3rd way of estimating spectrum

Predictors, Yule Walker equations, Wold theorem

lunedì 25 maggio 2020 16:24

[1]

$x \in \mathbb{R}^n \qquad x \in \mathbb{C}^n$

$x \xrightarrow{\text{predict}} w \quad \in \mathbb{R}, \ \in \mathbb{C}$

$$\hat{w} = a^{+} x \qquad a = \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} \qquad \hat{w} = \sum_{j=0}^{n-1} a_j^* x_j$$

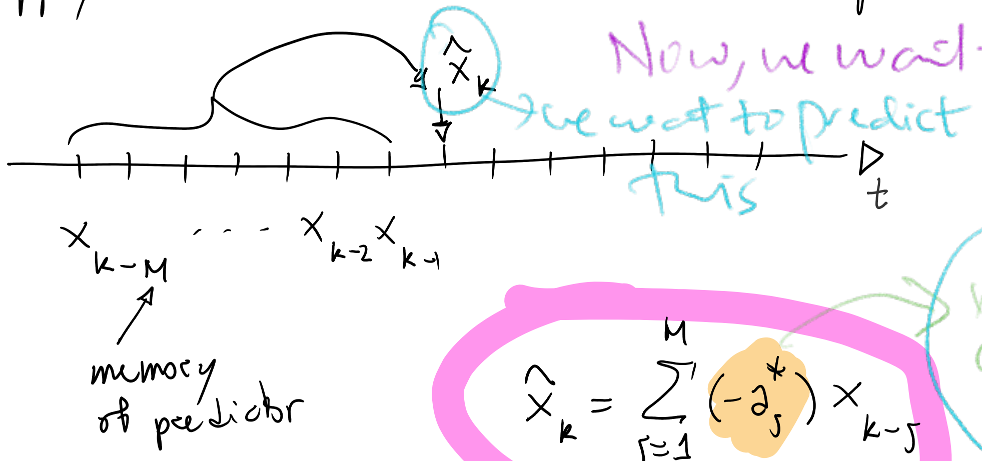
$$\mathcal{E} = E\left[|\varepsilon|^2\right] \qquad \varepsilon = w - \hat{w}$$

Orthogonality principle

$$E[\varepsilon^* x] = \begin{bmatrix} E[\varepsilon^* x_0] \\ E[\varepsilon^* x_1] \\ \vdots \\ E[\varepsilon^* x_{n-1}] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\boxed{E[\varepsilon^* w] = \mathcal{E}_{min}}$$

Apply this to discrete-time stochastic processes

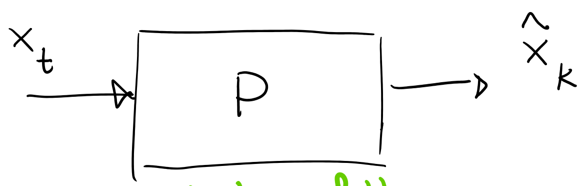


Now, we want to predict the Future!

$$\hat{x}_k = \sum_{j=1}^{M} (-a_j^*)\, x_{k-j}$$

we define the coefficient like this!

This is the equation of a filter



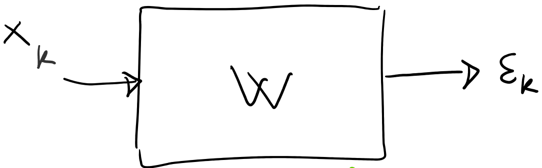
prediction filter

$$\varepsilon_k = x_k - \hat{x}_k = x_k - \sum_{j=1}^{M} (-a_j^*)\, x_{k-j}$$

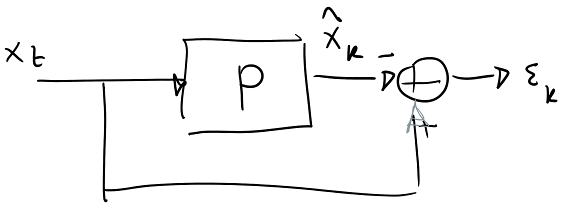
we calculate error after 1 clk time

$$= \sum_{j=0}^{M} a_j^*\, x_{k-j} \qquad a_0 = 1$$

We reproduce a sequence of ERRORS if we continue with the process & the sequence of ERRORS is a Random Process



