11. Wide-sense stationary processes

- definition
- properties of correlation function
- power spectral density (Wiener – Khinchin theorem)
- cross-correlation
- cross spectrum
- linear system with random inputs
- designing optimal linear filters

Definition

the second-order joint cdf of an RP $X(t)$ is

$$F_{X(t_1),X(t_2)}(x_1, x_2)$$

(joint cdf of two different times)

we say $X(t)$ is wide-sense (or second-order) stationary if

$$F_{X(t_1),X(t_2)}(x_1, x_2) = F_{X(t_1+\tau),X(t_2+\tau)}(x_1, x_2)$$

the second-order joint cdf do not change for all t_1, t_2 and for all τ

results:

- $\mathbf{E}[X(t)] = m$ (mean is constant)
- $R(t_1, t_2) = R(t_2 - t_1)$ (correlation depends only on the time gap)

Properties of correlation function

let $X(t)$ be a wide-sense scalar real-valued RP with correlation function $R(t_1, t_2)$

- since $R(t_1, t_2)$ depends only on $t_1 - t_2$, we usually write $R(\tau)$ where $\tau = t_1 - t_2$ instead
- $R(0) = \mathbf{E}[X(t)^2]$ for all t
- $R(\tau)$ is an even function of τ

$$R(\tau) \triangleq \mathbf{E}[X(t+\tau)X(t)] = \mathbf{E}[X(t)X(t+\tau)] \triangleq R(-\tau)$$

- $|R(\tau)| \leq R(0)$ (correlation is maximum at lag zero)

$$\mathbf{E}[(X(t+\tau)-X(t))^2] \geq 0 \Longrightarrow 2\mathbf{E}[X(t+\tau)X(t)] \leq \mathbf{E}[X(t+\tau)^2]+\mathbf{E}[X(t)^2]$$

- the autocorrelation is a measure of **rate of change** of a WSS

$$
\begin{aligned}
&P(|X(t+\tau) - X(t)| > \epsilon) \\
&= P(|X(t+\tau) - X(t)|^2 > \epsilon^2) \leq \frac{\mathbf{E}[|X(t+\tau) - X(t)|^2]}{\epsilon^2} = \frac{2(R(0) - R(\tau))}{\epsilon^2}
\end{aligned}
$$

- for complex-valued RP, $R(\tau) = R^*(-\tau)$

$$
\begin{aligned}
R(\tau) &\triangleq \mathbf{E}[X(t+\tau)X^*(t)] \\
&= \mathbf{E}[X(t)X^*(t-\tau)] \\
&= \overline{\mathbf{E}[X(t-\tau)X^*(t)]} \\
&\triangleq R^*(-\tau)
\end{aligned}
$$

- if $R(0) = R(T)$ for some T then $R(\tau)$ is **periodic** with period T and $X(t)$ is **mean square periodic**, *i.e.*,

$$
\mathbf{E}[(X(t+T) - X(t))^2] = 0
$$

$R(\tau)$ is periodic because

$$
\begin{aligned}
&(R(\tau+T) - R(\tau))^2 \\
&\quad = \{\mathbf{E}[(X+t+\tau+T) - X(t+\tau))X(t)]\}^2 \\
&\quad \leq \mathbf{E}[(X(t+\tau+T) - X(t+\tau))^2]\mathbf{E}[X^2(t)] \quad \text{(Cauchy-Schwarz ineq)} \\
&\quad = 2[R(0) - R(T)]R(0) = 0
\end{aligned}
$$

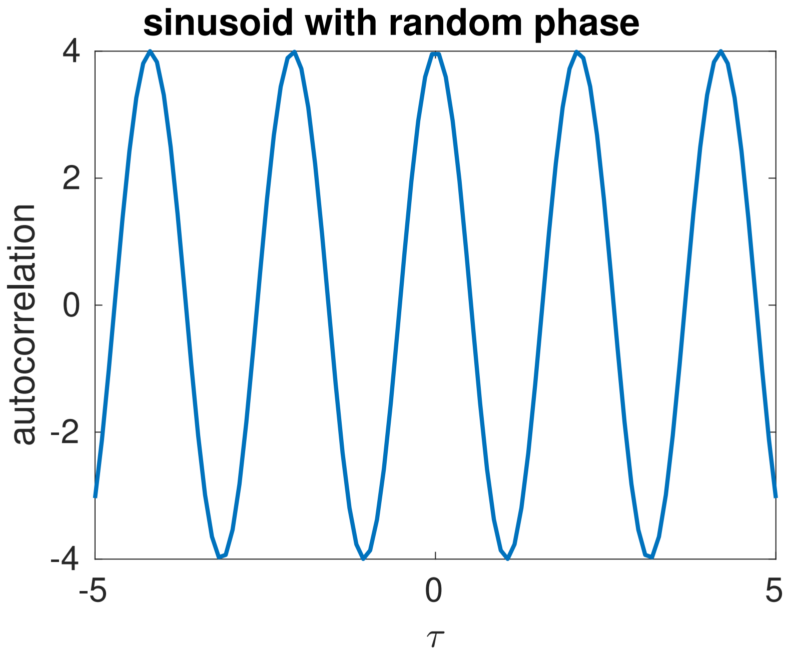
$X(t)$ is mean square periodic because

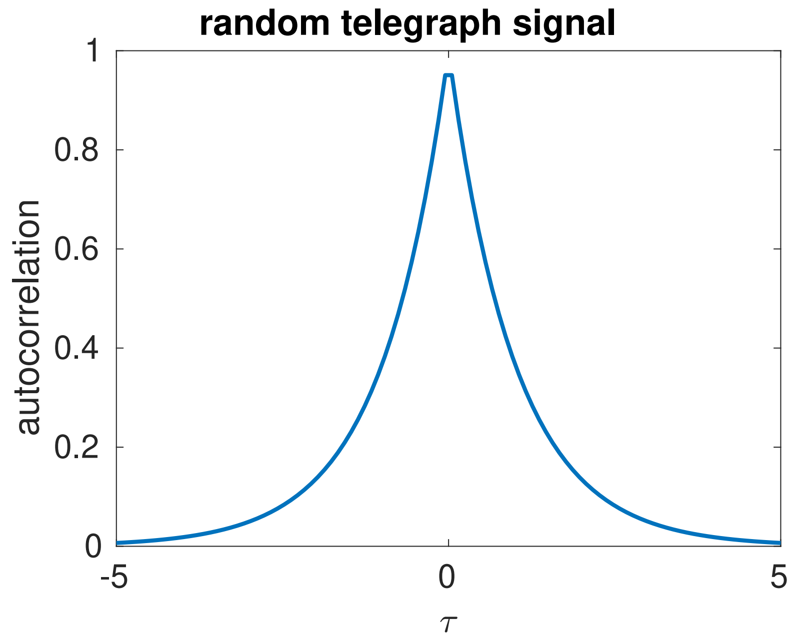
$$\mathbf{E}[(X(t+T) - X(t))^2] = 2(R(0) - R(T)) = 0$$

- let $X(t) = m + Y(t)$ where $Y(t)$ is a zero-mean process

$$R_x(\tau) = m^2 + R_y(\tau)$$

examples:





- sinusoide with random phase: $R(\tau) = \frac{A^2}{2}\cos(\omega\tau)$
- random telegraph signal: $R(\tau) = e^{-2\alpha|\tau|}$

Nonnegativity of correlation function

let $X(t)$ be a real-valued WSS and let $Z = (X(t_1), X(t_2), \ldots, X(t_N))$

the correlation matrix of Z, which is always nonnegative, takes the form

$$\mathbf{R} = \begin{bmatrix} R(0) & R(t_1 - t_2) & \cdots & R(t_1 - t_N) \\ R(t_2 - t_1) & R(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & R(t_{N-1} - t_N) \\ R(t_N - t_1) & \cdots & R(t_N - t_{N-1}) & R(0) \end{bmatrix} \quad \text{(symmetric)}$$

since by assumption,

- $X(t)$ can be either CT or DT random process
- N (the number of time samples) can be any number
- the choice of t_k's are arbitrary

we then conclude that $\mathbf{R} \succeq 0$ holds for all sizes of $\mathbf{R}$ $(N = 1, 2, \ldots)$

the nonnegativity of $\mathbf{R}$ can also be checked from the definition:

$$a^T \mathbf{R} a \geq 0, \quad \text{for all} \quad a = (a_1, a_2, \ldots, a_N)$$

which follows from

$$\begin{aligned} \sum_{i=1}^{N} \sum_{j=1}^{N} a_i^T R(t_i - t_j) a_j &= \sum_{i}^{N} \sum_{j}^{N} \mathbf{E}[a_i^T X(t_i) X(t_j)^T a_j] \\ &= \mathbf{E}\left[\left(\sum_{i=1}^{N} a_i^T X(t_i) \right)^2 \right] \geq 0 \end{aligned}$$

important note: the value of $R(t)$ at some fixed t can be negative !

example: $R(\tau) = e^{-|\tau|/2}$ and let $t = (t_1, t_2, \ldots, t_5)$

```
k=5; t = abs(randn(k,1)); t = sort(t); % t = (t1,...,tk)
R = zeros(k);
for i=1:k
    for j=1:k
        R(i,j) = exp(-0.5*abs(t(i)-t(j)));
    end
end
R =
    1.0000    0.6021    0.4952    0.4823
    0.6021    1.0000    0.8224    0.8011
    0.4952    0.8224    1.0000    0.9740
    0.4823    0.8011    0.9740    1.0000

eig(R) =
0.0252    0.2093    0.6416    3.1238
```

showing that $\mathbf{R} \succeq 0$ (try with any k)

