A6523
Signal Modeling, Statistical Inference and Data Mining in Astrophysics
Spring 2011

Reading
Lecture 11

• Nongaussianity examples
• Fourier Transform Spectral Estimators

Projects: need to choose a topic

• can be theoretical, numerical simulation, data processing
• e.g. explore an algorithm and underlying statistics
\[ X(t) = \sum_{j} a_j h(t-t_j) \quad \gamma = \text{rate of events} \]
\[ w_h = \text{width of } h(t) \]

Dense Random Walk: \( \gamma w_h \gg 1 \)

\[ n(t) = \text{white, Gaussian noise} \]

Amplitude vs. Time

- First Integral of \( n(t) \)
- Second Integral of \( n(t) \)
- Third Integral of \( n(t) \)
\( \eta(t) = \text{white, non-gaussian} \)

Sparse Random Walk

- Amplitude
- Time

Graph showing the behavior of a white, non-gaussian process over time.
White noise + three orders of random walk
- generate white noise
- integrate once to get $RW_1$
- integrate again to get $RW_2$
- integrate again to get $RW_3$
Multiple realizations: Nonstationary statistics
RMS(t) for WN, RM$_k$, k=1,3

RMS vs time calculated from 100 realizations

- WN, $\sigma =$ constant
- RW$_1$, $\sigma \propto t^{1/2}$
- RW$_2$, $\sigma \propto t^{3/2}$
- RW$_3$, $\sigma \propto t^{5/2}$

Time (steps)
Correlation Function Practicalities

Consider the autocorrelation function (ACF) of a zero-mean WSS process. We want to consider how the ACF converges as well as understand what the ACF actually quantifies.

Figure 1 shows a time series and its ACF along with an ACF averaged over 10 realizations of the time series.

![Time series and ACF](image)

Fig. 1.— Top panel: Time series of Gaussian noise with unity variance and a correlation time of 21 steps. Bottom panel: ACF of the time series in the top panel along with the ACF averaged over 10 realizations of the time series.
The time series was created by taking a realization of white, Gaussian noise and smoothing it with a boxcar filter with width of 21 samples.

Features of the ACF include:

1. The maximum at zero lag has a value equal to the variance in the time series (set to unity).
2. The feature that maximizes at zero lag is the same in both ACFs.
3. The decay from the maximum is on a time scale $\sim 20$ steps, which is order of the smoothing time used to create the time series.
4. There are statistical variations centered on zero correlation. These variations are larger in the ACF calculated from single-realization and are estimation errors in the ACF.

The width of the persistant feature in the ACF is the autocorrelation time of the process, which we define as $W_y$. This quantifies the time interval over which the process decorrelates.

The estimation error in the ACF at larger lags is determined by the number of independent fluctuations $N_i$ in the time series used. In a single time series of length $T$, this number is

$$N_i \approx \frac{T}{W_y}$$

For the example, $N_i \approx 1024/21 \approx 50$. The estimation error in the ACF for a single time series will be

$$\delta C_y \approx C_y(0)/\sqrt{N_i} \approx 0.14$$

so we expect variations about zero at approximately this level.

For the 10-realization average, we expect the estimation errors to decrease by another factor of $1/\sqrt{10}$ to about 0.045.
Calculating Correlation Functions

Consider a discrete data set with $N$ equally spaced samples, $x_i, i = 1, N$.

Two slightly different estimators can be used to calculate the ACF. The first normalizes by the number of terms (lagged products) in the sum for each lag, $N - |\tau|:

$$
\hat{R}_\tau = \begin{cases} 
\frac{1}{(N - |\tau|)} \sum_{i=1}^{N-|\tau|} x_i x_{i+\tau}, & 0 \leq \tau < N \\
R^*_\tau & \tau < 0
\end{cases}
$$

This normalization yields an unbiased result because it can be shown that $\langle \hat{R}_\tau \rangle = R_\tau$, i.e. the ensemble average of the estimator equals the ensemble-average ACF. However, for $\tau \sim N$, the estimation errors are large because there are few terms in the sum and their departure from the true ACF is amplified by the $1/(N - |\tau|)$ factor.

An alternative normalization divides by $N$ instead of $N - |\tau|$. This biases the estimator, but only at large lags owing to the presence of a triangle function, $1 = |\tau|/N$ that multiples the true ACF. The advantage of this normalization is that it keeps the estimation errors at large lags small.

