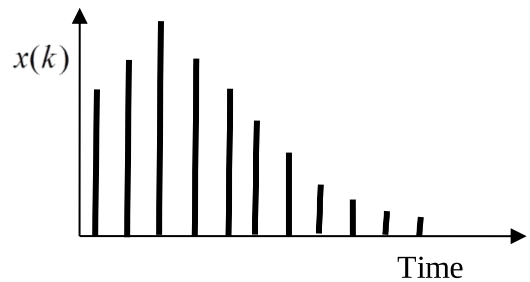
12 Frequency Domain Analysis

As an alternative to the time domain approach using auto- and cross-correlation functions, frequency-domain methods based on spectral analysis have been developed for obtaining frequency transfer functions of linear time-invariant systems.

12.1 Discrete-Time Fourier Transform and Power Spectrum

Discrete-Time Fourier transform of a sampled-data system: $x(k)$,

$$X(\omega) = \sum_{k=-\infty}^{+\infty} x(k) e^{-i\omega k} \qquad (1)$$



Note that $X(\omega)$ is a 2π-periodic function:

$$X(\omega + 2\pi n) = \sum_{k=-\infty}^{\infty} x(k) e^{-i(\omega+2\pi n)k} = \underbrace{\sum_{k=-\infty}^{\infty} x(k) e^{-i\omega k}}_{X(\omega)} \underbrace{e^{-i\cdot 2\pi nk}}_{1} = X(\omega) \qquad (2)$$

$$\therefore X(\omega + 2\pi n) = X(\omega) \qquad (3)$$

A periodic function can be expanded to a Fourier series expansion. Therefore, we can write

$$\int_{-\pi}^{\pi} X(\omega) e^{i\omega\ell} d\omega = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} x(k) e^{-i\omega k} e^{i\omega\ell} d\omega = \sum_{k=-\infty}^{\infty} x(k) \int_{-\pi}^{\pi} e^{i\omega(\ell-k)} d\omega = 2\pi x(\ell) \qquad (4)$$

Note that the last equation was derived from $\int_{-\pi}^{\pi} e^{i\omega(\ell-k)} d\omega = \begin{cases} 2\pi : k = \ell \\ 0 \quad : k \neq \ell \end{cases}$. Directly from (4) the inverse transform of $X(\omega)$ is obtained as

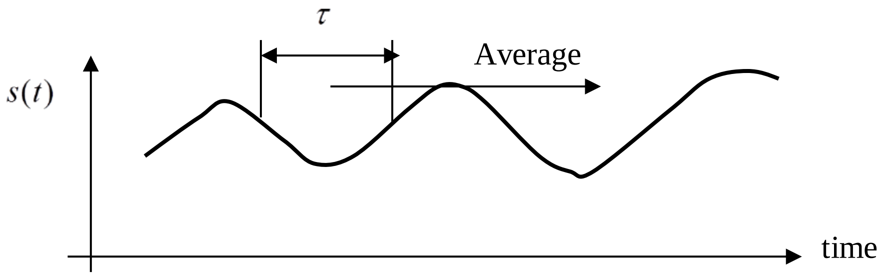
$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{i\omega n} d\omega \qquad (5)$$

Power Spectrum

Consider a wide-sense stationary sequence $\{s(t)\}$ for which the following limit exists:

$$R_s(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} s(t)s(t+\tau) : \qquad \text{Auto-correlation} \tag{6}$$

(Auto-covariance if zero mean)



The power spectrum of $\{s(t)\}$ is defined as the Fourier transform of auto-correlation function $R_s(\tau)$

$$\Phi_s(\omega) = \sum_{\tau=-\infty}^{\infty} R_s(\tau) e^{-i\tau\omega} \tag{7}$$

Inverse Transform:

$$R_s(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_s(\omega) e^{i\tau\omega} d\omega \tag{8}$$

A special case for $\tau = 0$ is:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} s^2(t) = R_s(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_s(\omega) d\omega \tag{9}$$

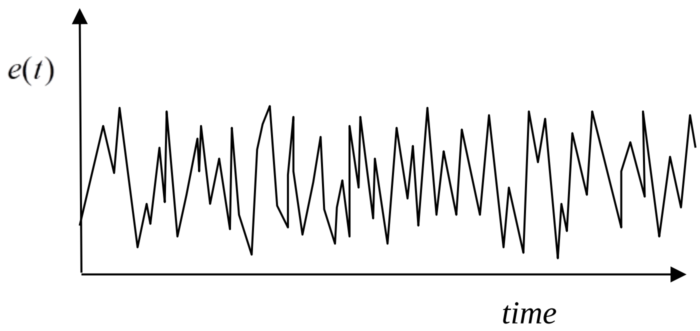
Important! We will often use this formula to obtain the mean of a squared signal $s(t)$, a special case of $\tau = 0$. The sum of the square of signal $s(t)$ represents a type of "energy" of the signal. The above expression represents that the signal energy in the time domain and that in the frequency domain are the same. This is referred to as Parseval's Theorem, or Rayleigh's Energy Theorem.