Block toeplitz structure of correlation matrix

CT process: if $X(t)$ are sampled as $Z = (X(t_1), X(t_2), \dots, X(t_N))$ where

$$t_{i+1} - t_i = \text{constant} = s \quad , i = 1, \dots, N-1$$

(times have **constant spacing**, $s > 0$ and no need to be an integer)

we see that $\mathbf{R} = \mathbf{E}[ZZ^T]$ has a symmetric **block toeplitz structure**

$$\mathbf{R} = \begin{bmatrix} R(0) & R(-s) & \cdots & R(-(N-1)s) \\ R(s) & R(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & R(-s) \\ R((N-1)s) & \cdots & R(s) & R(0) \end{bmatrix} \quad \text{(symmetric)}$$

if $X(t)$ is WSS then $\mathbf{R} \succeq 0$ for any integer N and any $s > 0$

example: $R(\tau) = e^{-|\tau|/2}$

```
>> t=0:0.5:2; R = exp(-0.5*abs(t)); T = toeplitz(R)
R =

    1.0000    0.7788    0.6065    0.4724    0.3679

T =

    1.0000    0.7788    0.6065    0.4724    0.3679
    0.7788    1.0000    0.7788    0.6065    0.4724
    0.6065    0.7788    1.0000    0.7788    0.6065
    0.4724    0.6065    0.7788    1.0000    0.7788
    0.3679    0.4724    0.6065    0.7788    1.0000

eig(T) =

    0.1366    0.1839    0.3225    0.8416    3.5154
```

DT process: time indices are integers, so $Z = (X(1), X(2), \ldots, X(N))$

times also have **constant spacing**

$\mathbf{R} = \mathbf{E}[ZZ^T]$ also has a symmetric **block toeplitz structure**

$$\begin{bmatrix} R(0) & R(-1) & \cdots & R(1-N) \\ R(1) & R(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & R(-1) \\ R(N-1) & \cdots & R(1) & R(0) \end{bmatrix}$$

if $X(t)$ is WSS then $\mathbf{R} \succeq 0$ for any positive integer N

example: $R(\tau) = \cos(\tau)$

```
>> t=0:2; R = cos(t); T = toeplitz(R)

R =
    1.0000    0.5403   -0.4161

T =
    1.0000    0.5403   -0.4161
    0.5403    1.0000    0.5403
   -0.4161    0.5403    1.0000

eig(T) =

    0.0000
    1.4161
    1.5839
```

$R(\tau)$ at some τ can be negative !

Power spectral density

Wiener-Khinchin Theorem: if a process is wide-sense stationary, the autocorrelation function and the power spectral density form a Fourier transform pair:

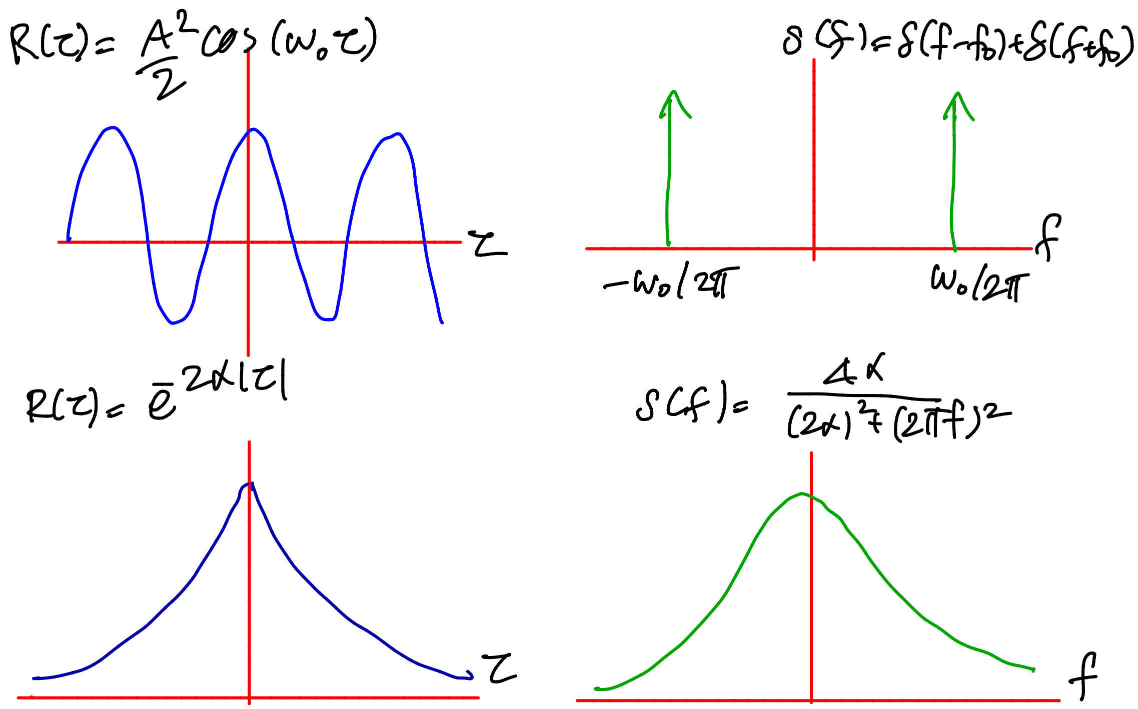
$$S(\omega) = \int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau \qquad \text{continuous-time FT}$$

$$S(\omega) = \sum_{k=-\infty}^{\infty} R(k) e^{-i\omega k} \qquad \text{discrete-time FT}$$

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega \qquad \text{continuous-time IFT}$$

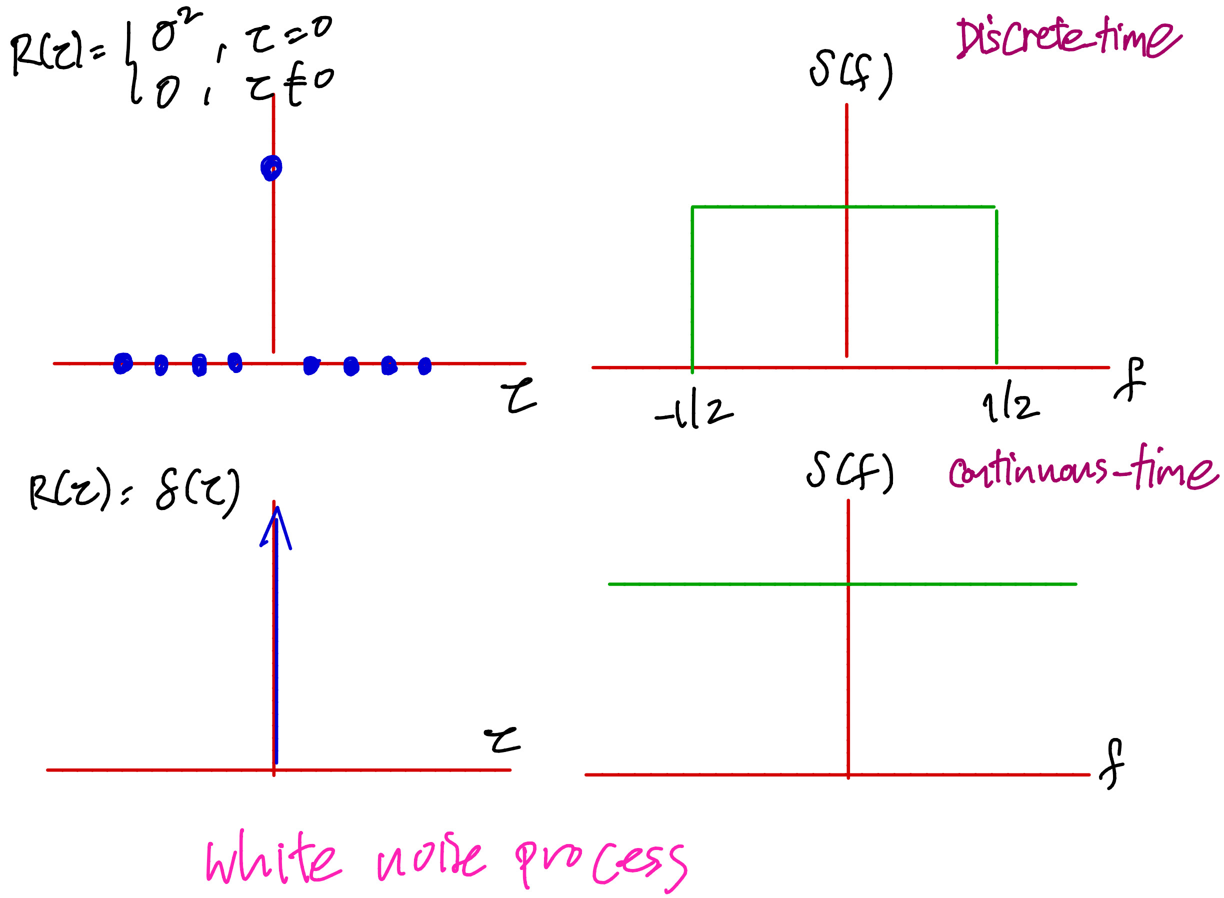
$$R(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega\tau} S(\omega) d\omega \qquad \text{discrete-time IFT}$$

$S(\omega)$ indicates a density function for average power versus frequency examples: sinusoid with random phase and random telegraph



- (left) $X(t) = A\sin(\omega_0 t + \phi)$ and $\phi \sim \mathcal{U}(-\pi, \pi)$
- (right) $X(t)$ is random telegraph signal

examples: white noise process



- (left) DT white noise process has a spectrum as a rectangular window
- (right) CT white noise process has a flat spectrum