whitener filter



$x_k \rightarrow \hat{x}_k, \varepsilon_k$ are stationary

$$\mathcal{E} = E[|\varepsilon_k|^2]$$

error is orthogonal to previous values of 'x'

$$\begin{cases} E[\varepsilon_k^* x_{k-l}] = 0 & l = 1, 2, \dots, M \\ E[\varepsilon_k^* x_k] = \mathcal{E}_{min} \end{cases}$$

data considered in the prediction

$$\underbrace{E[x_0^* x_{k-l-k+j}]}$$

M * M .

$$E\left[\left(\sum_{j=0}^{M} a_j^* x_{k-j}\right)^* x_{k-\ell}\right] = \sum_{j=0}^{M} a_j \underbrace{E\left[x_{k-j}^* \, x_{k-\ell}\right]}_{C_x(k-\ell-(k-j))} = 0$$

$$C_x(k-\ell-(k-j)) = C_x(j-\ell)$$

$$\sum_{j=0}^{M} a_j \, C_x(j-\ell) = 0 \qquad \ell = 1, 2, \dots, M$$

$$E\left[\left(\sum_{j=0}^{M} a_j^* x_{k-j}\right)^* x_k\right] = \sum_{j=0}^{M} a_j \, C_x(j-\ell) = \mathcal{E}_{min} \qquad \ell = 0$$

Correlation Matrix

$$\underbrace{\begin{bmatrix} C_x(0) & C_x(1) & \cdots & C_x(M) \\ C_x(-1) & C_x(0) & \cdots & C_x(M-1) \\ \vdots & & & \vdots \\ C_x(-M) & & & C_x(0) \end{bmatrix}}_{R_x} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} \mathcal{E}_{min} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Yule - Walker equations

$$R_x \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} \mathcal{E}_{min} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

For perfect prediction the gist lies in the NULLSPACE

Case 1

$\det R_x \neq 0$

Non Singular

R_x is non-singular

$\Downarrow$

only $0 \in \ker R_x$

$\Downarrow$

$\nexists\ p \neq 0 \quad R_x p = 0$

$\Downarrow$

no perfect prediction exists

all the Unknowns are Here

$\mathcal{E}_{min} > 0$

$$R_x \begin{bmatrix} \frac{1}{\mathcal{E}_{min}} \\ \frac{a_1}{\mathcal{E}_{min}} \\ \vdots \\ \frac{a_m}{\mathcal{E}_{min}} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$b_j = \frac{a_j}{\mathcal{E}_{min}} \quad \left(b_0 = \frac{1}{\mathcal{E}_{min}} \right)$$

$$R_x \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} b_0 \\ \vdots \\ b_m \end{bmatrix} = R_x^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Case 2

$\det R_x = 0$

$\exists\ p \quad R_x p = 0 \qquad p \neq \emptyset$

$\Downarrow$

there is the chance of a perfect prediction of a perfect prediction

$$p = \begin{bmatrix} p_0 \\ \vdots \\ p_M \end{bmatrix} \qquad p \in \ker R_x$$

$p_0 \neq 0$

$$a = \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \frac{p}{p_0}$$

cannot be zero

$$R_x \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

perfect prediction

$p_0 = 0$

$\frac{p}{p_0}$ does not exist

try again

$$p \in \ker R_x \Rightarrow p_0 = 0$$

$$\begin{bmatrix} C_x(0) & C_x(1) & C_x(2) & \cdots & C_x(M) \\ C_x(-1) & C_x(0) & C_x(1) & & \\ C_x(-2) & C_x(-1) & C_x(0) & & \\ \vdots & & & \ddots & \\ C_x(-M) & & & & C_x(0) \end{bmatrix} \begin{bmatrix} 0 \\ p_1 \\ \vdots \\ \vdots \\ p_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