White Noise

We have seen "White Noise" in many sections of the previous lectures.
We defined $\{e(t)\}$ as a sequence of independent random variables with zero mean values and covariance λ :

$$E[e(t)e(s)] = \begin{cases} \lambda & t = s \\ 0 & t \neq s \end{cases} \tag{10}$$

Now its properties are characterized by using Power Spectrum.



$$E\left[e(t)e(s)\right] = \begin{cases} \lambda & t = s \\ 0 & t \neq s \end{cases}$$

The auto-covariance of the random process $e(t)$ is given by

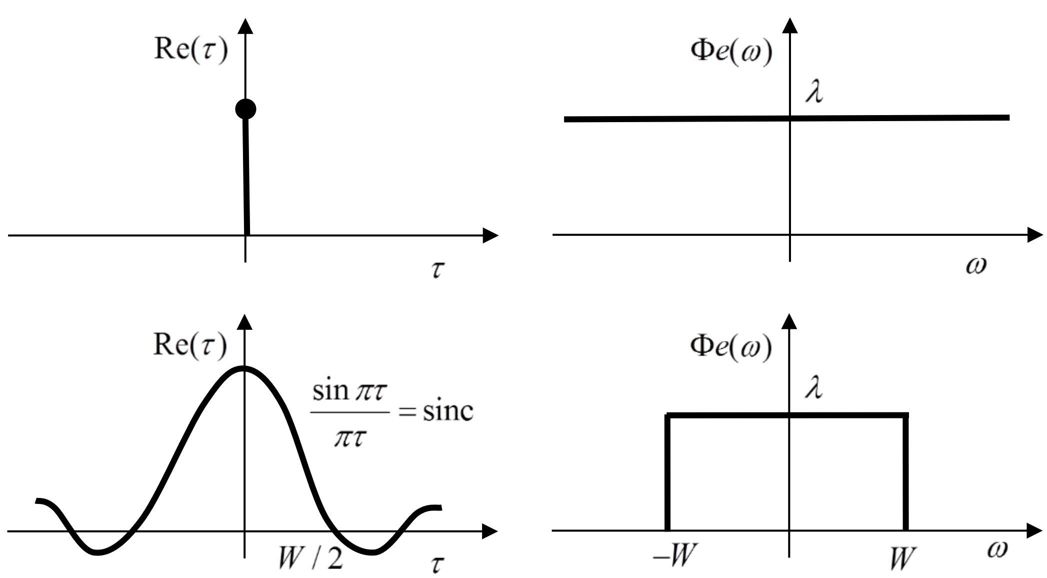
$$R_e(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} e(t)e(t+\tau) \Rightarrow E\left[e(t)e(t+\tau)\right] = \lambda\, \delta(\tau) \qquad (11)$$

The left hand side of the above expression is a time average, while the right hand side is an ensemble average. If these two averages are the same, the process is called ***ergodic.*** We assume this ergodicity for most of the processes. See the discussion at the end of this lecture notes.

For the above equation the Power spectrum is given by

$$\Phi_e(\omega) = \sum_{\tau=-\infty}^{\infty} R_e(\tau) e^{-i\tau\omega} = \sum_{\tau=-\infty}^{\infty} \lambda\, \delta(\tau) e^{-i\tau\omega} = \lambda \qquad (12)$$

The figures below show the plots of the auto-covariance $\mathrm{R_e}(\tau)$ against time and the corresponding power spectrum $\Phi_e(\omega)$ against frequency. Note that the power spectrum plot is constant for the entire frequency. In optics, this means that the light has a uniform distribution over the entire wave length, that is, "White". This is why the random process $e(t)$ is called "White Noise".



When the white noise is band-limited, the auto-covariance becomes a sinc function. See the figures.

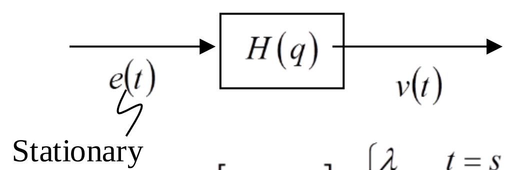
12.2 Applying spectral Analysis to System Identification

A colored random signal can be created with the White noise going through a dynamical process. The following theorem plays a major role in many of the analyses involved in system identification.