Spectrum of a moving average process

let $X(n)$ be a DT white noise process with variance σ^2

$$Y(n) = X(n) + \alpha X(n-1), \quad \alpha \in \mathbf{R}$$

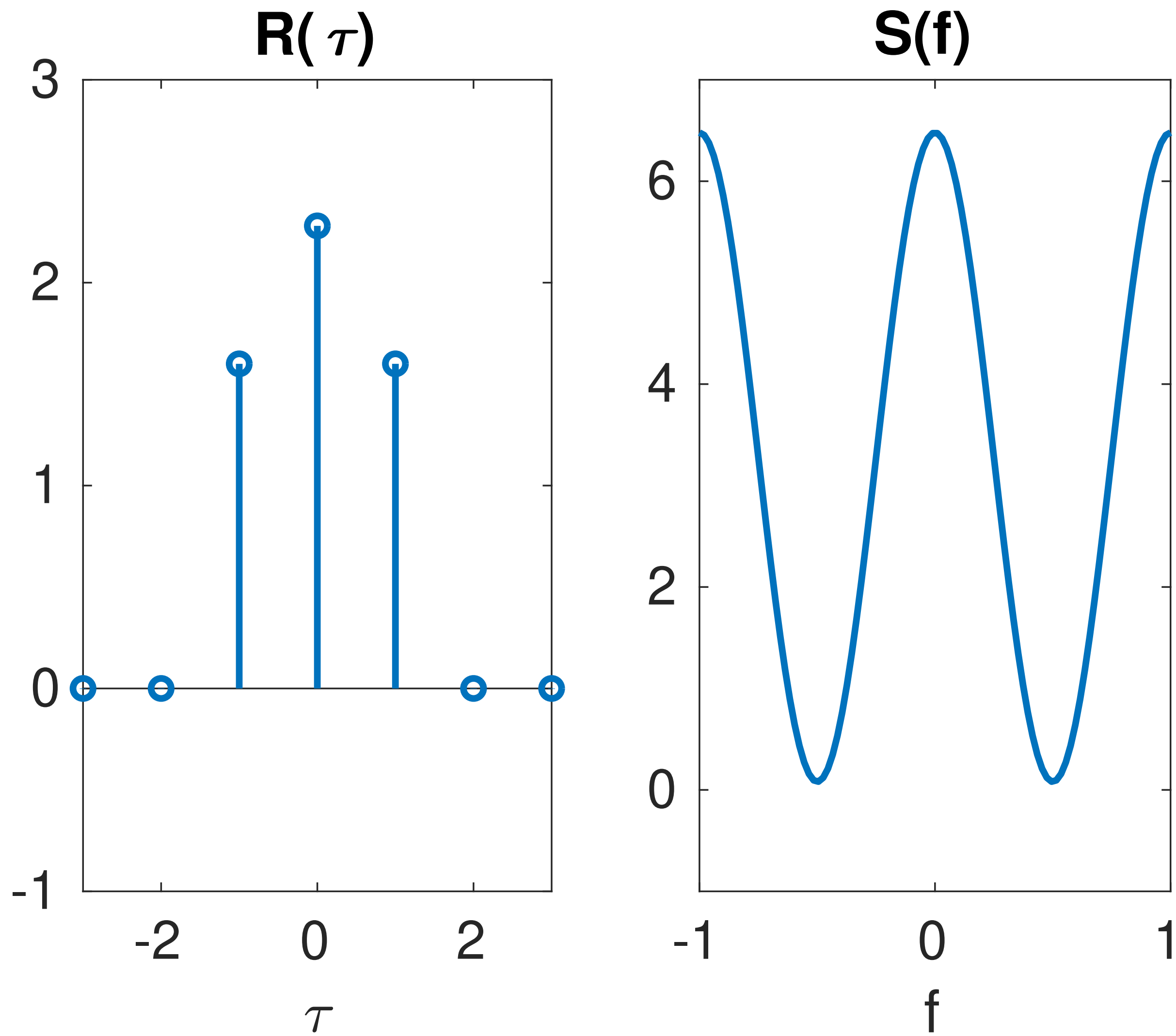
then $Y(n)$ is an RP with autocorrelation function

$$R_Y(\tau) = \begin{cases} (1+\alpha^2\sigma^2), & \tau = 0, \\ \alpha\sigma^2, & |\tau| = 1, \\ 0, & \text{otherwise} \end{cases}$$

the spectrum of DT process (is periodic in $f \in [-1/2, 1/2]$) is given by

$$\begin{aligned} S(f) &= \sum_{k=-\infty}^{\infty} R_Y(k) e^{-i2\pi f k} \\ &= (1+\alpha^2\sigma^2) + \alpha\sigma^2(e^{i2\pi f} + e^{-i2\pi f}) \\ &= \sigma^2(1+\alpha^2 + 2\alpha\cos(2\pi f)) \end{aligned}$$

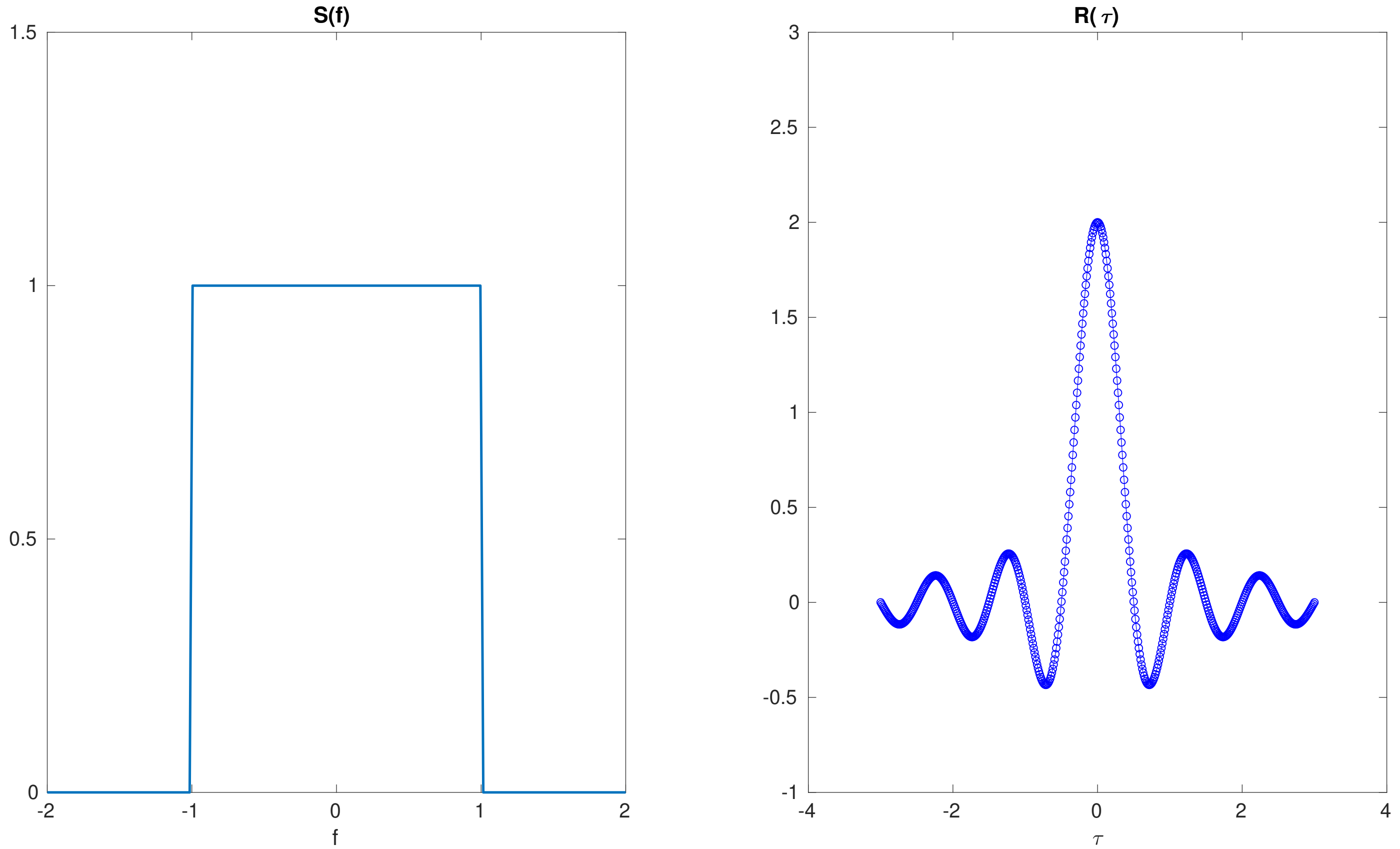
examples: moving average process with $\sigma^2 = 2$ and $\alpha = 0.8$



spectrum is periodic in $f \in [-1/2, 1/2]$

Band-limited white noise

given a (white) process whose spectrum is *flat* in the range $-B \le f \le B$



the magnitude of the spectrum is $N/2$

what will the (continuous-valued) process look like ?

autocorrelation function is obtained from IFT

$$\begin{aligned}
R(\tau) &= (N/2) \int_{-B}^{B} e^{i2\pi f\tau} df \\
&= \frac{N}{2} \cdot \frac{e^{i2\pi B\tau} - e^{-i2\pi B\tau}}{i2\pi\tau} \\
&= \frac{N \sin(2\pi B\tau)}{2\pi\tau} = NB\mathrm{sinc}(2\pi B\tau)
\end{aligned}$$

- $X(t)$ and $X(t+\tau)$ are uncorrelated at $\tau = \pm k/2B$ for $k = 1, 2, \ldots$
- if $B \to \infty$, the band-limited white noise becomes a white noise

$$S(f) = \frac{N}{2}, \quad \forall f, \quad R(\tau) = \frac{N}{2}\delta(\tau)$$

Properties of power spectral density

consider real-valued RPs, so $R(\tau)$ is real-valued

- $S(\omega)$ is real-valued and even function ($\because R(\tau)$ is real and even)
- $R(0)$ indicates the **average power**

$$R(0) = \mathbf{E}[X(t)^2] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)d\omega$$

- $S(\omega) \geq 0$ for all ω and for all $\omega_2 \geq \omega_1$

$$\frac{1}{2\pi}\int_{\omega_1}^{\omega_2} S(\omega)d\omega$$

is the average power in the frequency band (ω_2, ω_1)

(see proof in Chapter 9 of H. Stark)

