Lecture 8 **Fourier Transforms and Application to the PSD**

Definition 1. Let $x(t)\,;\,t\in(-\infty,\infty)$ be a deterministic function that satisfies $\int_{-\infty}^{\infty}|\,x(t)\,|\,dt<\infty$. Then its ***Fourier Transform*** is defined as $X(\omega)\overset{\Delta}{=}\int_{-\infty}^{\infty}x(t)\,e^{-i\omega t}dt\overset{\Delta}{=}\Im(x)(\omega)$.

Theorem 1. The inverse Fourier transform of $X(\omega)$ is $x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)\,e^{i\omega t}d\omega\overset{\Delta}{=}\Im^{-1}(X)(t)$.

Proof: Begin with $\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{i\omega t}d\omega=\frac{1}{2\pi}\int_{\omega=-\infty}^{\infty}\left[\int_{\tau=-\infty}^{\infty}x(\tau)e^{-i\omega\tau}d\tau\right]e^{i\omega t}d\omega=\frac{1}{2\pi}\int_{\tau=-\infty}^{\infty}x(\tau)\left[\int_{\omega-\infty}^{\infty}e^{-i\omega(\tau-t)}d\omega\right]d\tau$.
Let $s=\tau-t$. Then $ds=d\tau$ and $\tau=s+t$, so that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{i\omega t}d\omega=\frac{1}{2\pi}\int_{s=-\infty}^{\infty}x(s+t)\left[\frac{1}{2\pi}\int_{\omega-\infty}^{\infty}e^{-i\omega s}d\omega\right]ds$$

We now need to compute the inner integral. Rather than using a limit-based argument associated with the infinite limits of integration, we will use the following method:

The Fourier transform of $\delta(s)$ is $\int_{s-\infty}^{\infty}\delta(s)e^{-i\omega s}ds=1$. Hence, the inverse Fourier transform of 1 is

$\frac{1}{2\pi}\int_{s-\infty}^{\infty}1\bullet e^{-i\omega s}d\omega=\delta(s)$. But this gives $\int_{s-\infty}^{\infty}e^{-i\omega s}d\omega=2\pi\delta(s)$. Hence:

$\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{i\omega t}d\omega=\int_{s=-\infty}^{\infty}x(s+t)\delta(s)ds=x(t)$. This completes the proof. □

Remark 1. As a result of this proof, we have discovered the following Fourier transform pairs:

$$\delta(t)\quad\leftrightarrow\quad 1\qquad\text{and}\qquad 1\quad\leftrightarrow\quad 2\pi\delta(\omega).$$

Remark 2. The units of ω are typically radians/second, though radians per hour, day or year are also common. Let $f=\omega/2\pi$ have units of cycles/second (i.e. Hz). Then the Fourier transform pair becomes:

$$\breve{X}(f)=\int_{-\infty}^{\infty}x(t)e^{-i2\pi f t}dt\qquad\text{and}\qquad x(t)=\int_{-\infty}^{\infty}\breve{X}(f)\,e^{i2\pi f t}df\,.$$

The Fourier Transform of the Autocorrelation Function-

***Example* 1.** Consider a Gauss-Markov random process with $R_x(\tau)=e^{-\beta|\tau|}$. Then using calculus,

$$\Im(R_x)(\omega)=\int_{-\infty}^{\infty}e^{-\beta|\tau|}e^{-i\omega\tau}d\tau=2\Re e\left(\int_{0}^{\infty}e^{-(\beta+i\omega)\tau}d\tau\right)=2\Re e\left(\left.\frac{e^{-(\beta+i\omega)\tau}}{-(\beta+i\omega)}\right|_{0}^{\infty}\right)=2\Re e\left(\left.\frac{e^{-(\beta+i\omega)\tau}}{\beta+i\omega}\right|_{\infty}^{0}\right)=2\Re e\left(\frac{1}{\beta+i\omega}\right)=2\Re e\left(\frac{\beta-i\omega}{\beta^2+\omega^2}\right)$$

Hence, $\Im(R_x)(f) = \dfrac{2\beta}{\beta^2 + (2\pi f)^2}$ in units of power/Hz. □

Definition 2. For a *wss* random process with $S_x(\omega) \overset{\Delta}{=} \Im(R_x)(\omega)$ is called the ***power spectral density*** (psd) for $x(t)$. Even so, its units are in power per Hz; not power per rad/sec.