doesn't count

discard first column of R_x

first row p

first row of R_x

first row of the result

$$
\begin{bmatrix} \cancel{c_x(0)} & \cancel{c_x(1)} & \cdots & \cancel{c_x(M)} \\ \cancel{c_x(1)} & c_x(0) & c_x(1) \cdots & c_x(M-1) \\ \vdots & c_x(-1) & c_x(0) & \vdots \\ \cancel{c_x(M)} & c_x(-M+1) & \cdots & c_x(0) \end{bmatrix}
\begin{bmatrix} \cancel{0} \\ p_1 \\ \vdots \\ p_M \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\qquad \varepsilon_{min}
$$

the same R_x

of order M

$\Downarrow$

Yule-Walker equation

for a filter of order $M-1$ dimention reduced by '1'

— look-up vectors in ker reduced R_x

— if every vector of the kernel starts with $\emptyset$

$\Downarrow$

discard smaller row

discard smaller colum

$$\Downarrow$$

Yule-Walker equations for a filter with $M-2$ taps

proceed until you find a vector in ker reduced R_x starting with $\neq 0$

Whitener filter (filter that produces sequence of ERRORS)

$$\hat{x}_k = \sum_{j=1}^{\infty} (-a_j^*)\, x_{k-j}$$

$$\varepsilon_k = \sum_{j=0}^{\infty} a_j^*\, x_{k-j} \qquad a_0 = 1$$

The process of errors in a Best predictor is a WHITE PROCESS

"whitener" why?

$$x_t \longrightarrow \boxed{w} \longrightarrow \varepsilon_k$$

$$C_\varepsilon(\tau) = E[\varepsilon_0^* \varepsilon_\tau] \qquad \tau > 0$$

$$= E\left[\left(\sum_{j=0}^{\infty} a_j^* x_{0-j}\right)^* \varepsilon_\tau\right]$$

$$= \sum^{\infty} a_j\, E[x_{0-j}^* \varepsilon_\tau] = \sum^{\infty} a_j \cdot 0 = 0$$

orthogonal

Non Random $= \sum_{j=0}^{\infty}{}' a_j \, E\left[\underbrace{x^*_{0-j}\, \varepsilon_\tau}\right] = \sum_{j=0}^{\infty}{}' a_j \, 0 = 0$

$\tau > 0 \qquad \tau > -j$

$\varepsilon_\tau = x_\tau - \hat{x}_\tau \qquad \tau > 0$

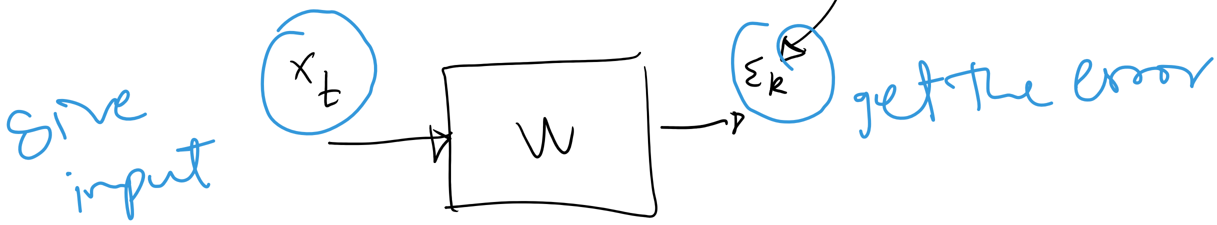
$$\hat{x}_\tau = \sum_{l=1}^{\infty}{}' (-a_j^*)\, x_{\tau - l}$$