Power spectral density as a time average

let $X[0], X[1], \ldots, X[N-1]$ be N observations from DT WSS process

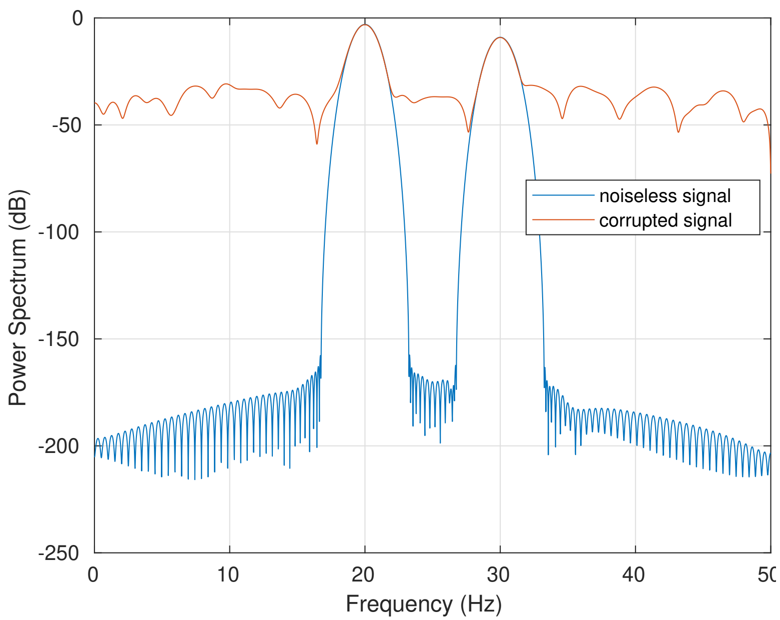
discrete Fourier transform of the time-domain sequence is

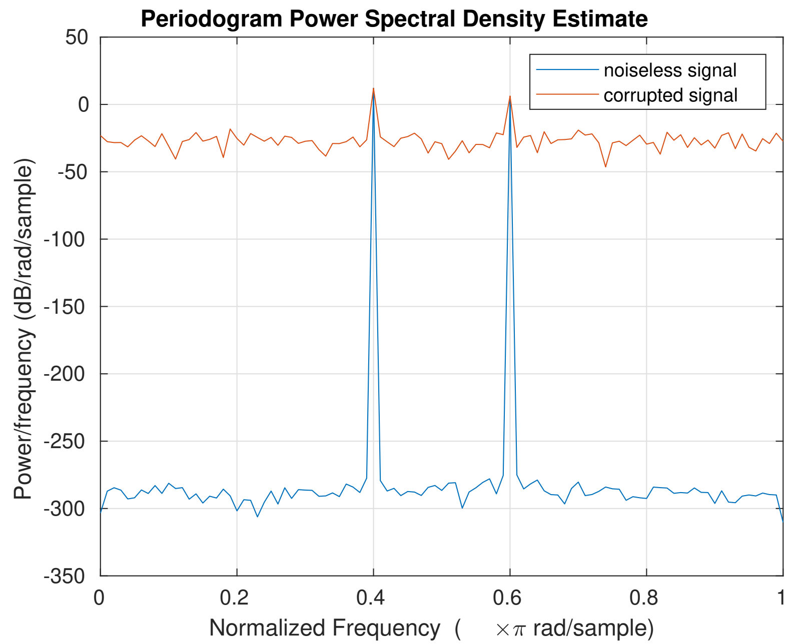
$$\tilde{X}[k] = \sum_{n=0}^{N-1} X[n] e^{-\frac{i2\pi}{N}kn}, \quad k = 0, 1, \ldots, N-1$$

- $\tilde{X}[k]$ is a complex-valued sequence describing DT Fourier transform with only *discrete frequency points*
- $\tilde{X}[k]$ is a measure of *energy* at frequency $2\pi k/N$
- an estimate of *power* at a frequency is then

$$\tilde{S}(k) = \frac{1}{N} |\tilde{X}[k]|^2$$

and is called **periodogram estimate** for the power spectral density example: $X(t) = \sin(40\pi t) + 0.5\sin(60\pi t)$

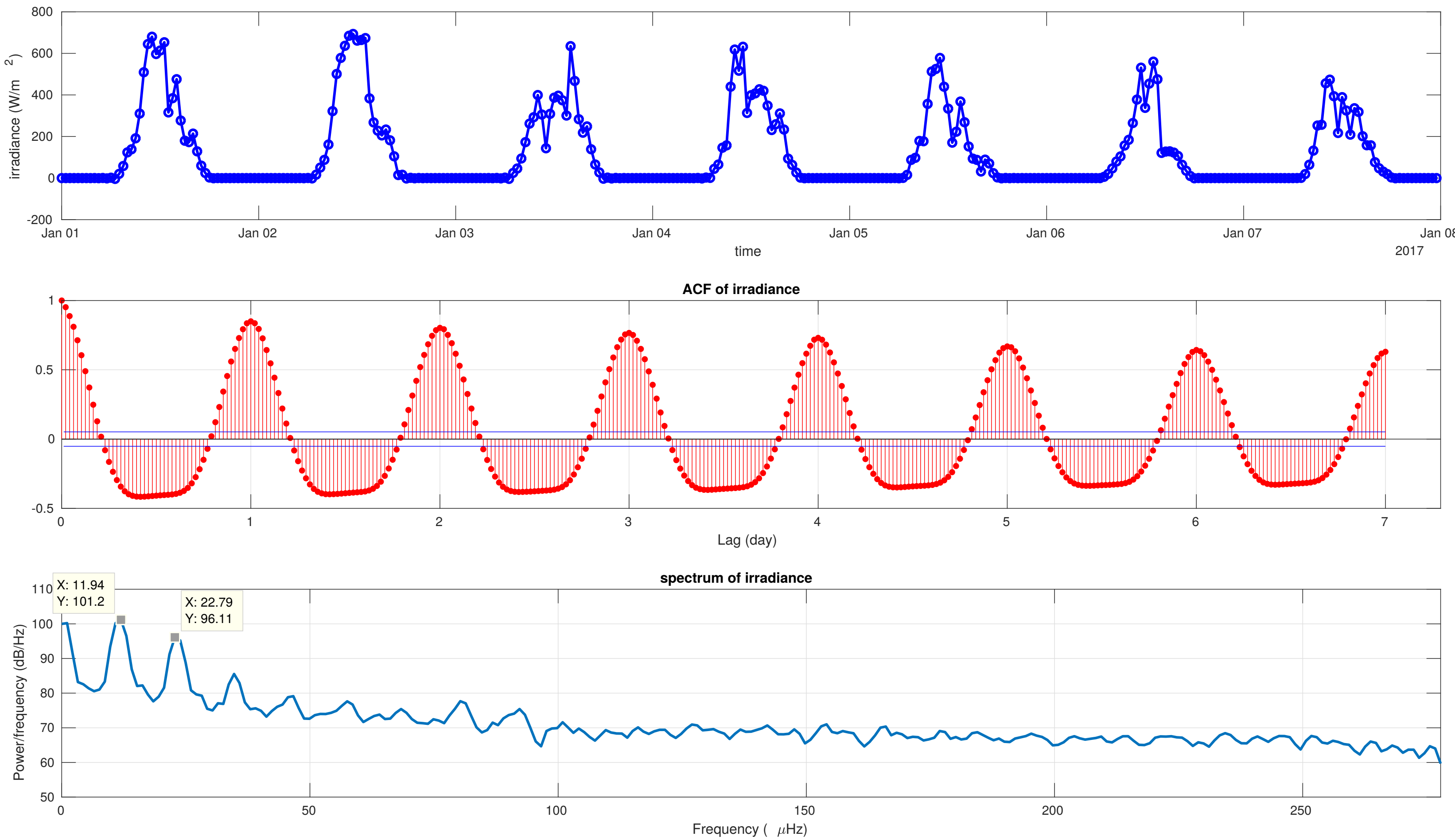




- signal has frequency components at 20 and 30 Hz
- peaks at 20 and 30 Hz are clearly seen
- when signal is corrupted by noise, spectrum peaks can be less distinct
- the plots are done using `pspectrum` and `periodogram` in MATLAB

Frequency analysis of solar irradiance

data are irradiance with sampling period of $T = 30$ min



- ACF is a normalized autocorrelation function (by $R(0)$) and appears to be periodic
- spectral density appears to have three peaks corresponding to $0, 12, 24$ μHz
- the frequencies of $12, 24$ μHz correspond to the periods of one day and half day respectively
- ACF and spectral density are computed by `autocorr` and `pwelch` commands in MATLAB
- more details on spectrum estimation methods can be further studied in signal processing

Cross correlation and cross spectrum

cross correlation between processes $X(t)$ and $Y(t)$ is defined as

$$R_{XY}(\tau) = \mathbf{E}[X(t+\tau)Y(t)]$$

cross-power spectral density between $X(t)$ and $Y(t)$ is defined as

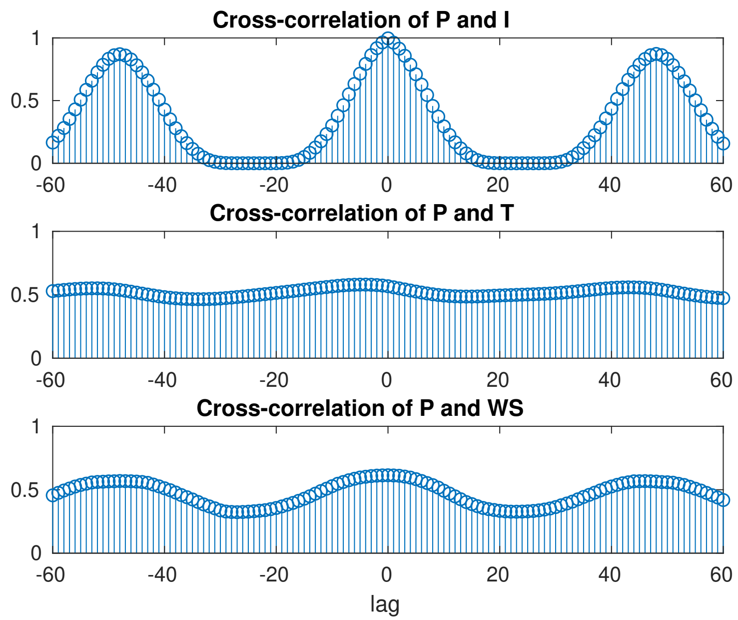
$$S_{XY}(\omega) = \int_{-\infty}^{\infty} e^{-\mathrm{i}\omega\tau} R_{XY}(\tau) d\tau$$