**Unequal sampling:** The ACF of a time series that isn’t sampled uniformaly can be calculated by binning in lag. For example, we can write (where now $\tau$ is in time units rather than being an index)

$$
\hat{R}_\tau = \frac{1}{N_\tau} \sum_i \sum_j x_i x_j^*, \quad N_\tau = \sum_i \sum_j 1.
$$

Here $N_\tau$ is the number of lagged products summed in a bin of size $\Delta \tau$. Clearly the bin size $\Delta \tau$ needs to be chosen carefully so that structure in the ACF is not lost (bins too large) or that there are too few counts per bin (bins too small). Note also that (a) bins do not have to be equal and (b) in fact, they can be spaced logarithmically, e.g. $\Delta \tau/\tau = \text{constant}$.
Estimation Errors in Correlation Functions

Errors in the ACF have several causes:

1. Stochasticity of the “signal” part of the time series

2. Additive noise.

Let the measurement model be the sum of a signal $s$ and contaminating zero-mean, white noise $n$ (both real):

$$x(t) = s(t) + n(t)$$

The ACF of $x$ has an ensemble average

$$R_x(\tau) = \langle [s(t) + n(t)][s(t + \tau) + n(t + \tau)] \rangle$$

$$= \langle s(t)s(t + \tau) \rangle + \langle n(t)n(t + \tau) \rangle + \text{cross terms}$$

$$= R_s(\tau) + R_n(\tau) \quad \text{(variances add)}$$

What determines the errors in the estimated ACF?

Both the signal (if stochastic) and the additive noise. Many results in the literature consider only the errors from additive noise. As we have seen in Figure 1, however, a signal with no additive noise produces an ACF estimate with errors.

Considering again the unbiased estimate $\hat{R}_x$, as with any estimator we calculate the mean and variance of the estimated quantity. The mean is equal to the true ensemble average, as before. The mean square is (for $\tau \geq 0$ and $x$ still real)

$$\langle \hat{R}_x^2 \rangle = \frac{1}{(N-\tau)^2} \sum_i \sum_j \langle x_ix_{i+\tau}x_jx_{j+\tau} \rangle.$$ 

If $x$ has Gaussian statistics, it is easy to work out the fourth moment in the summand. How many terms does the fourth moment expand into? In this case, the mean square ACF estimate and the estimation error can be worked out.
FOURIER TRANSFORM AND POWER SPECTRUM ESTIMATE FOR A STOCHASTIC PROCESS

Another stochastic integral is the Fourier transform. As stated before, the FT of a random process (WSS) $X(t)$

$$\tilde{X}(f) = \int_{-\infty}^{\infty} dt \ x(t) \ e^{-i\omega t} \quad \text{recall WSS} \Rightarrow \langle X^2(t) \rangle = \text{constant in } t$$

cannot exist because the integral diverges. Luckily we need consider windowed or finite integrals in order to model experimental situations:

$$\tilde{X}_T(f) = \int_{-T}^{T} dt \ x(t) \ e^{-i\omega t}$$

What should an estimator of the power spectrum $S_x(f)$ look like? Recall for deterministic functions that

$$\int_{-\infty}^{\infty} dt \ f^*(t) \ f(t + \tau) \Leftrightarrow |\hat{F}(f)|^2$$

So we expect that an estimator (denoted with a carat) for the power spectrum of a process $x(t)$ would have the form:

$$\hat{S}_x(f) = \text{const} \ |\tilde{X}_T(f)|^2$$

and an appropriate value of the constant is $= \frac{1}{2T}$ so

$$\hat{S}_x(f) = \frac{1}{2T} |\tilde{X}_T(f)|^2$$

This ensures that the corresponding ACF has the correct units:
It can be shown that the estimator satisfies the Wiener-Khinchin theorem (which applies to ensemble average quantities), where

\[ \hat{C}_x(\tau) \Leftrightarrow \hat{S}_x(f) \]

is

\[ \hat{C}_x(\tau) = \frac{1}{2T} \int_{-\infty}^{\infty} df \ e^{i\omega \tau} |\hat{X}_T(f)|^2 \]

\[ = \frac{1}{2T} \int_{-\infty}^{\infty} df \ e^{i\omega \tau} \int_{-T}^{T} dt \ dt' \ x(t) \ x^*(t') \ e^{-i\omega(t-t')} \]

\[ = \frac{1}{2T} \int_{-T}^{T} dt \ dt' \ x(t) \ x^*(t') \ \int_{-\infty}^{\infty} df \ e^{i\omega(t'+\tau-t)} = \delta(t'+\tau-t) \]

\[ \hat{C}_x(\tau) = \frac{1}{2T} \int_{-T}^{T} dt \ x(t) \ x^*(t - \tau) \]

so \[ \hat{C}_x(0) = \frac{1}{2T} \int_{-T}^{T} dt \ |x(t)|^2 \]