NOTE 1: The reason for this 'weirdness' of the units stems from the $1/2\pi$ factor included in the inverse Fourier transform. It is that factor that changes the units from power per rad/sec. to power per Hz.]

NOTE 2: One might well ask about the use of the term 'power'. After all, the units of $R_x(\tau) \overset{\Delta}{=} E[x(t)x(t+\tau)]$ is the square of the units of $x(t)$. The answer is that often the squared units of $x(t)$ are related to the units of power. For example, if $x(t)$ is the voltage across a resistor having resistance *R*, then the power associated with this voltage is $x^2(t)/R$, which has units of volts^2/Ohm. If $R = 1\Omega$, then the numerical value of the power is $x^2(t)$.]

***Example* 1 continued.** For $\beta = 0.5$, $S_x(\omega) = \dfrac{1}{0.25 + \omega^2}$.

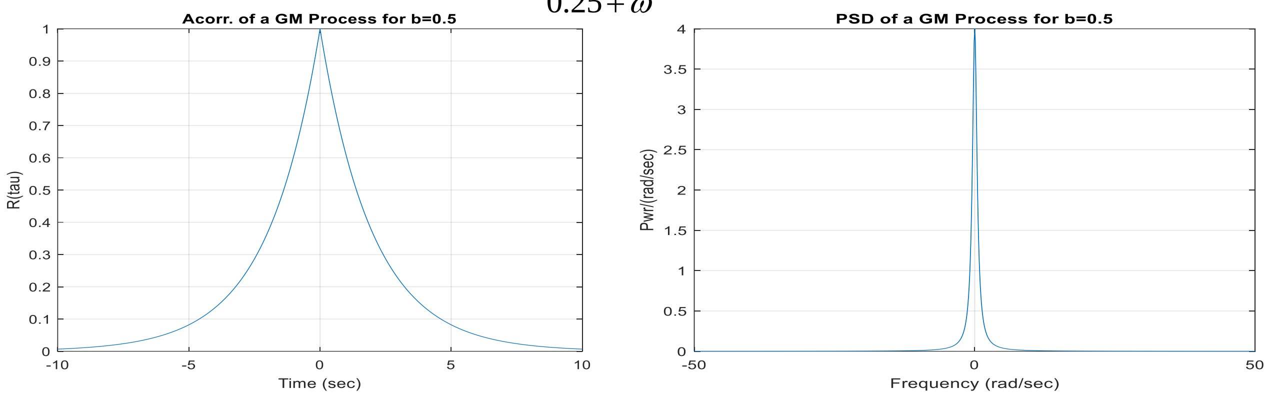

Figure 1. The autocorrelation and *psd* functions for a GM process with $\beta = 0.5\, rad/\sec$.

***Example* 2.** Continuous-time white noise has $R_x(\tau) = \sigma^2 \delta(\tau)$ where $\delta(\tau)$ is the Dirac-delta function. Hence, the *psd* is $S_x(\omega) = \sigma^2$ for all $\omega \in (-\infty, \infty)$. Since the total power is the area associated with $S_x(\omega) = \sigma^2$, the white noise process has *infinite* variance. The parameter σ^2 is called the *variance intensity*. Clearly, such a process does not exist. On the other hand, consider *band-limited* white noise that has $S_x(\omega) = c$ for all $\omega \in (-\omega_{bw}, \omega_{bw})$. We will now recover $R_x(\tau)$ for this process.

$$R_x(\tau) = \frac{1}{2\pi}\int_{-\omega_{bw}}^{\omega_{bw}} c e^{i\omega\tau} d\omega = \frac{\sigma^2}{2\pi}\left(\frac{e^{i\omega\tau}}{i\tau}\right)_{-\omega_{bw}}^{\omega_{bw}} = \frac{c}{2\pi}\left(\frac{e^{i\omega_{bw}\tau} - e^{i\omega_{bw}\tau}}{i\tau}\right) = \frac{c}{\pi\tau}\left(\frac{e^{i\omega_{bw}\tau} - e^{i\omega_{bw}\tau}}{2i}\right).$$

Recalling that $\dfrac{e^{i\theta} - e^{-i\theta}}{2i} = \sin\theta$ gives $R_x(\tau) = \dfrac{c}{\pi\tau}\sin(\omega_{bw}\tau) = \dfrac{c\omega_{bw}}{\pi}\left(\dfrac{\sin(\omega_{bw}\tau)}{\omega_{bw}\tau}\right) \overset{\Delta}{=} \dfrac{c\omega_{bw}}{\pi} Sa(\omega_{bw}\tau)$.