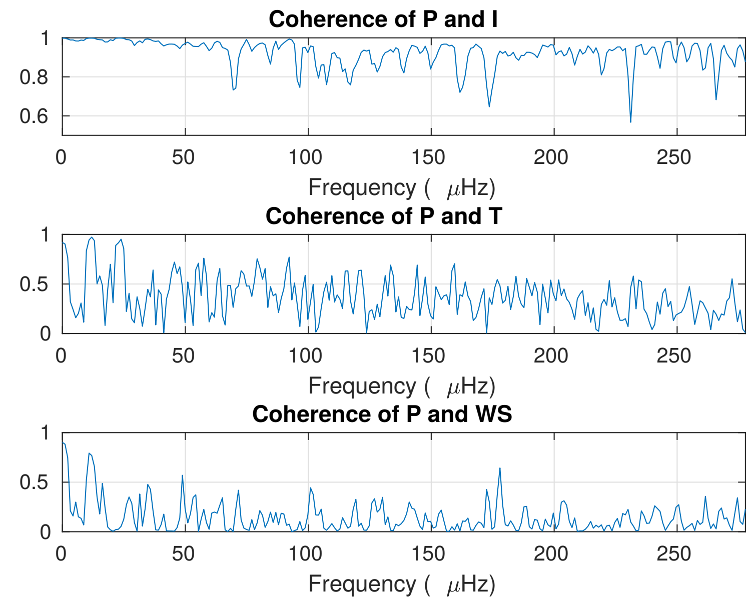
properties:

- $S_{XY}(\omega)$ is complex-valued in general, even $X(t)$ and $Y(t)$ are real
- $R_{YX}(\tau) = R_{XY}(-\tau)$
- $S_{YX}(\omega) = S_{XY}(-\omega)$

Examples from solar variables

solar power (P), solar irradiance (I), temperature (T), wind speed (WS)

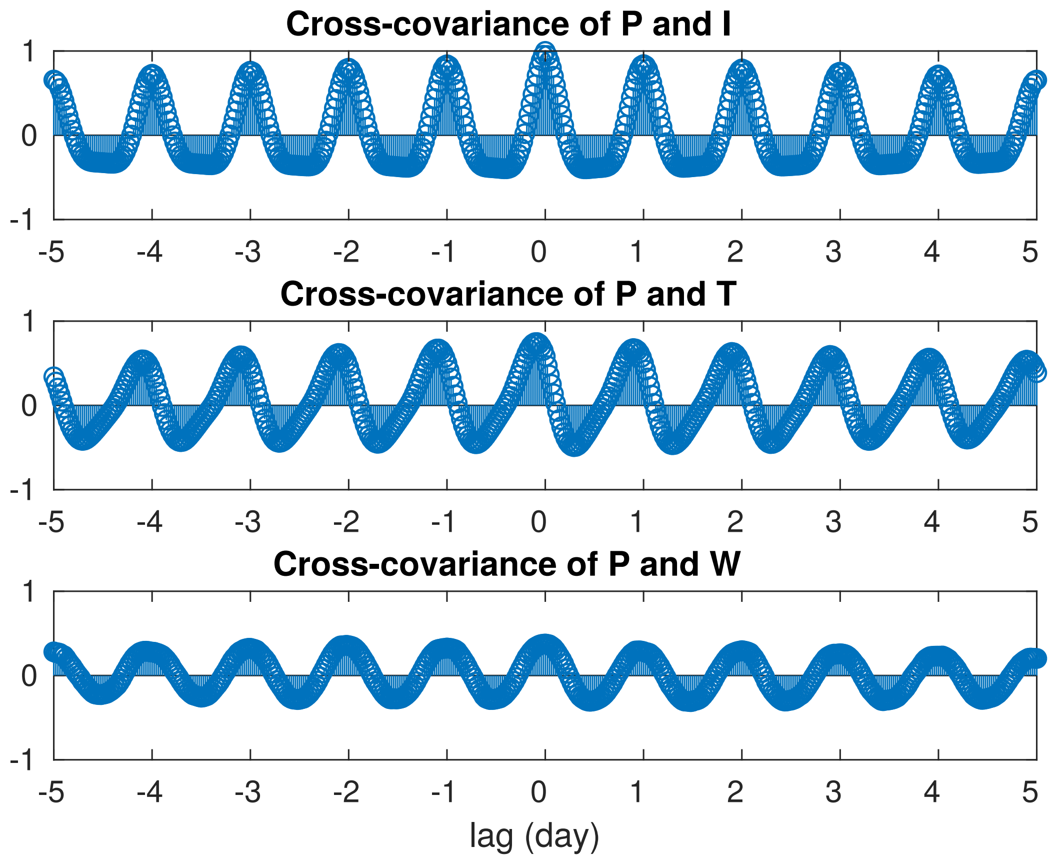




- (normalized) cross correlations are computed by `xcorr` in MATLAB
- (normalized) coherence functions are computed by `mscohere`:

$$C_{xy}(f) = \frac{|S_{xy}(f)|^2}{S_x(f)S_y(f)}$$

cross covariance function:



- P and I are highly correlated while P and WS are least correlated
- cross covariance functions are almost periodic (daily cycle) with slightly decaying envelopes

Extended definitions

extension: let $X(t)$ be a *complex-valued vector* random process

- denote $*$ Hermittian transpose, $i.e.$, $X^* = \overline{X}^T$
- correlation function: $R(\tau) = \mathbf{E}[X(t+\tau)X(t)^*]$
- covariance function: $C(\tau) = R(\tau) - \mu\mu^*$
- $R_{YX}(\tau) = R^*_{XY}(-\tau)$
- $S_{YX}(\omega) = S^*_{XY}(-\omega)$
- $S(\omega)$ is self-adjoint, $i.e.$, $S(\omega) = S^*(\omega)$ and $S(\omega) \succeq 0$

(cross) correlation and (cross) spectral density functions are *matrices*

Theorems on correlation function and spectrum

Theorem 1: a necessary and sufficient condition for $R(\tau)$ to be a correlation function of a WSS is that it is positive semidefinite

- proof of sufficiency part: if $R(\tau)$ is positive semidefinite then there exists a WSS whose correlaction function is $R(\tau)$
  - if $R(\tau)$ is psdf then its Fourier transform is positive semidefinite (a proof is not obvious)
  - let us call $S(\omega) = \mathcal{F}(R(\tau)) \succeq 0$
  - by spectral factorization theorem, there exists a stable filter $H(\omega)$ such that $S(\omega) = H(\omega)H^*(\omega)$ – more advanced topic
  - the existence of a WSS is given by applying a white noise to the filter $H(\omega)$ – the topic we will learn next on page 11-38
- proof of necessity part: if a process is WSS then $R(\tau)$ is positive semidefinite – shown on page 11-7

Theorem 2: let $S(\omega)$ be a self-adjoint and nonnegative matrix and

$$\int_{-\infty}^{\infty} \mathbf{tr}(S(\omega))d\omega < \infty$$

then its inverse Fourier transform:

$$R(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega t}S(\omega)d\omega$$

is **nonnegative**, *i.e.*, $\sum_{j=1}^{N}\sum_{k=1}^{N} a_j^* R(t_j - t_k)a_k \geq 0$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}^* \begin{bmatrix} R(0) & R(t_1 - t_2) & \cdots & R(t_1 - t_N) \\ R(t_2 - t_1) & R(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & R(t_{N-1} - t_N) \\ R(t_N - t_1) & \cdots & R(t_N - t_{N-1}) & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \succeq 0$$

proof: let us consider $N = 3$ case (can be extended easily)

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}^* \begin{bmatrix} R(0) & R(t_1 - t_2) & R(t_1 - t_3) \\ R(t_2 - t_1) & R(0) & R(t_2 - t_3) \\ R(t_3 - t_1) & R(t_3 - t_2) & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$= \int_{-\infty}^{\infty} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}^* \begin{bmatrix} e^{i\omega(t_1-t_1)}S(\omega) & e^{i\omega(t_1-t_2)}S(\omega) & e^{i\omega(t_1-t_3)}S(\omega) \\ e^{i\omega(t_2-t_1)}S(\omega) & e^{i\omega(t_2-t_2)}S(\omega) & e^{i\omega(t_2-t_3)}S(\omega) \\ e^{i\omega(t_3-t_1)}S(\omega) & e^{i\omega(t_3-t_2)}S(\omega) & e^{i\omega(t_3-t_3)}S(\omega) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} d\omega$$

$$= \int_{-\infty}^{\infty} \begin{bmatrix} e^{-i\omega t_1}a_1 \\ e^{-i\omega t_2}a_2 \\ e^{-i\omega t_3}a_3 \end{bmatrix}^* \begin{bmatrix} S^{1/2}(\omega) \\ S^{1/2}(\omega) \\ S^{1/2}(\omega) \end{bmatrix} \begin{bmatrix} S^{1/2}(\omega) & S^{1/2}(\omega) & S^{1/2}(\omega) \end{bmatrix} \begin{bmatrix} e^{-i\omega t_1}a_1 \\ e^{-i\omega t_2}a_2 \\ e^{-i\omega t_3}a_3 \end{bmatrix} d\omega$$

$$\triangleq \int_{-\infty}^{\infty} Y^*(\omega)Y(\omega)d\omega \succeq 0$$

because the integrand is nonnegative definite for all ω

(we have used the fact that $S(\omega) \succeq 0$ and has a square root)