= estimate for \( \langle |x(t)|^2 \rangle \) as expected for \( C_x(0) \)
How good an estimator is \( \hat{S}_x(f) \) for the power spectrum \( S_x(f) \)?

Answer: terrible! Recall that \( \hat{S}(f) \) is itself a random process (since \( \hat{S}_x \) for fixed \( f \) is a random variable). Thus, we may fairly ask what the convergence properties of \( \hat{S}_x(f) \) are just as we investigated \( Y \equiv \frac{1}{T} \int_0^T dt \ X(t) \) and found that \( \langle Y \rangle = \langle X \rangle \) and \( \sigma_y \to 0 \) as \( T \to \infty \) so long as the correlation function of \( X(t) \) decayed sufficiently quickly.

Mean \( \langle \hat{S}_x(f) \rangle \underset{T \to \infty}{\to} S_x(f) \) so \( \hat{S}_x(f) \) converges in the mean.

Variance \( \sigma_{\hat{S}}^2(f) \equiv \langle \hat{S}_x^2(f) \rangle - \langle \hat{S}_x(f) \rangle^2 \) does not decay to zero as \( T \to \infty \):

\[
\lim_{T \to \infty} \sigma_{\hat{S}}^2(f) \neq 0.
\]

Conclusion: The squared magnitude of a finite Fourier transform of a WSS process is a poor estimate for the power spectrum \( S_x(f) \) (an ensemble average quantity).

Why \( \hat{S}_x(f) \) is a poor estimator:

1. \( \hat{S}_x(f) \) is a \( \chi^2 \) r.v. in the limit where \( W_x \ll T \)

\[
\Rightarrow \frac{\sigma_{\hat{S}_x}}{\langle \hat{S}_x \rangle} \equiv 1 \text{ independent of } T.
\]

2. From the point of view of number of degrees of freedom, the number of degrees of freedom in the data \( N_{dof} \sim \frac{T}{W_x} \) may be large, but the number of degrees of freedom in the spectral estimate (per independent frequency bin of width \( \Delta f \approx \frac{1}{2T} \)) is small.
Intuitive Approaches

We can see the same result by bypassing the brute force details by being clever.

Method 1: Consider

\[ \tilde{X}_T(\omega) \equiv \int_{-T}^{T} dt \ x(t) \ e^{-i\omega t} \]

infinite sum of random variables

View this as the sum of many random variables. How many?

Let \( W_x \) = autocorrelation width of \( x(t) \) as in discussion on ergodicity of \( Y(t) = T^{-1} \int_0^T dt \ x(t) \).

By definition, \( W_x \) is the time scale over which two samples \( X(t_1) \) and \( X(t_2) \) become independent.

Therefore \( N = \frac{2T}{W_x} \approx \) number of independent samples of \( x(t) \)

If \( N \gg 1 \) then we invoke the Central Limit Theorem and say that \( \tilde{X}_T(\omega) \) becomes a Gaussian random variable with zero mean if \( x(t) \) is zero mean.

Break \( \tilde{X}_T(\omega) \) into real and imaginary parts:

\[ \tilde{X}_T(\omega) = R(\omega) + i \ I(\omega) \]

It can be shown that \( R(\omega) \) and \( I(\omega) \) are independent r.v.’s; therefore, they are zero mean, independent Gaussian r.v.’s.
Proof:

\[
R(\omega) = \int_{-2T}^{2T} dt \, x(t) \, \omega s t
\]

\[
I(\omega) = -\int_{-2T}^{2T} dt \, x(t) \, \sin \omega t
\]

\[
\Rightarrow \quad \langle R(\omega) \, I(\omega) \rangle = -\int \int dt_1 \, dt_2 \, \langle x(t_1) \, x(t_2) \rangle \, \cos \omega t_1 \sin \omega t_2 \left\{ \frac{1}{2} \left[ \sin \omega (t_1 + t_2) + \sin \omega (t_2 - t_1) \right] \right\}
\]

\[
= -\frac{1}{2} \left\{ \int \int dt_1 \, dt_2 \, R_x(t_1 - t_2) \sin \omega (t_1 + t_2) + \int \int dt_1 \, dt_2 \, R_x(t_1 - t_2) \sin \omega (t_2 - t_1) \right\}.
\]