correlation is zero from $-\infty$ to τ for $\tau > 0$

stationarity

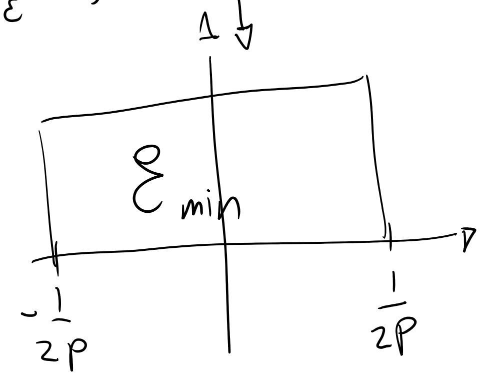
$$C_\varepsilon(\tau) = \begin{cases} 0 & \tau > 0 \\ E[|\varepsilon_0|^2] = \mathcal{E}_{min} & \tau = 0 \\ C_\varepsilon^*(-\tau) = 0 & \tau < 0 \end{cases}$$

correlation of a white process



transfer function

$$S_x^W(f) \, |W(f)|^2 = S_\varepsilon^W(f)$$



$$S_x^w(f)\,|W(f)|^2 = P\,\mathcal{E}_{min}$$

original Power Spectrum — squared T/F of whitener

All pole estimation – maximum entropy

$$\boxed{S_x^w(f) = \frac{P\,\mathcal{E}_{min}}{|W(f)|^2}}$$

This is one way of making Estimation

Recipe (take a large M)

1). collect samples

2). estimate $\hat{C}_x(\tau) \to$ fill $\hat{R}_x$

3). solve Yule–Walker equations $\to$ coefficients a_j $\mathcal{E}_{min}$

4). $$\varepsilon_k = \sum_{j=0}^{M} a_j^* x_{k-j}$$

impulse response of whitener

$$W(f) = \sum_{j=0}^{M-1} a_j^* e^{-2\pi i p f_j}$$

- $$\hat{S}_x^w(f) = \frac{P\,\mathcal{E}_{min}}{|W(f)|^2}$$

completely different way of to estimate power spectrum

$$S_x^w(f)\,|W(f)|^2 = p\,\mathcal{E}_{min}$$

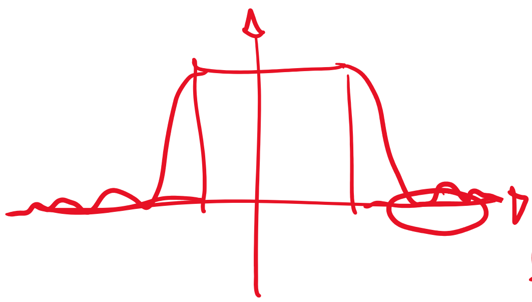
$$\int_{-\frac{1}{2p}}^{\frac{1}{2p}} S_x^w(f)\,|W(f)|^2\,df = \int_{-\frac{1}{2p}}^{\frac{1}{2p}} p\,\mathcal{E}_{min}\,df$$

$$\boxed{\int_{-\frac{1}{2p}}^{\frac{1}{2p}} S_x^w(f)\,|W(f)|^2\,df = \mathcal{E}_{min}}$$

→ perfect predictability $\Rightarrow$ $\mathcal{E}_{min} = 0$

This is Theoretically fundamental formulation of stochastic process

$$\int_{-\frac{1}{2p}}^{\frac{1}{2p}} \underbrace{S_x^w(f)\,|W(f)|^2}_{\geq 0}\,df = 0$$



$$\Downarrow$$

either $S_x^w(f) = 0$

or $W(f) = 0$ → possible only for some isolated frequencies

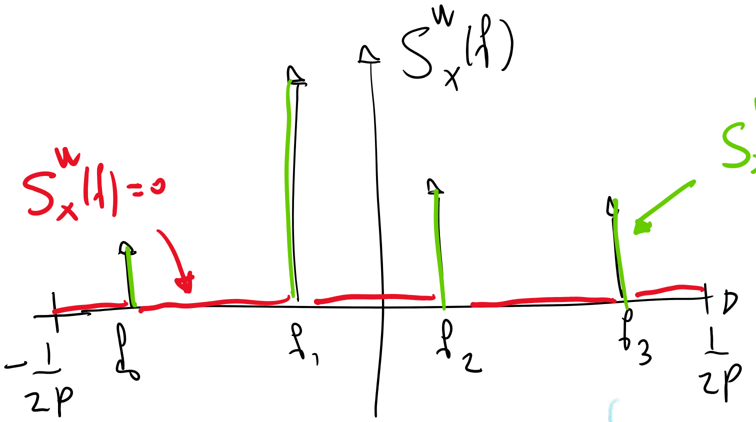
Teorema di Paley-Wiener

→ Predictable processes

$\mathcal{E}_{min} = 0$

To be perfectly predictable when we multiply our process by whitener it should be equal to zero!