Theorem 3: let $R(t)$ be a continuous correlation matrix function such that

$$\int_{-\infty}^{\infty} |R_{ij}(t)| dt < \infty, \quad \forall i, j$$

then the spectral density matrix

$$S(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} R(t) dt$$

is self-adjoint and positive semidefinite

- matrix case: proof by Balakrishnan, Introduction to Random Process in Engineering, page 79
- scalar case: proof by Starks and Woods, page 607 (need to learn the topic on page 11-38 first)

simple proof (from Starks): let $\omega_2 > \omega_1$, define a filter transfer function

$$H(\omega) = 1, \quad \omega \in (\omega_1, \omega_2), \qquad H(\omega) = 0, \quad \text{otherwise}$$

let $X(t)$ and $Y(t)$ be input/output to this filter, then

$$S_{YY}(\omega) = S_{XX}(\omega), \quad \omega \in (\omega_1, \omega_2), \qquad 0, \quad \text{else}$$

since $\mathbf{E}[Y(t)^2] = R_y(0)$ and it is nonnegative, it follows that

$$R_y(0) = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} S_x(\omega) d\omega \geq 0$$

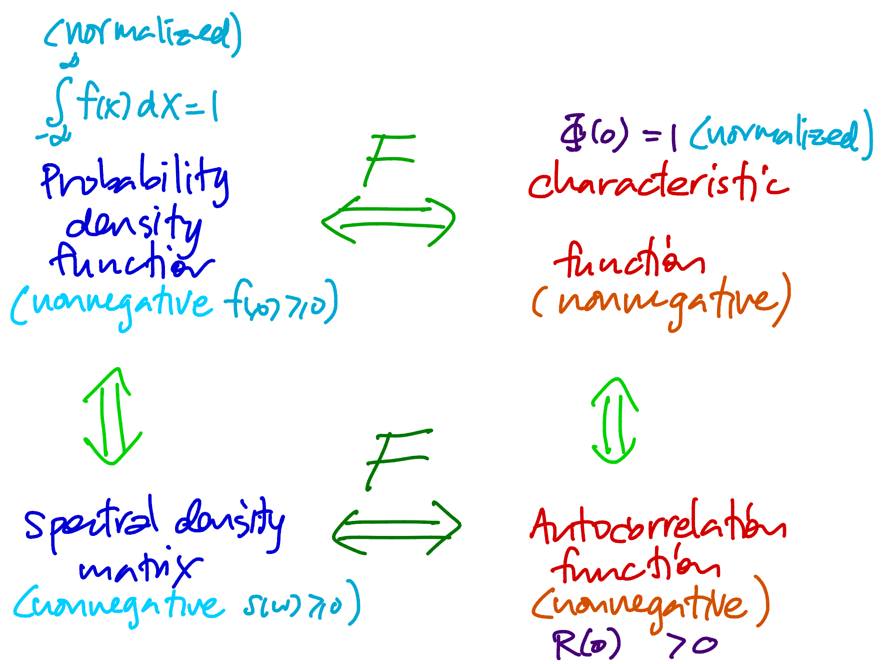
this must holds for any $\omega_2 > \omega_1$

hence, choosing $\omega_2 \approx \omega_1$ we must have $S_x(\omega) \geq 0$ — the power spectral density must be nonnegative

conclusion: a function $R(\tau)$ is nonnegative if and only if

it has a nonnegative Fourier transform

- a valid spectral density function therefore can be checked by its nonnegativity and it is easier than checking the nonnegativity condition of $R(\tau)$
- analogy for probability density function



Linear system with random inputs

consider a linear system with input and output relationship through

$$y = Hx$$

which represents many applications (filter, transformation of signals, etc.)

questions regarding this setting:

- if x is a random signal, how can we explain about randomness of y?
- if x is wide-sense stationary, how about y? under what condition on H?
- if y is also wide-sense, how about relations between correlation/power spectral density of x and y?

recall the definitions

- **linear system:**

$$H(x_1 + \alpha x_2) = Hx_1 + \alpha Hx_2$$

- **time-invariant system:** it commutes with shift operator

$$Hx(t-T) = y(t-T)$$

(time shift in the input causes the same time shift in the output)

- response of linear time-invariant system: denote h the impulse response

$$y(t) = h(t) * x(t) = \begin{cases} \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau & \text{continous-time} \\ = \sum_{k=-\infty}^{\infty} h(t-k)x(k) & \text{discrete-time} \end{cases}$$

- **stable:** poles of H are in stability region (LHP or inside unit circle)
- **causal system:** response of y at t depends only on *past* values of x

$$\text{impulse response } \; h(t) = 0, \quad \text{for } t < 0$$

Properties of output from LTI system

let $Y = HX$ where H is linear time-invariant system and **stable**

if $X(t)$ is wide-sense stationary then

- $m_Y(t) = H(0)m_X(t)$
- $Y(t)$ is also wide-sense stationary
  (in steady-state sense if $X(t)$ is applied when $t \geq 0$)
- correlations and spectra are given by

time-domain	frequency-domain
$R_{YX}(\tau) = h(\tau) * R_X(\tau)$	$S_{YX}(\omega) = H(\omega)S_X(\omega)$
$R_{XY}(\tau) = R_X(\tau) * h^*(-\tau)$	$S_{XY}(\omega) = S_X(\omega)H^*(\omega)$
$R_Y(\tau) = R_{YX}(\tau) * h^*(-\tau)$	$S_Y(\omega) = S_{YX}(\omega)H^*(\omega)$
$R_Y(\tau) = h(\tau) * R_X(\tau) * h^*(-\tau)$	$S_Y(\omega) = H(\omega)S_X(\omega)H^*(\omega)$

using $\mathcal{F}(f(t) * g(t)) = F(\omega)G(\omega)$ and $\mathcal{F}(f^*(-t)) = F^*(\omega)$

proof of mean of Y: $m_Y(t) = H(0)m_X(t)$

$$\begin{aligned} Y(t) &= \int_{-\infty}^{\infty} h(s)X(t-s)ds \\ \mathbf{E}[Y(t)] &= \int_{-\infty}^{\infty} h(s)\mathbf{E}[X(t-s)]ds \\ &= \int_{-\infty}^{\infty} h(s)ds \cdot m_x \qquad \text{(since } X(t) \text{ is WSS)} \\ &= H(0)m_x \end{aligned}$$

mean of Y is transformed by the DC gain of the system

proof of WSS of Y

$$\begin{aligned}
R_y(t+\tau, t) &= \mathbf{E}[Y(t+\tau)Y(t)^T] \\
&= \mathbf{E}\left[\left(\int_{-\infty}^{\infty} h(\sigma)X(t+\tau-\sigma)ds\right)\left(\int_{-\infty}^{\infty} h(s)X(t-s)ds\right)^T\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\sigma)\mathbf{E}[X(t+\tau-\sigma)X(t-s)^T]h(s)^T d\sigma ds \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\sigma)R_x(\tau+s-\sigma)h(s)^T d\sigma ds \quad (X \text{ is WSS})
\end{aligned}$$

we see that $R_y(t+\tau, t)$ does not depend on t anymore but only on τ

- we have shown that $Y(t)$ has a constant mean and the autocorrelation function depends only on the time gap τ
- hence, $Y(t)$ is also a wide-sense stationary process

proofs of cross-correlation: using $Y(t) = \int_{-\infty}^{\infty} h(\alpha)X(t-\alpha)d\alpha$

- $R_{YX}(\tau) = h(\tau) * R_X(\tau)$

$$\begin{aligned} R_{YX}(\tau) = \mathbf{E}[Y(t)X^*(t-\tau)] &= \int_{-\infty}^{\infty} h(\alpha)\mathbf{E}[X(t-\alpha)X^*(t-\tau)]d\alpha \\ &= \int_{-\infty}^{\infty} h(\alpha)R_X(\tau-\alpha)d\alpha \end{aligned}$$

- $R_Y(\tau) = R_{YX}(\tau) * H^*(-\tau)$

$$\begin{aligned} R_Y(\tau) = \mathbf{E}[Y(t)Y^*(t-\tau)] &= \int_{-\infty}^{\infty} \mathbf{E}[Y(t)X^*(t-(\tau+\alpha))]h^*(\alpha)d\alpha \\ &= \int_{-\infty}^{\infty} R_{YX}(\tau+\alpha)h^*(\alpha)d\alpha = \int_{-\infty}^{\infty} R_{YX}(\tau-\sigma)h^*(-\sigma)d\sigma \end{aligned}$$

Power of output process

the relation $S_Y(\omega) = H(\omega)S_X(\omega)H^*(\omega)$ reduces to

$$S_Y(\omega) = |H(\omega)|^2 S_X(\omega)$$

for *scalar* processes $X(t)$ and $Y(t)$

- average power of the output depends on the input power at that frequency multiplied by power gain at the same frequency
- we call $|H(\omega)|^2$ the **power spectral density (psd) transfer function**

this relation gives a procedure to estimate $H(\omega)$ when signals $X(t)$ and $Y(t)$ can be observed

example: a random telegraph signal with transition rate α is passed thru an RC filter with

$$H(s) = \frac{\tau}{s+\tau}, \quad \tau = 1/RC$$

question: find psd and autocorrelation of the output

random telegraph signal has the spectrum: $S_x(f) = \dfrac{4\alpha}{4\alpha + 4\pi^2 f^2}$

from $S_y(f) = |H(f)|^2 S_x(f)$ and $R_y(t) = \mathcal{F}^{-1}[S_y(f)]$

$$S_y(f) = \left(\frac{\tau^2}{\tau^2 + 4\pi^2 f^2}\right) \frac{4\alpha}{4\alpha + 4\pi^2 f^2}$$

$$= \frac{4\alpha\tau^2}{\tau^2 - 4\alpha^2} \left\{ \frac{1}{4\alpha^2 + 4\pi^2 f^2} - \frac{1}{\tau^2 + 4\pi^2 f^2} \right\}$$

$$R_y(t) = \frac{1}{\tau^2 - 4\alpha^2} \left(\tau^2 e^{-2\alpha|t|} - 2\alpha\tau e^{-\tau|t|} \right)$$