The first term integrates to zero for \( W_x \ll T \) while the second term is the integral of the product of odd and even functions. Therefore

\[
\hat{S}_T(\omega) = \frac{1}{2T} |\hat{X}_T(\omega)|^2 = \frac{1}{2T} \left[ R^2(\omega) + I^2(\omega) \right]
\]

Where the sum of squares is of two independent Gaussian r.v.’s \( \equiv \chi^2_2 \) implying that \( \hat{S}_T(\omega) \) is a \( \chi^2_2 \) r.v. with mean \( \langle \hat{S}_T(\omega) \rangle = S(\omega) = \text{true power spectrum.} \)

Let

\[
p(\omega) \equiv \hat{S}_T(\omega).
\]

Then the PDF of \( p(\omega) \) is

\[
f_p(p) = \frac{1}{\langle p \rangle} e^{-p/\langle p \rangle} U(p).
\]

It can be shown that \( \langle p^2 \rangle = 2 \langle p \rangle^2 \Rightarrow \varepsilon = \sigma_p/\langle p \rangle = 1 \) as before.
Method 2: Another way of understanding why \( \hat{S}_T(\omega) \) does not converge as \( T \to \infty \) is to consider the number of degrees of freedom in each independent frequency bin.

- Note that
  \[
  \tilde{X}_T(\omega) \equiv \int_{-T}^{T} dt \, x(t) \, e^{-i\omega t} = \int_{-\infty}^{\infty} dt \, x(t) \, W_T(t) \, e^{-i\omega t}
  \]
  where \( W_T \) is a window function,
  \[
  W_T(t) = \begin{cases} 
  1 & |t| \leq T \\
  0 & |t| \geq T 
  \end{cases}
  \]
  or
  \[
  \tilde{X}_T(\omega) = \tilde{X}(\omega) * \frac{2\sin \omega T}{\omega}
  \]
  As usual, multiplication by \( W_T(t) \) in the time domain corresponds to convolution by \( \frac{\sin \omega T}{\omega} \) in the frequency domain.

  A frequency cell or bin has a width \( \Delta \omega \approx \frac{\pi}{T} \) or \( \Delta f \approx \frac{1}{2T} \).

- Let \( W_t = \) correlation time in the time domain. Then the number of independent fluctuations in the time series is \( N_t = \frac{2T}{W_t} \).

- Let \( W_\omega = \) width of spectrum (bandlimited). Then the number of frequency cells into which the variance is divided is
  \[
  N_\omega = \frac{W_\omega}{\Delta \omega} = \frac{T}{\pi} \cdot W_\omega
  \]

- The number of degrees of freedom (d.o.f.) per frequency cell is
  \[
  N_{d.o.f.} = \frac{N_t}{N_\omega} = \frac{\# \text{ of independent data points}}{\# \text{ of frequency cells}} = \frac{2T/W_t}{TW_\omega/\pi} = \frac{2\pi}{W_tW_\omega}.
  \]
But the uncertainty principle \( \Rightarrow W_t W_\omega \geq 2\pi \), so \( N_{d.o.f.} \approx 1 \) for each part of F.T.

**Interpretation:**

As \( T \to \infty \) more and more independent fluctuations contribute to the integral, but these are being spread into more and more frequency bins so that the \# d.o.f. per bin remains the same and small \( \Rightarrow \) large errors.

\( N_{d.o.f.} \approx 1 \Rightarrow |\tilde{X}_T(\omega)|^2 \) will have 2 d.o.f. per cell, as before.

**Solution to the convergence problem: increase the number of degrees of freedom in a frequency cell.**

The simplest approach to take is to average spectral estimates. Obtain spectra from \( L \) realizations of length \( 2T \) and average; i.e. find \( \hat{S}_T(\omega) \) for a block of data, repeat for \( L \) blocks of data, and average.

\[
\hat{S}_{T,L}(\omega) = L^{-1} \sum_{j=1}^{L} \hat{S}_T;_j(\omega)
\]

\[
\text{error} : \frac{\text{Var}[\hat{S}_{T,L}(\omega)]^{1/2}}{\langle \hat{S}_{T,L}(\omega) \rangle} = L^{-1/2}
\]

10% \( \Rightarrow L = 100 \)

(Best method if unlimited amount of data are available.) We will talk about this and other methods later.
Direct Calculation of the Mean and Variance of the Spectral Estimate

We will use continuous notation for now.