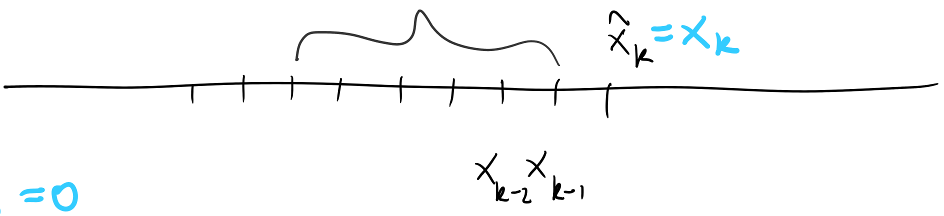
$$S_x^w(f) = \sum_j{}' \alpha_j \, \delta(f - f_j) \qquad \sum_j \alpha_j = E[|x_0|^2]$$



This is The shape of P.S of a RP that can predicted perfectly

Sequence of Delta function

predictability = way of generating the process



$\mathcal{E}_{min} = 0$

$$x_k = \hat{x}_k = \sum_{j=1}^{M}{}' (-a_j^*) \, x_{k-j}$$

initial condition

Texas Instrument

Frequency Synthesizers

look in The INTERNET

* Frequency Synthesis by Recursion

Regular processes

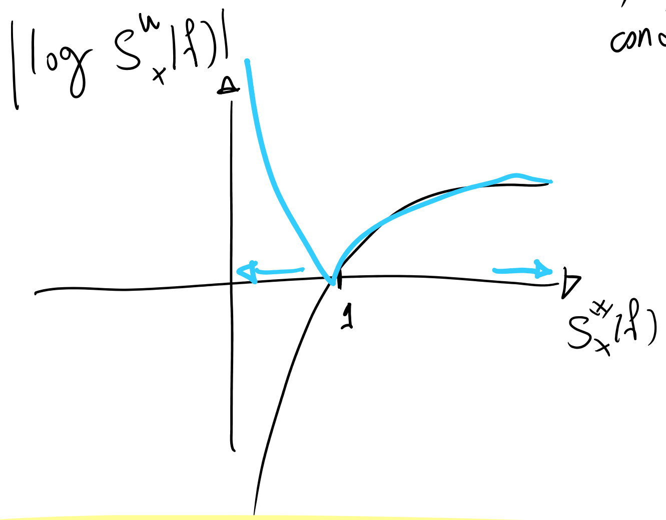
complementary of predictable Process

i.e They rarely touch Zero & Almost Never go to Infinity

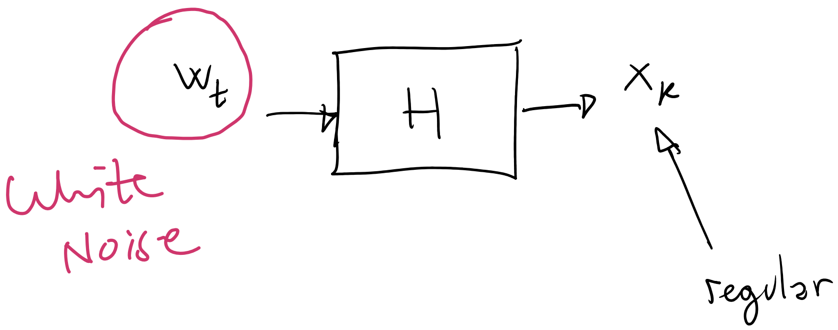
$S_x^u(f)$ stays away from ϕ and ∞

$$\int_{-\frac{1}{2p}}^{\frac{1}{2p}} |\log S_x^u(f)|\, df < +\infty$$

Paley - Wiener condition



* generate a regular process by filtering white noise



Regular process & predictable process are enough to Generate all The Stationary Processes