Notice that $R_x(0) = \dfrac{1}{2\pi}\displaystyle\int_{-\omega_{bw}}^{\omega_{bw}} c e^{i\omega 0} d\omega = \dfrac{c}{2\pi}(2\omega_{bw}) = \dfrac{c\omega_{bw}}{\pi}$. Hence, for a specified $R_x(0) = \sigma^2$, the *psd* constant is $c = \sigma^2\pi/\omega_{bw}$. We can also write this as $c = \sigma^2\pi/2\pi f_{bw} = \sigma^2/2 f_{bw}$.

Hence, the constant c has units of pwr/Hz and the two-sided bandwidth $2f_{bw}$ has units of Hz. A plot of $R_x(\tau)$ for $\sigma^2 = 1$ and $\omega_{bw} = 5$ is shown at right. Notice that the first null in $Sa(\omega_{bw}\tau)$ occurs when $\omega_{bw}\tau = \pi$, or when $\tau = \pi / \omega_{bw}$. Hence, the wider the bandwidth, the narrower the autocorrelation function and the larger $R_x(0)$. □

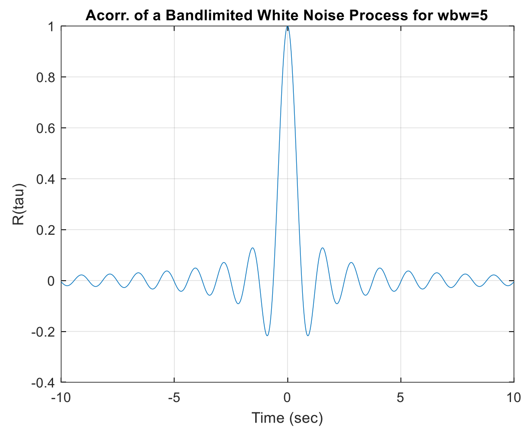

Figure 2. Plot of $R_x(\tau)$ for $\omega_{bw} = 5$.

From *Example* 2, we have the following Fourier transform pair:

$$X(\omega) = \begin{cases} a & for\ |\omega| < \omega_0 \\ 0 & otherwise \end{cases} \quad \leftrightarrow \quad x(t) = \frac{a\omega_0}{\pi} Sa(\omega t)\,.$$

Example* 3.** The $Sa(ax)$ function described in *Example* 2 is an extremely important function. It is called the ***sinc function. For this reason, we will pursue it in more detail in this example. Specifically, we will develop the Fourier transform of $x(t) = \begin{cases} a & for\ |t| < T/2 \\ 0 & otherwise \end{cases}$:

$$X(\omega) = \int_{-T/2}^{T/2} a\, e^{-i\omega t} dt = a\left(\frac{e^{-i\omega t}}{-i\omega}\right)_{-T/2}^{T/2} = a\left(\frac{e^{i\omega T/2} - e^{-i\omega T/2}}{i\omega}\right) = \frac{2a}{\omega}\left(\frac{e^{i\omega T/2} - e^{-i\omega T/2}}{2i}\right) = \frac{2a}{\omega}\sin(\omega T/2)\,.$$

Writing this as $X(\omega)\dfrac{2aT}{\omega T}\sin(\omega T/2) = (aT)\dfrac{\sin(\omega T/2)}{\omega T/2} = (aT)Sa(\omega T/2)$, we have the Fourier transform pair:

$$x(t) = \begin{cases} a & for\ |t| < T/2 \\ 0 & otherwise \end{cases} \quad \leftrightarrow \quad X(\omega) = (aT)Sa(\omega T/2)\,. \quad \square$$

The Importance of the Rectangular Window and Its Fourier Transform- the Sinc Function-

To appreciate the importance of the rectangular window and its Fourier transform, it helps to understand the following property of Fourier transform pairs:

Property 1: *Multiplication in one domain is equivalent to convolution in the other domain*.

