EE 123: Discrete-time Signal Processing	Spring 2007

Lecture 25 — April 12

Lecturer: Prof. Anant Sahai *Scribe: Evan Battaglia*

25.1 Outline

In previous lectures we used a model for random noise signals that assumed wide-sense stationary, ergodic, and white random signals. Today we explore a little more as to what exactly these mean.

- Wide-sense Stationarity
- Autocorrelation and Ergodicity
- Power Spectral Density (PSD) and White Signals
- PSD thru an LTI System
- Periodogram

25.2 Wide-sense Stationarity

The model we are using for random signals is that they are **wide-sense stationary (WSS)**. This term has two parts to its meaning:

- **Wide-sense** – we only care about the first and second moments, or, equivalently, we only care about the mean and the variance. The first moment is $E[x[n]]$, the expectation of x[n], or the average of x[n] across all universes. In other words, the first moment is the mean value across all universes. The second moment is of the same order of the variance. Variance is defined as:

$$Variance\,[x[n]] = \sigma_x^2 = E\left[|x[n] - m_{x[n]}|^2\right] \tag{25.1}$$

  Where $m_{x[n]}$ is $E[x[n]]$, the mean of $x[n]$. We can have two vastly different probability densities – for instance, a discrete probability density with Dirac deltas representing a coin toss, or a continuous flat probability density – and as long as they have the same mean and variance they are equivalent for our purposes under this model.

- **Stationary** – stationarity is a bit like time-invariance for random signals. A random signal is stationary if it doesn't matter where the clock starts, i.e. at which position in the signal we choose to be zero.

A wide-sense stationary signal is then a random signal for which the expectation is constant in time, i.e.

$$E\left[x[n]\right] = m_x \quad \forall n \tag{25.2}$$

and also the variance σ_x^2 is constant in time, i.e.

$$E\left[x[n]\right] = \sigma_x^2 \quad \forall n \tag{25.3}$$

In addition, as we detail further in the next section, the **autocorrelation** ϕ_{xx} of wide-sense stationary signals is dependent only on the time difference m:

$$\phi_{xx}[n+m, n] = \phi_{xx}[m] = E\left[x[n+m]x^*[n]\right] \tag{25.4}$$

where this equation holds for all n.
Also, for the signals we will be looking at, we will assume that the mean is zero, i.e.

$$E\left[x[n]\right] = 0 \tag{25.5}$$

25.3 Autocorrelation and ergodicity

Sometimes wide-sense stationarity is not a strong enough condition for random signals to be useful for our purposes. Consider a random signal represented by the universes-vs-time table in Table 25.1:

	Time									
Universes	+1	+1	+1	+1	+1	+1	+1	+1	+1	...
	-1	-1	-1	-1	-1	-1	-1	-1	-1	...
	+1	+1	+1	+1	+1	+1	+1	+1	+1	...
	+1	+1	+1	+1	+1	+1	+1	+1	+1	...
	-1	-1	-1	-1	-1	-1	-1	-1	-1	...
	-1	-1	-1	-1	-1	-1	-1	-1	-1	...
	...	...	...	...	...	...	...	...	...	

Table 25.1. Universes vs. time for an example random signal

Here, the only signals possible are all positive one or all negative one. The random signal clearly "doesn't care what time is". The mean is always 0 for any point in time, and it can be seen that the product of any two values in one signal is always one – i.e. $E\left[x[n+m]x^*[n]\right] = 1$ for all n and even for all m. Thus it is wide-sense stationary, but somehow "not nice" for our purposes and at times hard to work with. For instance, if we are just dealing with one universe we cannot tell anything about the probability in other universes. It is a bit like asking the question "What is the probability that China is communist?" Seeing that

it was communist today, yesterday, and the day before doesn't tell us anything about the probability that the Nationalists would have won in the civil war in another universe.