→ Wold decomposition theorem

If x_k is discrete time and stationary then

1) $\boxed{x_k = x_k^p + x_k^R}$

predictable   regular

2) x_k^p x_k^R are orthogonal $\forall j,k$ $E[(x_j^p)^* x_k^R]=0$

Consequence on power spectra

$$S_x^w(f) = \mathcal{F}[C_x](f)$$

$$C_x(\tau) = E[x_0^* x_\tau] \qquad x_k = x_k^p + x_k^R$$

$$= E[(x_0^p + x_0^R)^* (x_\tau^p + x_\tau^R)]$$

$$= E[(x_0^p)^* x_\tau^p] + E[(x_0^R)^* x_\tau^R] +$$

$$\underbrace{E[(x_0^p)^* x_\tau^R]}_{0} + \underbrace{\cancel{E[(x_0^R)^* x_\tau^p]}}_{0}$$

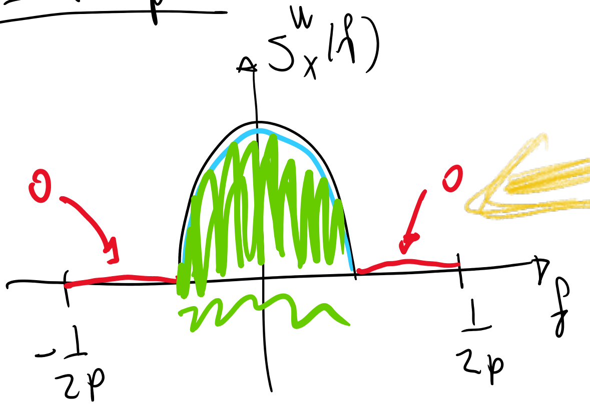
$$C_x(\tau) = C_{x^p}(\tau) + C_{x^R}(\tau)$$

$$\mathcal{F}[\ \downarrow\ ] = \mathcal{F}[\ \downarrow\ ] + \mathcal{F}[\ \downarrow\ ]$$

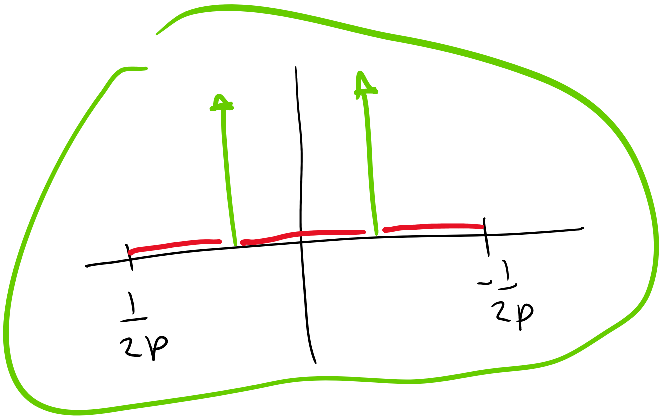
$$S_x^w(f) = S^w_{x^p}(f) + S^w_{x^R}(f)$$