Theorem

Let $H(q)$ be the transfer function of a (BIBO) stable process with a White noise input $e(t)$ of variance λ.

$$v(t) = H(q)e(t) \qquad (13)$$



$$E[e(t)e(s)] = \begin{cases} \lambda & t = s \\ 0 & t \neq s \end{cases}$$

The power spectrum of $v(t)$ is given by

$$\Phi_v(\omega) = \lambda \left| H\left(e^{i\omega}\right) \right|^2 \qquad (14)$$

where $|\bullet|$ is the norm of a complex number, and $H(e^{i\omega})$ is obtained by replacing q in $H(q)$ by $e^{i\omega}$.

<u>Proof</u> $v(k) = \sum_{l=0}^{\infty} h(l)e(k-l)$

Similarly, $v(k-\tau) = \sum_{m=0}^{\infty} h(m)e(k-\tau-m)$

Combining these two yields

$$\begin{aligned} R_v(\tau) &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} v(k)v(k-\tau) \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \sum_{l=0}^{\infty} h(l)e(k-l) \sum_{m=0}^{\infty} h(m)e(k-\tau-m) \end{aligned} \qquad (15)$$

$$R_v(\tau) = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} h(l)h(m) \underbrace{\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} e(k-l)e(k-\tau-m)}_{\|}$$

$$\lambda\delta(l-\tau-m) = \begin{cases} \lambda : & l = \tau + m \\ 0 : & otherwise \end{cases}$$

$$\therefore \quad R_v(\tau) = \lambda \sum_{l=\max(0,\tau)}^{\infty} h(l)h(l-\tau) \quad : \quad h(l) = 0 \quad for \quad l < 0$$

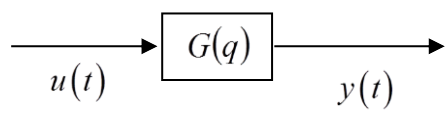
The power spectrum is then given by

$$\begin{aligned}
\Phi_v(\omega) &= \sum_{\tau=-\infty}^{\infty} R_v(\tau) e^{-i\omega\tau} \\
&= \lambda \sum_{\tau=-\infty}^{\infty} \sum_{l=\max(0,\tau)}^{\infty} h(l) e^{-i\omega\ell} h(l-\tau) e^{i\omega(l-\tau)} \\
&\text{Replacing} \quad l-\tau \quad by \quad s \qquad\qquad (16) \\
&= \lambda \sum_{l=0}^{\infty} h(l) e^{-il\omega} \sum_{s=0}^{\infty} h(s) e^{i\omega s} \\
&= \lambda H\left(e^{i\omega}\right) \cdot H\left(e^{-i\omega}\right) = \lambda \left| H\left(e^{i\omega}\right) \right|^2
\end{aligned}$$

Here, $H(e^{i\omega})$ is a complex number, and $H(e^{-i\omega})$ is its complex conjugate. The product of the two gives the squared norm of $H(e^{i\omega})$. Q.E.D.

Cross Spectrum

Consider a time-invariant linear system $G(q)$ and a wide-sense stationary signal $\{u(t)\}$ with spectrum $\Phi_u(\omega)$



$$y(t) = G(q)u(t)$$

The cross spectrum is defined as

$$\Phi_{yu}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{yu}(\tau) e^{-i\tau\omega} \qquad (17)$$

$$R_{yu}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u(t) y(t+\tau) \qquad (18)$$

Following a procedure similar to the above theorem, we can show that the cross spectrum is related to the input power spectrum as

$$\Phi_{yu}(\omega) = G\left(e^{i\omega}\right) \Phi_u(\omega) \qquad (19)$$

Note that the cross spectrum is a complex function, since $G(e^{i\omega})$ is complex. The following properties of power spectrum and cross spectrum can be proven without difficulty:

$$\Phi_u(\omega) = \Phi_u(-\omega) \geq 0 \tag{20}$$
$$\Phi_{uy}(\omega) = \Phi_{yu}(-\omega) \tag{21}$$

Eq.(19) suggests a way of obtaining a frequency transfer function.

$$G\left(e^{i\omega}\right) = \frac{\Phi_{yu}(\omega)}{\Phi_u(\omega)} \tag{22}$$

From an experimental data set $\{(y(t), u(t)) \mid t = 1, \cdots, N\}$, compute both input auto-correlation and cross correlation, and obtain their power spectra. Dividing the cross spectrum by the input power spectrum gives the frequency transfer function.