***Example* 4.** A random process $\{x(t); t \in (-\infty, \infty)\}$ can only be observed over a finite time $[0, T]$. Let $w(t) = \begin{cases} 1 & for\ t \in [0, T] \\ 0 & otherwise \end{cases}$ denote a rectangular window of length T and height 1.0. We can then write the finite-time observation of $\{x(t); t \in (-\infty,\infty)\}$ as $\{x_w(t) = x(t)w(t); t \in (-\infty,\infty)\}$. In words, we have assumed that $x_w(t)$ equals *zero* for $t \notin [0,T]$. It follows from Property 1 that $X_w(\omega) = X * W(\omega)$, where

$$W(\omega) = \int_0^T e^{-i\omega t}dt = \left.\frac{e^{-i\omega t}}{-i\omega}\right|_{t=0}^{T} = \frac{1-e^{-i\omega T}}{i\omega} = \frac{2Te^{-i\omega T/2}}{\omega T}\left(\frac{e^{i\omega T/2} - e^{-i\omega T/2}}{2i}\right) = 2TSa(\omega T/2)e^{-i\omega T/2}.$$

Now suppose that $x(t) = A\sin(\omega_0 t + \Theta)$ where $\Theta \sim Uniform[0, 2\pi)$. Then

$$X(\omega) \stackrel{\Delta}{=} \int_{-\infty}^{\infty} A\sin(\omega_0 t + \Theta)e^{-i\omega t}dt = A\int_{-\infty}^{\infty} A\sin(\omega_0 t + \Theta)e^{-i\omega t}dt.$$

Recalling the identity $\sin\alpha = (e^{i\alpha} - e^{-i\alpha})/2i$ gives

$$X(\omega) = \int_{-\infty}^{\infty}\left(\frac{e^{i(\omega_0 t+\Theta)} - e^{-i(\omega_0 t+\Theta)}}{2i}\right)e^{-i\omega t}dt = \frac{e^{i\Theta}}{2i}\int_{-\infty}^{\infty} e^{i\omega_0 t}e^{-i\omega t}dt - \int_{-\infty}^{\infty} e^{i(-\omega_0)t}e^{-i\omega t}dt.$$

Use the Fourier transform pair relation $e^{i\omega_0 t} \quad \leftrightarrow \quad 2\pi\,\delta(\omega - \omega_0)$ gives

$$X(\omega) = \frac{e^{i\Theta}\pi}{i}\left[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)\right].$$

This can be simplified by noting that $i = \cos(\pi/2) + i\sin(\pi/2) = e^{i\pi/2}$. Hence,

$$X(\omega) = e^{i(\Theta-\pi/2)}\pi\left[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)\right].$$

It follows that

$$X_w(\omega) = (X*W)(\omega)\sqrt{2\pi} = e^{i(\Theta-\pi/2)}\pi\sqrt{2\pi}\left[W(\omega-\omega_0) - W(\omega+\omega_0)\right].$$

Note: the $\sqrt{2\pi}$ factor is the appropriate scale factor in this *convolution theorem*. □

In words, the effect of the rectangular window on the Fourier transform of the sinusoid is that the two delta functions have been replaced by two sinc functions. Effectively, the spikes have smeared out. For a small window the smearing effect will be large; whereas for a large window it will be more localized.

Remark 3. We have just alluded to what, more generally is known as the Heisenberg Uncertainty principle: The more you attempt to localize in one domain, the more uncertainty you have in the associated Fourier transform domain. □

***Example* 5.** The 'ideal' low pass filter (LPF) is the rectangular frequency window $W(\omega) = \begin{cases} 1 \text{ } for \text{ } \omega \in [-\omega_{bw}, \omega_{bw}] \\ 0 \text{ } otherwise \end{cases}$.