(we have used $\mathcal{F}[e^{-at}] = 2a/(a^2 + \omega^2)$) and $\omega = 2\pi f$

example: spectral density of AR process

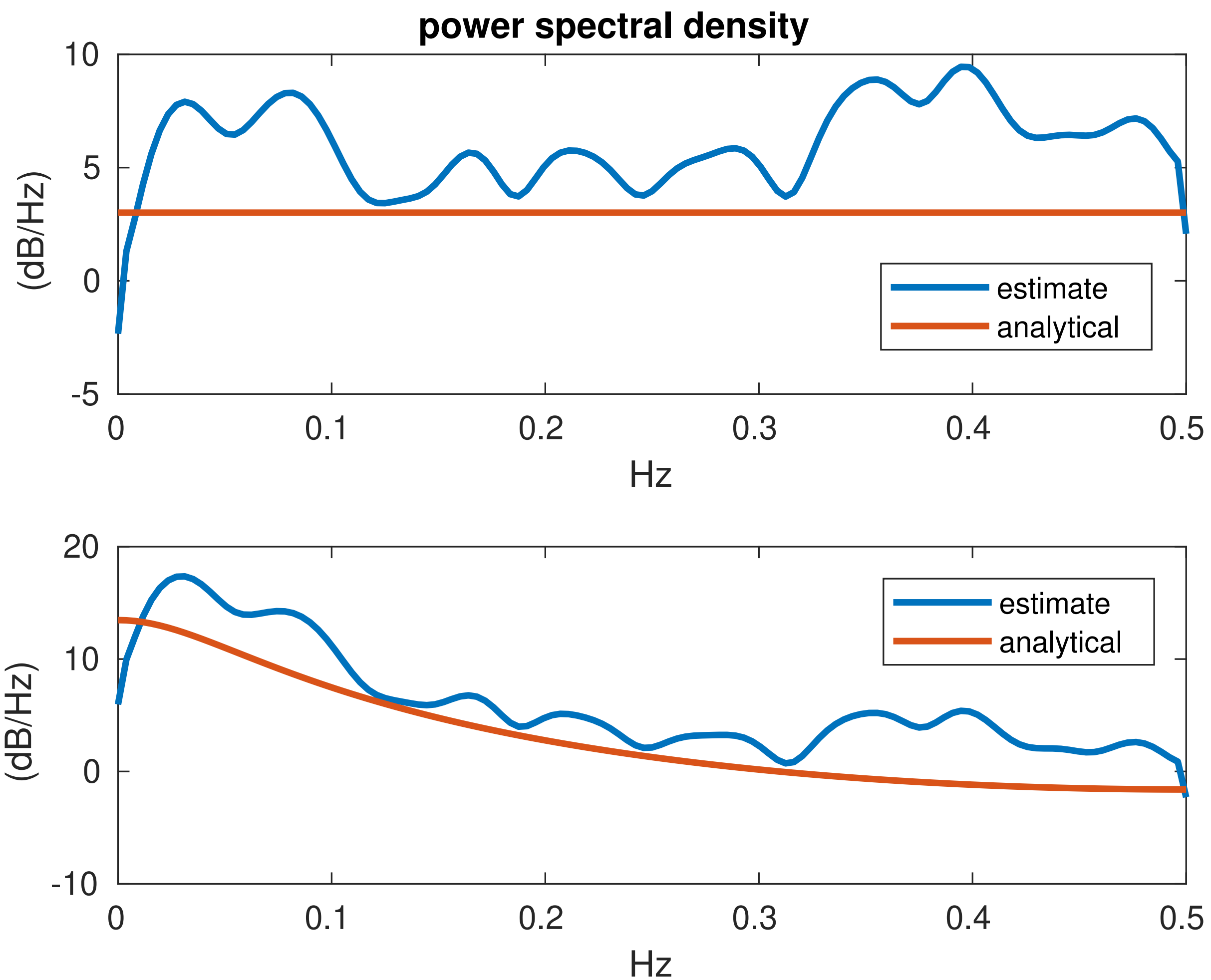
$$Y(n) = aY(n-1) + X(n)$$

$X(n)$ is i.i.d white noise with variance of σ^2

- $H(z) = \frac{1}{1-az^{-1}}$ or $H(e^{i\omega}) = \frac{1}{1-ae^{-i\omega}}$
- spectral density is obtained by

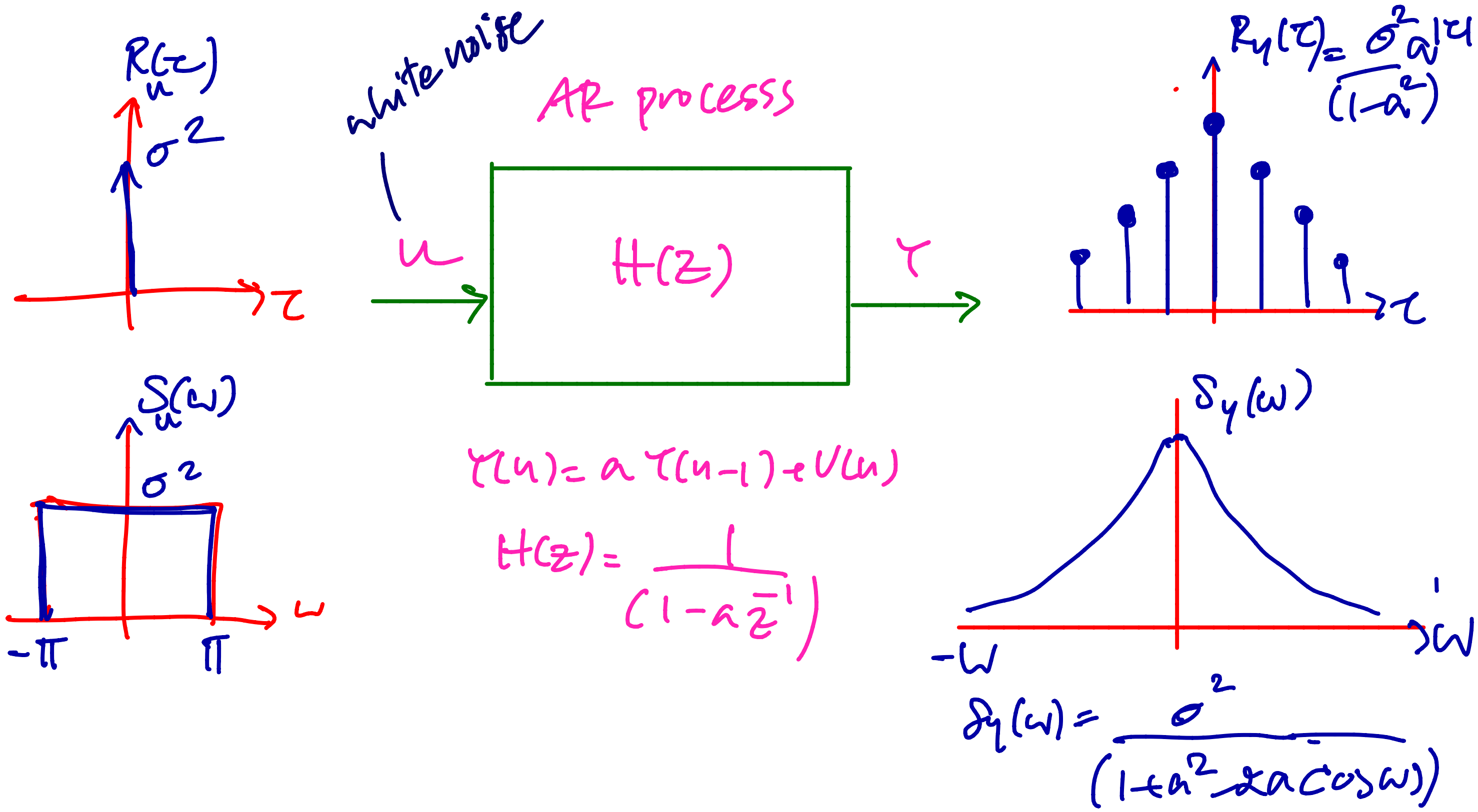
$$S_y(\omega) = |H(\omega)|^2 S_x(\omega) = \frac{\sigma^2}{(1-ae^{-i\omega})(1-ae^{i\omega})}$$
$$= \frac{\sigma^2}{1+a^2-2a\cos(\omega)}$$