To investigate another restriction on random signals, let us first look at the autocorrelation sequence, defined as:

$$\phi_{xx}[m] = E\left[x[n+m]x^*[n]\right] \tag{25.6}$$

The autocorrelation sequence is only defined for WSS signals, so that it is not dependent on n. The autocorrelation sequence is essentially a sequence of universe-averages for a given time difference. To parallel this idea we can also define averaging over time:

$$< x_n >= \text{average of } x_n \text{ over through time} = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{k=-L}^{L} x_k \tag{25.7}$$

This definition is tied to a particular universe – i.e. a particular instance of the random signal. We are also interested in averaging the product of a value times the conjugate of the value m samples before it:

$$< x_{n+m}x_n^* >= \lim_{L\to\infty} \frac{1}{2L+1} \sum_{k=-L}^{L} x_{k+m}x_k^* \tag{25.8}$$

We now consider the restriction that these two time averages are equal to the corresponding universe averages (assuming a WSS random signal so m_x and $\phi_{xx}[m]$ are defined):

$$< x_n >= m_x \tag{25.9}$$

$$< x_{n+m}x_n^* >= \phi_{xx}[m] \tag{25.10}$$

If these equations hold in all universes (or technically, "almost everywhere"), we say that the signal is **ergodic**.

Note that the example in Table 25.1 is clearly not ergodic, since the first-order time-average is universe-dependent (either 1 or -1) yet the universe-average m_x is zero.

25.4 Power Spectral Density (PSD) and White Signals

We define the power spectral density (PSD) of a random signal as the DTFT of the autocorrelation sequence:

$$\Phi_{xx}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} \phi_{xx}[m]e^{-j\omega n} \tag{25.11}$$

What does the PSD represent? We can get a hint from the fact that the average of the PSD over frequency is $\phi_{xx}[0]$, the mean-square value, or the energy of the random signal:

$$\phi_{xx}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{xx}(e^{j\omega})e^{j\omega(0)}d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{xx}(e^{j\omega})d\omega \tag{25.12}$$

In fact, the PSD represents how much power a random signal has in different frequencies. Thus it makes sense that the total power is found by integrating over all frequencies.

A random signal for which the PSD is constant over all frequencies is called **white** – in an analogy to white light, it contains equal amounts of "colors" or frequencies. Thus, for a white signal:

$$\phi_{xx}[n] = c\delta[n] \tag{25.13}$$

$$\phi_{xx}[m \neq 0] = 0 \tag{25.14}$$

Thus for a white signal $x[m]$ and $x[m+n]$ are completely independent for $n \neq 0$.

25.5 PSD thru an LTI System

When a random, WSS, ergodic signal $x[n]$ is sent through an LTI system $H(e^{j\omega})$, the random signal $y[n]$ that is produced has a PSD of:

$$\Phi_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2\Phi_{xx}(e^{j\omega}) \tag{25.15}$$

Thus, when a white signal is passed through a non all-pass system, the result is then colored. For instance, if we pass a white signal x through a tightly-notched BPF centered close to ω_f, as illustrated in Figure 25.1, the resulting signal y has all its power around the selecting frequency of the BPF.

Furthermore, if we have a non-white random signal x passing through a BPF, as we tune the BPF – i.e. change the selecting frequency – the resulting signal y completely changes, not just in which frequencies the power lies in, but also the magnitude (since the PSD of x is not flat). This is illustrated in Figure 25.2.

As an aside, if we take the Z-transform perspective, the PSD $\Phi_{yy}(z)$ is determined by:

$$\Phi_{yy}(z) = H(z)H^*(\frac{1}{z^*})\Phi_{xx}(z) \tag{25.16}$$

25.6 Periodogram

We can obtain an estimate of the power spectral density by finding what we will call the periodogram. Let us first window the random signal with an appropriate windowing function $w[n]$.