Applying it gives $X_w(\omega) = X(\omega)W(\omega)$. Its effect in the time domain is:

$$x_w(t) = (1/\sqrt{2\pi})(x * w)(t)$$

This time domain convolution will smear out the details in $x(t)$. To demonstrate this consider discrete-time white noise with sampling interval equal to 1 second. The analysis bandwidth is then $\omega \in [-\pi, \pi)$. We will apply an ideal LPF having a bandwidth frequency $\omega_{bw} = \pi / 4$ to this white noise. The results are shown at right. As expected, the details of the white noise have been reduced (i.e. smeared out). □

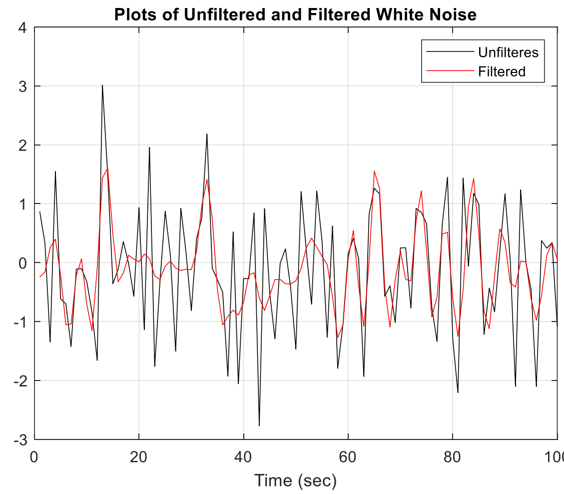

Figure 3. Plots of unfiltered and filtered white noise.

The Unbiased Versus the Biased Lagged-Product Autocorrelation Estimator-

The unbiased and biased lagged-product estimators of $R_x(\tau)$ are:

$$\widehat{R}_X^{(ub)}(\tau) = \frac{1}{T-\tau}\int_0^{T-\tau} x(t)x(t+\tau)dt \quad \text{and} \quad \widehat{R}_X^{(b)}(\tau) = \frac{1}{T}\int_0^{T-\tau} x(t)x(t+\tau)dt$$

Using the linearity of E(*) it is easy to show that:

$$E\left[\widehat{R}_X^{(ub)}(\tau)\right] = R_X(\tau) \quad \text{and} \quad E\left[\widehat{R}_X^{(b)}(\tau)\right] = \left(\frac{T-|\tau|}{T}\right)R_X(\tau) \overset{\Delta}{=} w_T(\tau)R_X(\tau).$$

The window $w_T(t) = \begin{cases} \dfrac{T-|t|}{T} & for\ t \in [-T,T] \\ 0\ otherwise \end{cases}$ is a *triangular window* of width 2*T*. In fact, it can be obtained by *convolving* the rectangular window $w(t) = \begin{cases} 1\ for\ t \in [-T/2, T/2] \\ 0\ otherwise \end{cases}$ with itself. Hence, by the *convolution theorem* we have $W_T(\omega) = \sqrt{2\pi}\left[TSa(\omega T/2)\right]^2$. The effect of convolving this window with the PSD $S_x(\omega)$ is to smear out the detail in $S_x(\omega)$. In other words, it *decreases the spectral resolution.*

Remark 4. In addition to decreasing the spectral resolution, the rectangular window also changes the nature of deterministic random processes. They are converted to regular random processes. For example, recall from *Example* 4 that for the deterministic random process $x(t) = A\sin(\omega_0 t + \Theta)$ $x(t) = A\sin(\omega_0 t + \Theta)$ where $\Theta \sim Uniform[0, 2\pi)$ we found that $X(\omega) = e^{i(\Theta - \pi/2)}\pi\left[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)\right]$. Hence, it is natural to assume the corresponding *psd* will include these two delta functions. We then found that $X_w(\omega)=e^{i(\Theta-\pi/2)}\pi\sqrt{2\pi}\left[W(\omega-\omega_0)-W(\omega+\omega_0)\right]$. And so, once again, it is natural to presume that the corresponding *psd* will no longer include these delta function. Rather, it will be more well-behaved, in the sense that this *psd* is actually a *function* of frequency; as opposed to the delta function that is not a regular function at all. Recall that the *psd* is a [power spectral density. Its units are power/Hz. The psd is analogous to a force density, which we call pressure. The units of pressure are force/unit-length. One need to compute force by integrating pressure over a given length. If that length includes a single point, this integral is zero. □