$$S_x^w(f) = S_{xP}^w(f) + S_{xR}^w(f)$$

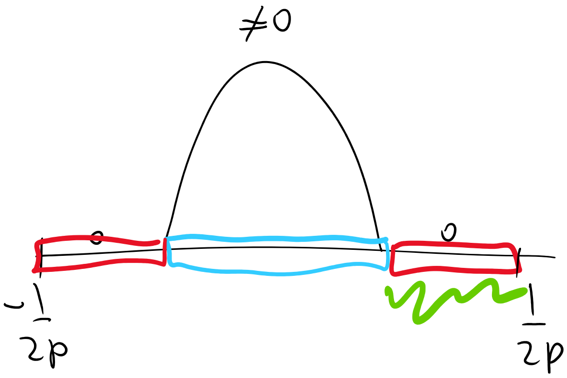
Examples



not a spectrum of a stationary process

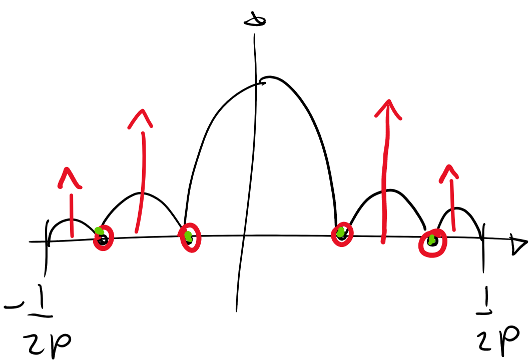


good spectrum of a stationary process

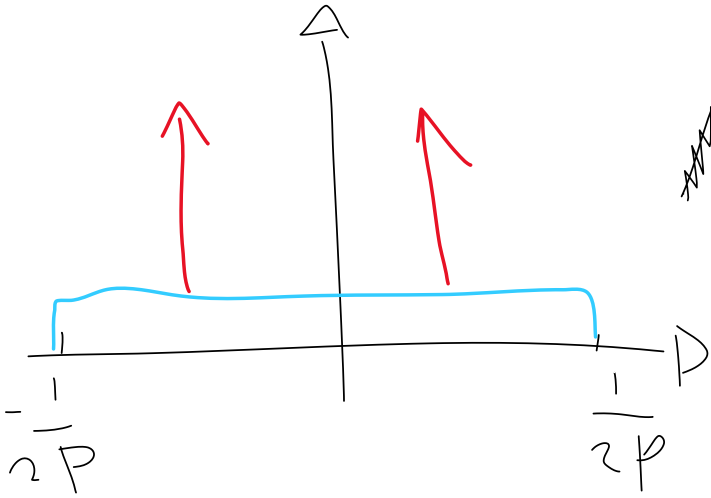


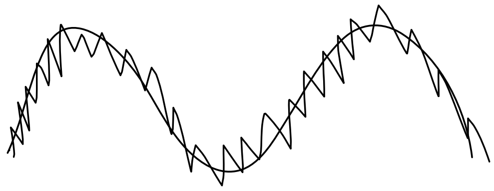
$$\int_{-\frac{1}{2p}}^{\frac{1}{2p}} |\log \underbrace{S_x^w(f)}_{=0}|\, df = +\infty$$

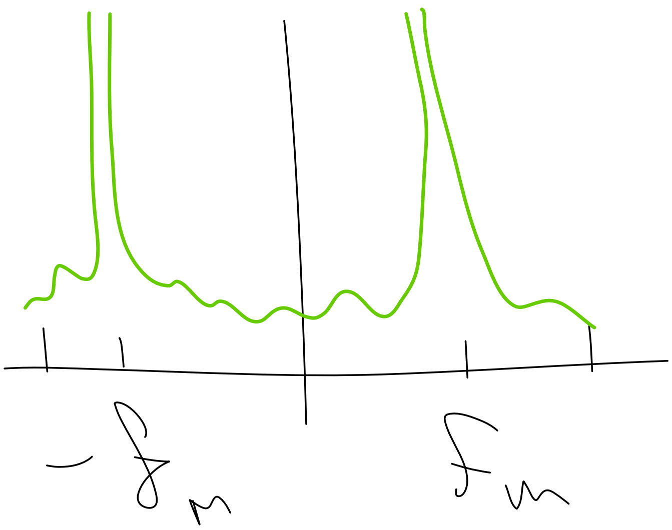
$\underbrace{}_{+\infty}$



$$\int_{-\frac{1}{2p}}^{\frac{1}{2p}} |\log S_x^w(f)|\, df < \infty$$


$-\frac{1}{2P}$
$\frac{1}{2P}$




$-f_m$
f_m