$$v[n] = w[n]x[n] \tag{25.17}$$

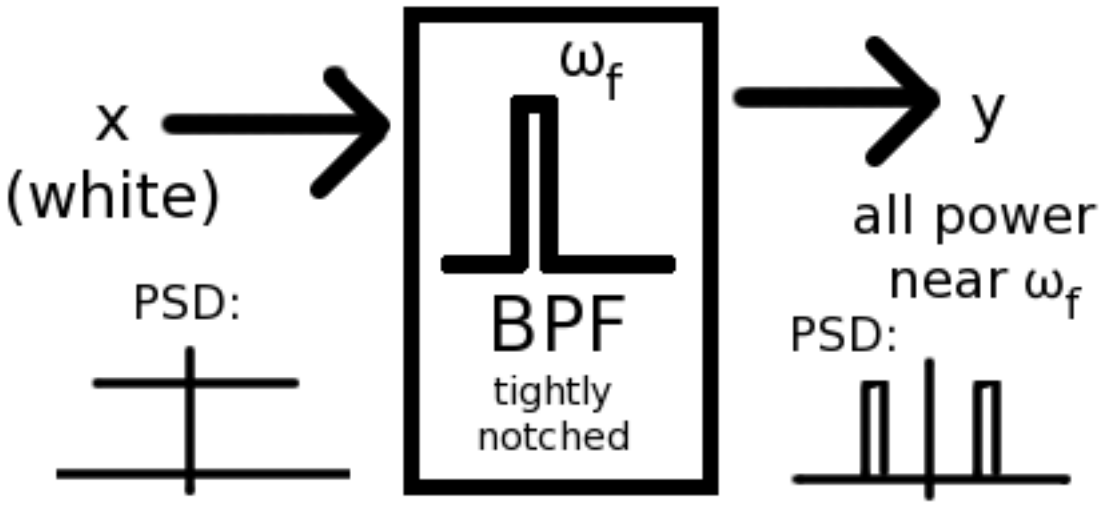

Figure 25.1. BPF coloring of white random signal

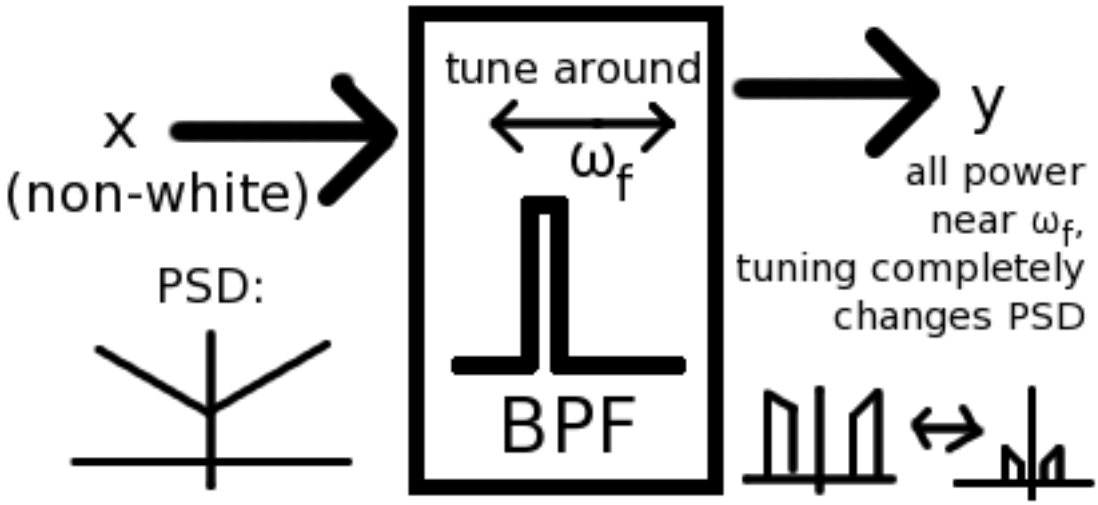

Figure 25.2. BPF effect on signal as notch is tuned

Let us for simplicity pick an L-point rectangular window. The DTFT of $v[n]$ is then:

$$V(e^{j\omega}) = \sum_{n=0}^{L-1} v[n]e^{-j\omega n} \tag{25.18}$$

We the define the **periodogram** as:

$$I(\omega) = \frac{1}{LU}\left|V(e^{j\omega})\right|^2 \tag{25.19}$$

Where U is a normalization constant (more on this later). It can be shown that the periodogram can also be found by the formula:

$$I(\omega) = \frac{1}{LU}\sum_{m=-(L-1)}^{L-1} c_{vv}[m]e^{-j\omega m} \tag{25.20}$$

where $c_{vv}[m]$ is defined by the following formula (assuming real signals):

$$c_{vv}[m] = \sum_{n=0}^{L-1} v[n]v[n+m] \tag{25.21}$$

Furthermore, it can be shown the the expectation of the periodogram is related to the PSD by the formula:

$$E\left[I(\omega)\right] = \frac{1}{2\pi LU}\int_{-\pi}^{\pi} \Phi_{xx}(e^{j\theta})C_{ww}(e^{j(\omega-\theta)})d\theta \tag{25.22}$$

where

$$C_{ww}(e^{j\omega}) = \left|W(e^{j\omega})\right|^2 \tag{25.23}$$

Thus, the expectation of the periodogram is a blurred version of the PSD, blurred by the square of the window function. Using Parseval's theorem we determine what the normalization factor U must be:

$$U = \frac{1}{L}\sum_{n=0}^{L-1}(w[n])^2 \tag{25.24}$$

As $L \to \infty$, $W(e^{j\omega})$ approaches a Dirac delta and the expectation of the periodogram $E[I(\omega)]$ approaches a constant times Φ_{xx}, the power spectral density.

However, we cannot just find the periodogram and use this as our estimate of the PSD, because the standard deviation of the periodogram is as large as the information conveyed. In other words:

$$Variance\left[I(\omega)\right] \simeq \Phi_{xx}^2(\omega) \tag{25.25}$$

Therefore if we want to obtain an estimate for the PSD $\Phi_{xx}(\omega)$ we must average the periodogram $I(\omega)$. How do we do this? As shown in Figure 25.3, we divide up time into K slices of length L, compute the periodogram for each of these, and average the results. As we take more slices the variance drops as $\frac{1}{K}$ and thus the standard deviation drops as $\frac{1}{\sqrt{K}}$. For example, we must find and average periodograms for 100 slices of length K to obtain a PSD estimate with a standard deviation (precision) of $\frac{1}{10}$.

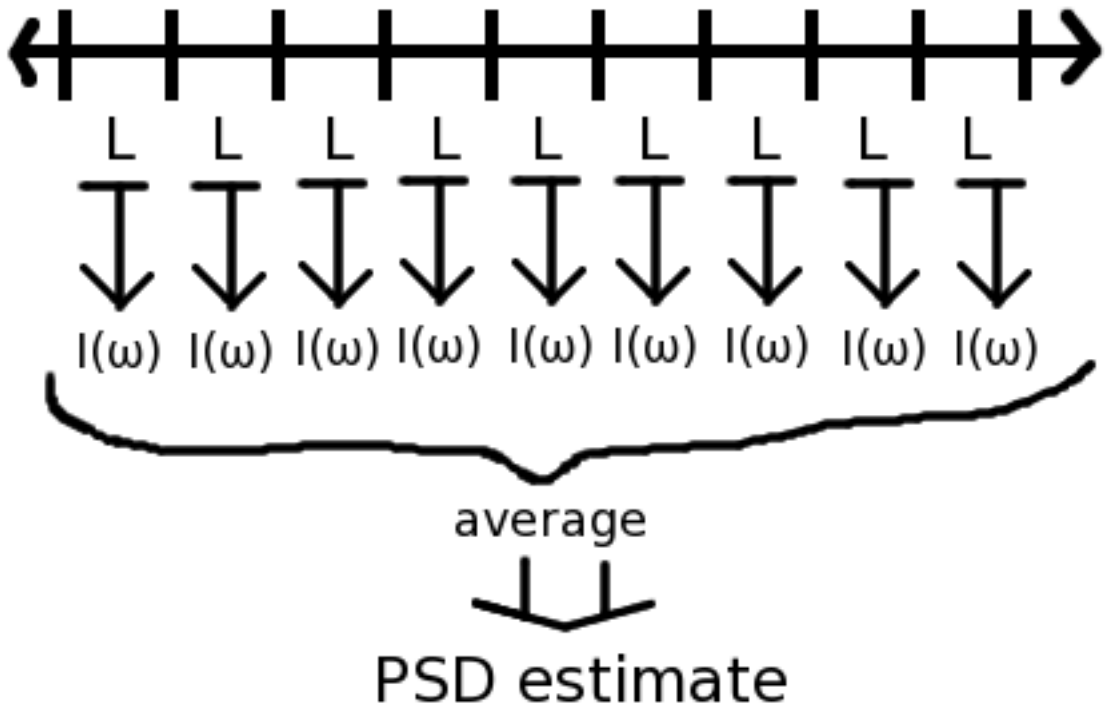

Figure 25.3. Finding the PSD by periodogram averaging