We will now address one of the most important theorems in relation to computation of the *psd*:

My Version of the Wiener-Kinchin Theorem. Let $x(t)$ be any regular random process with *psd* $S_x(\omega)$. Let $x^{(T)}(t)$ be a windowed version of $x(t)$, where the time window as width *T*. We showed that its *psd* is:

$$E\left[\widehat{R}_X^{(b)}(\tau)\right] \quad = \quad \left(\frac{T-|\tau|}{T}\right)R_X(\tau)\overset{\Delta}{=}w_T(\tau)R_X(\tau)\,.$$

For convenience, denote this triangular-windowed autocorrelation function as $R_x^{(T)}(\tau)$. The corresponding *psd* is:

$$S_x^{(T)}(\omega)=\int_{-T}^{T}R_x^{(T)}(\tau)\,e^{-i\omega\tau}d\tau\,. \tag{1}$$

On pp.71-72 of the book it is shown that: $S_x^{(T)}(\omega)=E\left[\left|\frac{1}{\sqrt{T}}X^{(T)}(\omega)\right|^2\right]$. (2)

This expression for the windowed *psd* is my version of the *Wiener-Kinchin Theorem.*

The relation (s) gives an alternative to taking the Fourier transform of the lagged-product autocorrelation function. Suppose that the window length *T* can be partitioned into *n* smaller windows, each having length Δ. Denote the process associated with the k^{th} window as $x^{(k)}(t)$. Its Fourier transform is $X^{(k)}(\omega)$. We can then estimate (2) via the average:

$$\widehat{S}_x^{(T)}(\omega)=\frac{1}{n}\sum_{k=1}^{n}\left|\frac{1}{\sqrt{\Delta}}X^{(k)}(\omega)\right|^2\,. \tag{3}$$

If we do not partition *T* and use the entire window, we have:

$$\widehat{S}_x^{(T)}(\omega)=\left|\frac{1}{\sqrt{T}}X(\omega)\right|^2\,. \tag{4}$$

The estimator (4) is called the *periodogram*. In words, you simply compute the Fourier transform of $\{x(t);t\in(0,T)\}$, and then take its magnitude-squared as the estimator of the *psd*. Now, an average of one is nothing to write home about. By partitioning *T* and taking an average of *n* mod-squared Fourier transforms we can significantly reduce the uncertainty of the *psd* estimator. However, this comes at a price. The sinc function associated with a window of size $\Delta=T/n$ will be wider than the sinc function associated with a window of size *T* by a factor of *n*. In words, we reduce the spectral resolution of the estimator. It must be emphasized:

The periodogram* (4) *is a <u>horrible</u> estimator of the psd!

The challenge is to determine the value of *n* such that the estimator will have acceptably low uncertainty, while at the same time having reasonable spectral resolution.

Finally, it should be mentioned that this problem has been around for a long time, and that many methods of attacking it have been proposed. One method is to use a specified overlap of the sub-windows. For a 50% overlap, the second half of a given window would be the first half of the next window. Having more sub-windows allows one to use a larger sub-window size. Matlab offers this method (as well as others). It is known as Welch's method, and is the code pwelch.m.

CONCLUSION: This lecture was packed with elements of *psd* analysis. There are entire graduate courses on this topic. You should not be discouraged if your head is now spinning. Re-read these lecture notes again and again. Go to Google to help improve your understanding. The *psd* is an indispensable topic in analysis of random processes. For this reason, we will return to it in the context of our next topic, which is linear systems.

Footnote (10/11/19): Consider the three Fourier Transforms:

(i): $X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-i\omega t} dt$; (ii) $X(\omega) = \lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x(t)\, e^{-i\omega t} dt$; $X(\omega) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, e^{-i\omega t} dt$.

The form (i) is the correct one for deterministic absolutely integrable $x(t)$.

The form (ii) is the correct one for a *regular wss* random process $x(t)$.

The form (iii) is the correct one for a periodic $x(t)$.

QUESTION: Suppose we have $x(t) = e^{-t} + w(t) + \sin(2t)$. Which transform is appropriate for $x(t)$?