The spectral estimator for a WSS process is

\[
\hat{S}_T(\omega) \equiv \frac{1}{2T} |\tilde{X}_T(\omega)|^2 \text{ where } \tilde{X}_T(\omega) = \int_{-T}^{T} dt \ x(t) \ e^{-i\omega t}.
\]

Properties of the estimator: As usual, we want to calculate the mean and variance of the estimator.

Mean:

The ensemble average is

\[
\langle \hat{S}_T(\omega) \rangle = \langle \frac{1}{2T} |\tilde{X}_T(\omega)|^2 \rangle
\]

\[
= \langle \frac{1}{2T} \int \int dt_1 \ dt_2 \ x(t_1) \ x^*(t_2) \ e^{-i\omega(t_1-t_2)} \rangle
\]

\[
= \frac{1}{2T} \int \int_{-T}^{T} dt_1 \ dt_2 \ R_X(t_1-t_2) \ e^{-i\omega(t_1-t_2)}.
\]

Use, as before, the coordinate transformation

\[
\tau = t_1 - t_2 \\
\tau = \frac{1}{2}(t_1 + t_2)
\]

(1)

where \( dz \ d\tau = dt_1 \ dt_2 \).
The integration limits are transformed as follows:

\[
\begin{align*}
\tau : 0 & \to 2T \quad z : -T + \frac{\tau}{2} & \to T - \frac{\tau}{2} \\
-2T & \to 0 \quad z : -T - \frac{\tau}{2} & \to T + \frac{\tau}{2}
\end{align*}
\]

\[
\langle \hat{S}_T(\omega) \rangle = \frac{1}{2T} \left[ \int_0^{2T} d\tau \ R_x(\tau) e^{-i\omega \tau} \int_{-T+\tau/2}^{-T+\tau/2} dz + \int_{-2T}^{0} d\tau \ R_x(\tau) e^{-i\omega \tau} \int_{-T-\tau/2}^{-T+\tau/2} dz \right]
\]

\[
= \frac{1}{2T} \left\{ \int_0^{2T} d\tau \ R_x(\tau) e^{-i\omega \tau} 2T \left(1 - \frac{\tau}{2T}\right) + \int_{-2T}^{0} d\tau \ R_x(\tau) e^{-i\omega \tau} 2T \left(1 + \frac{\tau}{2T}\right) \right\}
\]

\[
= \frac{1}{2T} \int_{-2T}^{2T} d\tau \ R_x(\tau) e^{-i\omega \tau} 2T \left[1 - \left|\frac{\tau}{2T}\right| \right]
\]

or

\[
\langle \hat{S}_T(\omega) \rangle = \int_{-2T}^{2T} d\tau \ R_x(\tau) e^{-i\omega \tau} \left[1 - \left|\tau\right|/2T\right]
\]

and

\[
\lim_{T \to \infty} \langle \hat{S}_T(\omega) \rangle = \int_{-\infty}^{\infty} d\tau \ R_x(\tau) e^{-i\omega \tau} \equiv S(\omega)
\]

\[\Rightarrow \hat{S}_T(\omega) \text{ is an unbiased estimator of } S(\omega) \text{ if the width of } R_x(\tau) \text{ is finite.}\]
Variance:

Now let’s look at the variance of the estimator:

$$\text{Var}[\hat{S}_T(\omega)] \equiv \langle \hat{S}_T^2(\omega) \rangle - \langle \hat{S}_T(\omega) \rangle^2$$

By definition,

$$\langle \hat{S}_T^2(\omega) \rangle = \left\langle \frac{1}{4T^2} |\tilde{X}_T(\omega)|^4 \right\rangle$$

$$= \frac{1}{4T^2} \int \int \int \int dt_1 dt_2 dt_3 dt_4 \langle x(t_1)x^*(t_2)x(t_3)x^*(t_4) \rangle e^{-i\omega[t_1-t_2+t_3-t_4]}$$

Assume the time series is real and assume a Gaussian process:

Then

$$\langle x(t_1)x(t_2)x(t_3)x(t_4) \rangle \equiv \langle x(t_1)x(t_2) \rangle \langle x(t_3)x(t_4) \rangle$$

$$+ \langle x(t_1)x(t_3) \rangle \langle x(t_2)x(t_4) \rangle$$

$$+ \langle x(t_1)x(t_4) \rangle \langle x(t_2)x(t_3) \rangle$$

$$\equiv R_X(t_1 - t_2) R_X(t_3 - t_4)$$

$$+ R_X(t_1 - t_3) R_X(t_2 - t_4)$$

$$+ R_X(t_1 - t_4) R_X(t_2 - t_3)$$
Plugging in, we get

\[
\frac{1}{4T^2} \iiint dt_1 \, dt_2 \, dt_3 \, dt_4 \left( e^{-i\omega[t_1-t_2+t_3-t_4]} \right)
\]

\[
= \frac{1}{4T^2} \left\{ \int dt_1 \, dt_2 \, R_X(t_1-t_2)e^{-i\omega[t_1-t_2]} \underbrace{\int dt_3 \, dt_4 \, R_X(t_3-t_4)e^{-i\omega[t_3-t_4]}}_{(\hat{S}_T(\omega)) \text{ as above}} \\
+ \int dt_1 \, dt_3 \, R_X(t_1-t_3)e^{-i\omega[t_1+t_3]} \int dt_2 \, dt_4 \, R_X(t_2-t_4)e^{-i\omega[t_2-t_4]} \\
+ \int dt_1 \, dt_4 \, R_X(t_1-t_4)e^{-i\omega[t_1-t_4]} \int dt_2 \, dt_3 \, R_X(t_2-t_3)e^{-i\omega[t_2-t_3]} \right\}
\]

Note the 1st and 3rd terms are of the form \(\langle \hat{S}_T(\omega) \rangle^2\). Therefore

\[
\langle \hat{S}_T^2(\omega) \rangle = 2\langle \hat{S}_T(\omega) \rangle^2 + \frac{1}{4T^2} \left| \int dt_1 \, dt_3 \, R_X(t_1-t_3)e^{-i\omega[t_1+t_3]} \right| \rightarrow 0
\]

\(-\rightarrow 0 \text{ as } T/W_x \rightarrow \infty \text{ (except for } \omega=0) \)

\(t_1, t_2 \rightarrow r, Z \text{ as before} \)
Looking at the double integral in the second term we have

\[
\int_0^T dt_1 \int_0^T dt_3 \, R_X(t_1 - t_3) e^{-i\omega t_1 + t_3} = \int_0^{2T} d\tau \, R_X(\tau) \left( \int_{-T+\tau/2}^{T-\tau/2} dz \, e^{-i\omega z} \right) + \int_{-2T}^0 d\tau \, R_X(\tau) \left( \int_{-2\tau}^{-T-\tau/2} dz \, e^{-i\omega z} \right)
\]

\[= \frac{1}{2T} \sin \omega T \sin \omega T + \frac{1}{\sin \omega T} \sin \omega T \sin \omega T \]

Therefore

\[
\frac{1}{2T} \int_0^T dt_1 \int_0^T dt_3 \, R_X(t_1 - t_3) e^{-i\omega t_1 + t_3} = \frac{1}{2T} \int_{-2T}^0 d\tau \, R_X(\tau) \sin[2\omega T(1 - |\tau|/2T)].
\]

Now this term \(\to 0\) as \(T \to \infty\) because

\[
\lim_{T \to \infty} \left| \frac{1}{2T} \int_{-2T}^0 d\tau \, R_X(\tau) \sin[2\omega T(1 - |\tau|/2T)] \right|^2 = \lim_{T \to \infty} \frac{1}{4T^2 \omega^2} \left| \int_{-\infty}^{\infty} d\tau \, R_X(\tau) \sin 2\omega T \right|^2
\]

finite if \(R_X(\tau)\) of finite width

\[= 0\]

In a slightly different approach for the second term we consider

\[
I_2 = \int_{-T}^T dt_1 \int_{-T}^T dt_3 \, R_X(t_1 - t_2) \, e^{-i\omega(t_1 + t_3)}
\]

Again let \(\tau = t_1 - t_3\) and \(z = \frac{1}{2}(t_1 + t_3)\): By calculating the integrals and taking a limit \(T \to 0\), it can be shown that \(I_2 \to 0\).
Consequently we have the fairly general result

$$\lim_{T \to \infty} \langle \hat{S}_T^2(\omega) \rangle = 2\langle \hat{S}_T(\omega) \rangle^2$$