spectral density of AR process: $a = 0.7$ and $\sigma^2 = 2$

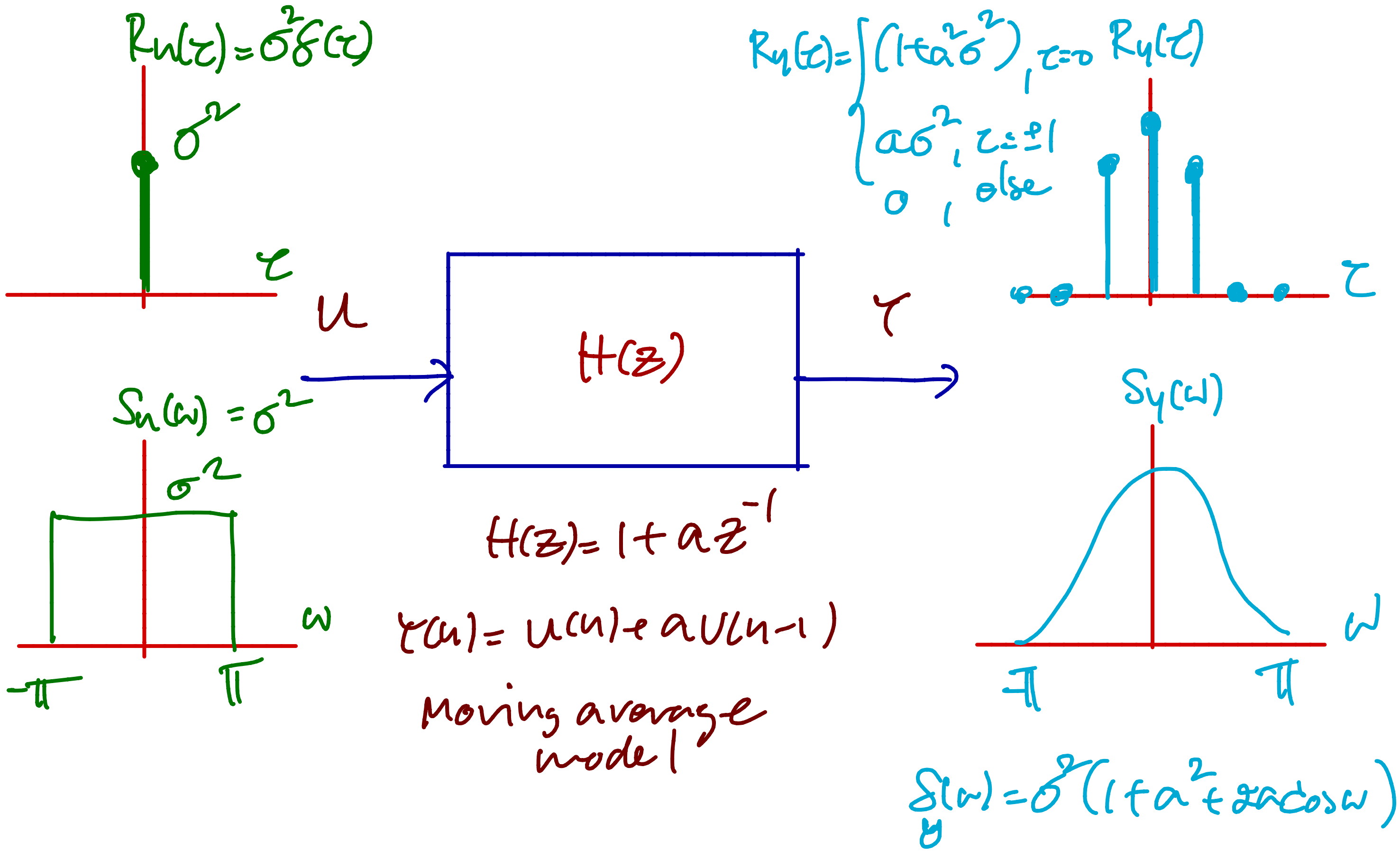


Input and output spectra

in conclusion, when input is white noise, the spectrum is flat



when white noise is passed through a filter, the output spectrum is no longer flat


Ru(τ) = σ²δ(τ)
σ²
τ
U
H(z)
Y
Su(ω) = σ²
σ²
ω
−π
π
H(z) = 1 + az⁻¹
Y(n) = U(n) + aU(n−1)
Moving average model
Ry(τ) = (1+a²σ²), τ=0
aσ², τ=±1
0, else
Ry(τ)
τ
Sy(ω)
ω
−π
π
Sy(ω) = σ²(1 + a² + 2a cos ω)

Response to linear system: state-space models

consider a discrete-time linear system via a state-space model

$$X(k+1) = AX(k) + BU(k), \quad Y(k) = HX(k)$$

where $X \in \mathbf{R}^n, Y \in \mathbf{R}^p, U \in \mathbf{R}^m$

known results:

- two forms of solutions of state and output variables are

$$\begin{aligned} X(t) &= A^t X(0) + \sum_{\tau=0}^{t-1} A^\tau BU(t-1-\tau), \quad Y(t) = CX(t) \\ &= A^{t-s} X(s) + \sum_{\tau=s}^{t-1} A^{t-1-s} BU(\tau), \quad Y(t) = CX(t) \end{aligned}$$

- the autonomous system (when $U = 0$) is stable if $|\lambda(A)| < 1$

State-space models: autocovariance function

Theorem: let U be a i.i.d white noise sequence with covariance Σ_u and if i) A is **stable** and ii) $X(0)$ is uncorrelated with $U(k)$ for all $k \geq 0$ then

- $\lim_{n\to\infty} \mathbf{E}[X(n)] = 0$
- $C(n, n) \to \Sigma$ as $n \to \infty$ where

$$\Sigma = A\Sigma A^T + B\Sigma_u B^T$$

  (Σ is a *unique* solution to the Lyapunov equation)

- $X(t)$ is wide-sense stationary in **steady-state** sense, $i.e.$,

$$\lim_{n\to\infty} C(n+k, n) = C(k) = \begin{cases} A^k\Sigma, & k \geq 0 \\ \Sigma(A^T)^{|k|}, & k < 0 \end{cases}$$

proof: the mean of $X(t)$ converges to zero

let $m(n) = \mathbf{E}[X(n)]$ and it's easy to see

$$m(n) = \mathbf{E}[X(n)] = A\mathbf{E}[X(n-1)] + B\mathbf{E}[U(n-1)] = Am(n-1)$$

hence, $m(n)$ propagates like a linear system:

$$m(n) = A^n m(0)$$

and goes to zero as $n \to \infty$ since A is stable ■

zero-mean system: $\tilde{X}(n) = X(n) - m(n)$

$$\tilde{X}(n) = A\tilde{X}(n-1) + BU(n-1)$$

mean-removed process also follow the same state-space equation

proof: $\lim_{n\to\infty} C(n,n) = \Sigma$ and satisfies the Lyapunov equation

- $\tilde{X}(n)$ is uncorrelated with $U(k)$ for all $k \geq n$

$$\tilde{X}(t) = A^t \tilde{X}(0) + \sum_{\tau=0}^{t-1} A^\tau B U(t-1-\tau)$$

  because $\tilde{X}(0)$ is uncorrelated with $U(t)$ for all t and $\tilde{X}(t)$ is only a function of $U(t-1), U(t-2), \ldots, U(0)$

- since $\tilde{X}(n-1)$ is uncorrelated with $U(n-1)$, we obtain

$$C(n,n) = AC(n-1,n-1)A^T + B\Sigma_u B^T$$

  from the state equation: $\tilde{X}(n) = A\tilde{X}(n-1) + BU(n-1)$

- then we can write $C(n,n)$ recursively

$$C(n,n) = \underbrace{A^n C(0,0)(A^T)^n}_{\text{go to zero}} + \underbrace{\sum_{k=0}^{n-1} A^k B \Sigma_u B^T (A^T)^k}_{\text{converges}}$$

and observe its asymptotic behaviour when $n \to \infty$

- if A is stable, there exists γ s.t. $\|A^k\| \leq \gamma^k < 1$ (requires a proof)

$$\|A^n C(0,0)(A^T)^n\| \leq \|A\|^{2n} \|C(0,0)\| \leq \gamma^{2n} \|C(0,0)\| \to 0, \quad n \to \infty$$

- let $\Sigma = \sum_{k=0}^{\infty} A^k B \Sigma_u B^T (A^T)^k$, we can check that

$$\Sigma = A \Sigma A^T + B \Sigma_u B^T$$

- Σ is unique, otherwise, by contradiction

$$\Sigma_1 = A \Sigma_1 A^T + B \Sigma_u B^T, \quad \Sigma_2 = A \Sigma_2 A^T + B \Sigma_u B^T$$

we can subtract one from another and see that

$$\Sigma_1 - \Sigma_2 = A(\Sigma_1 - \Sigma_2)A^T = A^2(\Sigma_1 - \Sigma_2)(A^T)^2 = \cdots = A^n(\Sigma_1 - \Sigma_2)(A^T)^n$$

this goes to zero since A is stable ($\|A^k\| \to 0$)

$$\|\Sigma_1 - \Sigma_2\| = \|A^n(\Sigma_1 - \Sigma_2)(A^T)^n\| \leq \|A\|^{2n}\|\Sigma_1 - \Sigma_2\| \to 0$$

this completes the proof ■

proof: $\tilde{X}(n)$ is wide-sense stationary in steady-state

- $\tilde{X}(k)$ is uncorrelated with $\{U(k), U(k+1), \ldots, U(n-1)\}$
- from the solution of $\tilde{X}(n)$

$$\tilde{X}(n) = A^{n-k}\tilde{X}(k) + \sum_{\tau=k}^{n-1} A^{n-1-\tau} BU(\tau), \quad k < n$$

  the two terms on RHS are uncorrelated

- the autocovariance function is obtained by (for $n > k$)

$$\begin{aligned} C(n,k) &= \mathbf{E}[\tilde{X}(n)\tilde{X}(k)^T] \\ &= A^{n-k}\mathbf{E}[\tilde{X}(k)\tilde{X}(k)^T] + \sum_{\tau=k}^{n-1} A^{n-1-\tau} B\mathbf{E}[U(\tau)\tilde{X}(k)^T] \\ &= A^{n-k}C(k,k) + 0 \end{aligned}$$

  which converges to $A^{n-k}\Sigma$ as $n, k \to \infty$ if A is stable ■

State-space models: autocovariance of output

output equation:

$$Y(n) = HX(n), \quad \tilde{Y}(n) = H\tilde{X}(n)$$

when $X(n)$ is wide-sense stationary (in steady-state) then

when $n, k \to \infty$, we have

$$C_y(n,k) = HC_x(n,k)H^T = HA^{n-k}C_x(k,k)H^T, \quad n \geq k$$

and

$$\lim_{n\to\infty} C_y(n,n) = \lim_{n\to\infty} HC_x(n,n)H^T = H\Sigma H^T$$

where Σ is the solution to the Lyapunov equation: $\Sigma = A\Sigma A^T + B\Sigma_u B^T$

example: AR process with $a = 0.7$ and U is i.i.d. white noise with $\sigma^2 = 2$

$$Y(n) = aY(n-1) + U(n-1)$$

1st-order AR process is already in state-space equation

- in steady-state, the covariance function at lag 0 converges to α where

$$\alpha = a\alpha^2 + \sigma^2 \quad \Longrightarrow \quad \alpha = \frac{\sigma^2}{1-a^2}$$

(we have solved the Lyapunov equation)

- in steady-state, the covariance function is given by

$$C(\tau) = \frac{\sigma^2 a^{|\tau|}}{1-a^2}$$

References

Chapter 9-10 in
A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*, 3rd edition, Pearson Prentice Hall, 2009

Chapter 9 in
H. Stark and J. W. Woods, *Probability, Statistics, and Random Processes for Engineers*, 4th edition, Pearson, 2012