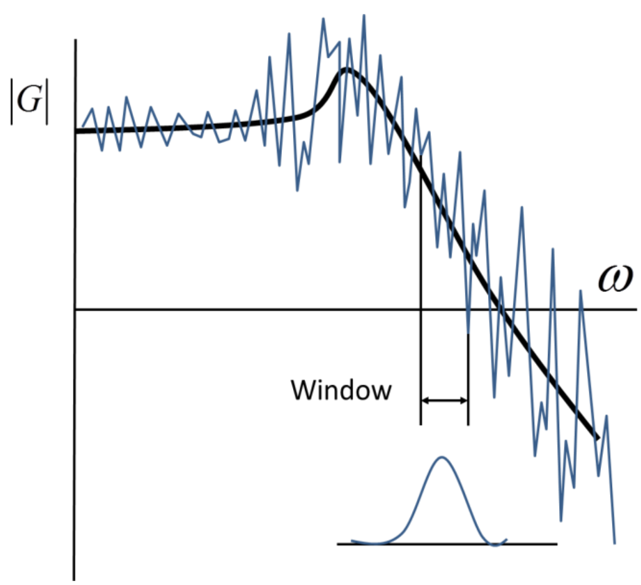
This method provides a consistent estimate of the transfer function, that is, the estimated transfer function $\hat{G}(e^{i\omega})$ approaches its true function $G_0(e^{i\omega})$ despite noise $v(t)$ as N tends to infinity. However, the estimation variance does not vanish although N tends to infinity. We can show that for a large N the variance is given by

$$\operatorname{var}\hat{G}(e^{i\omega}) = E\left[\,|\hat{G}(e^{i\omega}) - G_0(e^{i\omega})|^2\,\right] = \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \tag{23}$$

where $\Phi_v(\omega)$ is the power spectrum of noise $v(t)$.

Exercise Prove the above error variance in estimating a frequency transfer function.

It should be noted that the variance remains, no matter how many data points one can obtain for identifying the frequency transfer function. This often results in a jagged Bode plot, as illustrated in the figure below. To smoothen the curve, we can take a local average at each frequency, assuming that the true transfer function is smooth. Setting a window and a weight function in the vicinity of each frequency, we can take a weighted average of the frequency responses to obtain a smooth Body plot, as shown in the figure. There are various local averaging methods available, including the Hamming Window. For details see Chapter 6.4 of Ljung's textbook.



Coherence

Coherency is a measure often used to quantify the fidelity of a frequency transfer function obtained from experimental data.

$$\gamma^2(\omega) = \frac{\left|\Phi_{yu}(\omega)\right|^2}{\Phi_u(\omega)\Phi_y(\omega)}, \quad 0 \le \gamma^2 \le 1 \tag{24}$$

In an ideal case, the coherence takes 1, indicating the highest fidelity. This can be shown by considering a perfectly linear, noise-less system, where

$$G_0(e^{i\omega}) = \frac{\Phi_{yu}(\omega)}{\Phi_u(\omega)} \tag{25}$$

From the Theorem in the previous section we obtain

$$\Phi_y(\omega) = \left|G_0(e^{i\omega})\right|^2 \Phi_u(\omega) \tag{26}$$

Dividing (26) by $\Phi_y(\omega)$ and substituting (25) yields

$$1 = \frac{\left|\Phi_{yu}(\omega)\right|^2}{\Phi_u(\omega)\Phi_y(\omega)} = \gamma^2(\omega) \tag{27}$$

However, it becomes less than 1 due to various reasons, including:

- Other inputs and noise contributing to the output;
- Nonlinear distortion: the system has some nonlinearity; and
- Leakage errors of the Discrete Fourier Transfer used.

It is recommended to evaluate the coherence over the frequency range used. We can find in which frequency range the linear model given by the transfer function is valid.

<u>Notes on stationary processes</u>: We use

$$y(t) = \underbrace{G(q)u(t)}_{\text{Deterministic}} + \underbrace{H(q)e(t)}_{\text{Stochastic}} \tag{28}$$

One issue to clarify for mathematical rigor

Strictly speaking, the process is not stationary; input $u(t)$ drives the system. Therefore, the covariance function $R_s(\tau)$ cannot be defined in general.

$$R_s(\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} s(t)s(t-\tau)$$

The stationary property that we need is the existence of this covariance. Then, we extend the definition to the one assuming the existence of $R_s(\tau)$

Quasi-Stationary (wide sense stationary)

Important theories and techniques of system identification are dependent upon the spectra of the involved signals, i.e. only the second-order properties, and not on any higher-order properties.

Deterministic and stochastic processes are mixed. $\{e(t)\}, \{v(t)\}$ are stochastic processes. The covariance function for this type of variable must be given an ensemble mean:

$$R_s(\tau) = E\left[s(t)s(t-\tau)\right]$$

If this is equivalent to:

$$R_s(\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} s(t)s(t-\tau)$$

treatment will be very convenient, since we need to consider only one realization of the stochastic process, rather than considering the whole collection of ensemble average. This is an ergodicity problem. For dynamical systems, this ergodicity holds for signals generated through uniformly stable filters $G(q)$

$$s(t) = G(q)e(t)$$

Under this assumption, the theoretical boundary between stochastic and deterministic processes is low. More discussion will be given in the following chapter.