The error in the spectral estimate is

$$\varepsilon = \frac{\sqrt{\text{Var}[\hat{S}_T(\omega)]}}{\langle \hat{S}_T(\omega) \rangle} = \frac{\sqrt{2\langle \hat{S}_T(\omega) \rangle^2 - \langle \hat{S}_T(\omega) \rangle^2}}{\langle \hat{S}_T(\omega) \rangle} = 1$$

i.e. 100% error ⇒ \( \hat{S}_T(\omega) \) does not converge to \( S(\omega) \) as \( T \to \infty \).

How to fix the properties of the spectral estimator:

1. increase the number of degrees of freedom through averaging of multiple spectral estimates;

2. increase the number of degrees of freedom by smoothing the spectrum: convolve the spectral estimate with a window function; this procedure sacrifices resolution;

3. use another method that uses a priori information (e.g. Bayesian and maximum entropy approaches).
Power Spectrum of a Random Process Passed Through a Linear Filter

\[ X(t) \rightarrow h(t) \rightarrow Y(t) \]

\[ Y(t) = h(t) \ast X(t) = \int dt' X(t - t')h(t') \]

Now find the ACF of \( Y(t) \):

\[ \langle Y^*(t)Y(t + \tau) \rangle \equiv \left\langle \int dt' \int dt'' X^*(t - t')X(t + \tau - t'')h^*(t')h(t'') \right\rangle \]

\[ = \int \int dt' dt'' \left\langle X^*(t - t')X(t + \tau - t'') \right\rangle h^*(t')h(t'') \]

We have

\[ R_Y(\tau) = h^*(-\tau) \ast h(\tau) \ast R_X(\tau) \]

and the spectrum is

\[ S_y(\omega) = \tilde{H}^*(\omega)\tilde{H}(\omega)S_x(\omega) \]

Therefore

\[ S_Y(\omega) = |H(\omega)|^2 S_X(\omega) \]
Notes:

1. For linear filters with impulse responses $h(t)$ that have unit area,

$$\int dt \ h(t) \equiv \tilde{H}(0) = 1$$

then

$$\langle Y \rangle = \langle X \rangle$$

and the variance of $Y$ will be less than or equal to the variance of $X$:

$$\sigma_y^2 \leq \sigma_x^2 \quad \text{equality when } h(t) = \delta(t)$$

because

$$\sigma_x^2 - \sigma_y^2 = \langle x^2 \rangle - \langle y^2 \rangle = \frac{1}{2\pi} \int d\omega \ S_x(\omega)[1 - |H(\omega)|^2] \geq 0$$

2. e.g. Low pass filters: $|H(\omega)| \leq \tilde{H}(0) = 1$

Low pass filters that preserve the mean reduce the variance; they smooth the input.

In general,

$$\frac{\sigma_Y}{\langle Y \rangle} \leq \frac{\sigma_X}{\langle X \rangle} \quad \text{for low pass filters}$$
A Third Moment

**Bispectrum:** Let \( \gamma_X(t_1, t_2, t_3) = \langle X(t_1)X(t_2)X(t_3) \rangle \). If stationary to third order, then \( \gamma_x \) depends only on

\[
\begin{align*}
    t_2 - t_2 &\equiv \tau_1 \\
    t_3 - t_1 &\equiv \tau_2 \\
    t_3 - t_2 &\equiv \tau_2 - \tau_1
\end{align*}
\]

For a WSS process

\[
\gamma_x = \gamma_x(\tau_1, \tau_2)
\]

Now Fourier transform (2D) to find the bispectrum:

\[
S(\omega_1, \omega_2) = \int d\tau_1 e^{-i\omega_1 \tau_1} \int d\tau_2 e^{-i\omega_2 \tau_2} \cdot \gamma_X(\tau_1, \tau_2)
\]

It is the distribution of the third moment in frequency space because

\[
\langle x^3(\tau) \rangle = \frac{1}{(2\pi)^2} \iint d\omega_1 \, d\omega_2 \, S(\omega_1, \omega_2)
\]

The bispectrum:

1. contains phase information which the power spectrum does not
2. is useful in studies of **time asymmetries** of a random process
3. used in studying ocean waves
4. image formation in speckle interferometry (recall similarity of **power spectra** (second moments) of different kinds of music).