HSLU Hochschule Luzern

Time Series Analysis

Spectral analysis I

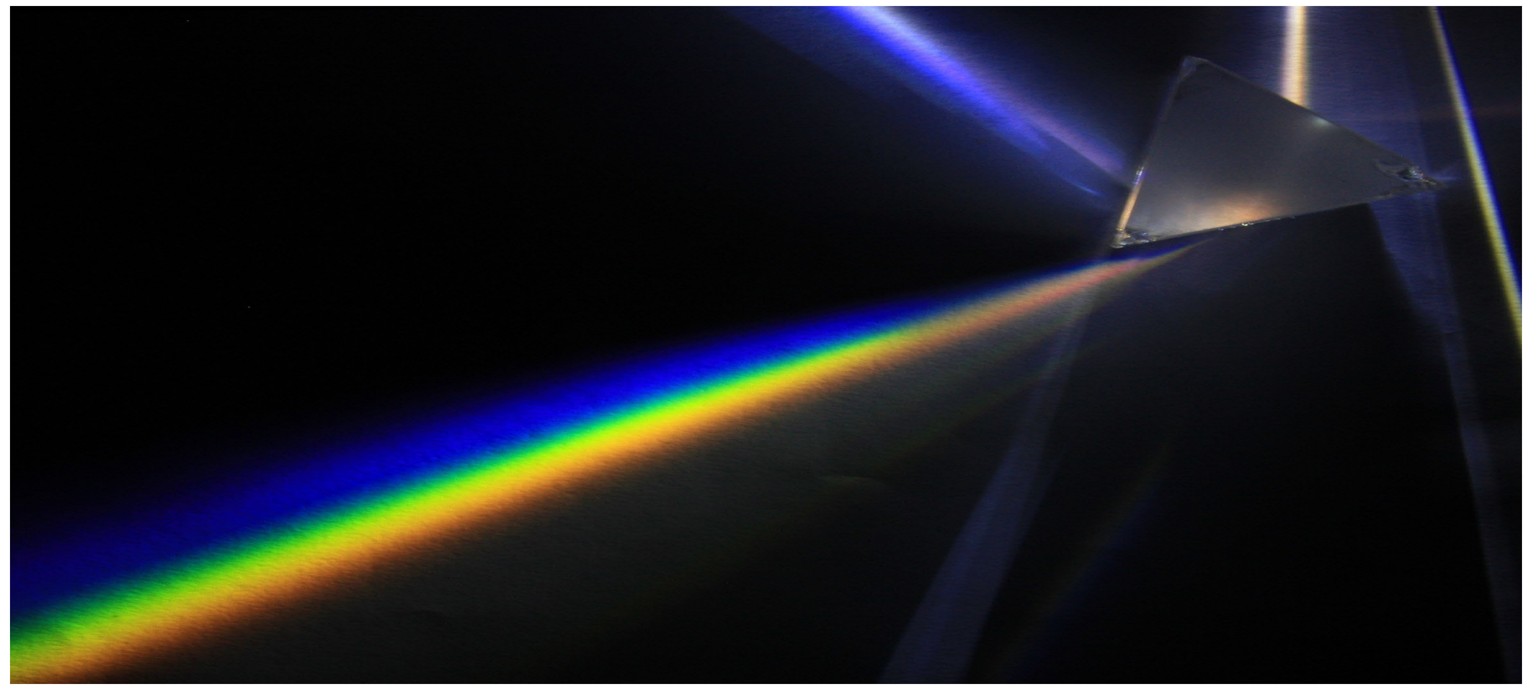
Dr. Ludovic Amruthalingam
ludovic.amruthalingam@hslu.ch

Informatik

Outline

- Spectral analysis
- Stochastic periodic process
- Periodogram
- Spectral density
- Spectral density of known processes
- Sample spectral density

Spectral analysis



D-Kuru, wikimedia

Spectral analysis is an application of **Fourier analysis**.

Fourier analysis decomposes a time series into sinusoidal components.

- Spectral analysis interprets these components in terms of their **power spectrum**.
- **Power** quantifies how much of the TS variance is contributed by a specific frequency.
- **Spectrum** refers to the distribution of frequencies that composes the TS.

Thus, spectral analysis studies the **distribution of variance across frequencies** within a time series.

- Periodic TS: identifies **dominant** frequencies (harmonics of the fundamental frequencies).
- Aperiodic TS: identifies a **continuum** of frequencies, representing the TS characteristics.

Like a prism for light, spectral analysis separates TS into their constituent frequencies.

Stochastic periodic process

Consider a **zero-mean periodic process** with period $\mathrm{P} = 1/\omega$ of the form

$$X_t = Rcos\big(2\pi(\omega \cdot t + \phi)\big)$$

where the magnitude (amplitude) R and the phase ϕ are both **independent random variables**.

We have seen that X_t could be rewritten as

$$\begin{aligned} X_t &= R(cos(2\pi\omega t)\cos(2\pi\phi) - sin(2\pi\omega t)\sin(2\pi\phi)) \\ &= Acos(2\pi\omega t) + \mathrm{B}sin(2\pi\omega t) = Ac_t + Bs_t \end{aligned}$$

($\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$ *)*

with $A = R\cos(2\pi\phi)$, $\mathrm{B} = -\mathrm{R}\sin(2\pi\phi)$, $c_t = cos(2\pi\omega t)$ and $s_t = sin(2\pi\omega t)$.

(recall that $\sin^2 x + \cos^2 x = 1$*)*

Then $R^2 = A^2 + B^2$ and $\phi = \arctan\left(-\frac{B}{A}\right)/2\pi$.

$$A, B \sim \mathcal{N}\left(0, \sigma^2\right), independent \quad \Leftrightarrow \quad R^2 \sim \chi_2^2, \qquad \phi \sim Unif\,(-\pi, \pi), \qquad independent$$

Stationarity of the periodic process

With $A, B \sim \mathcal{N}\left(0, \sigma^2\right)$ uncorrelated, the process $X_t = Ac_t + Bs_t$ is **stationary**.

Mean:

$$E[X_t] = E[Ac_t + Bs_t] = c_t E[A] + s_t E[B] = 0$$

Auto-covariance:

$$\begin{aligned} cov(X_{t+h}, X_t) &= cov(Ac_{t+h} + Bs_{t+h}, Ac_t + Bs_t) \\ &= c_{t+h} cov(A, Ac_t + Bs_t) + s_{t+h} cov(B, Ac_t + Bs_t) \\ &= c_{t+h} c_t cov(A, A) + c_{t+h} s_t cov(A, B) + s_{t+h} c_t cov(B, A) + s_{t+h} s_t cov(B, B) \\ &= c_{t+h} c_t \sigma^2 + s_{t+h} s_t \sigma^2 = \sigma^2 c_h = \gamma_h \end{aligned}$$

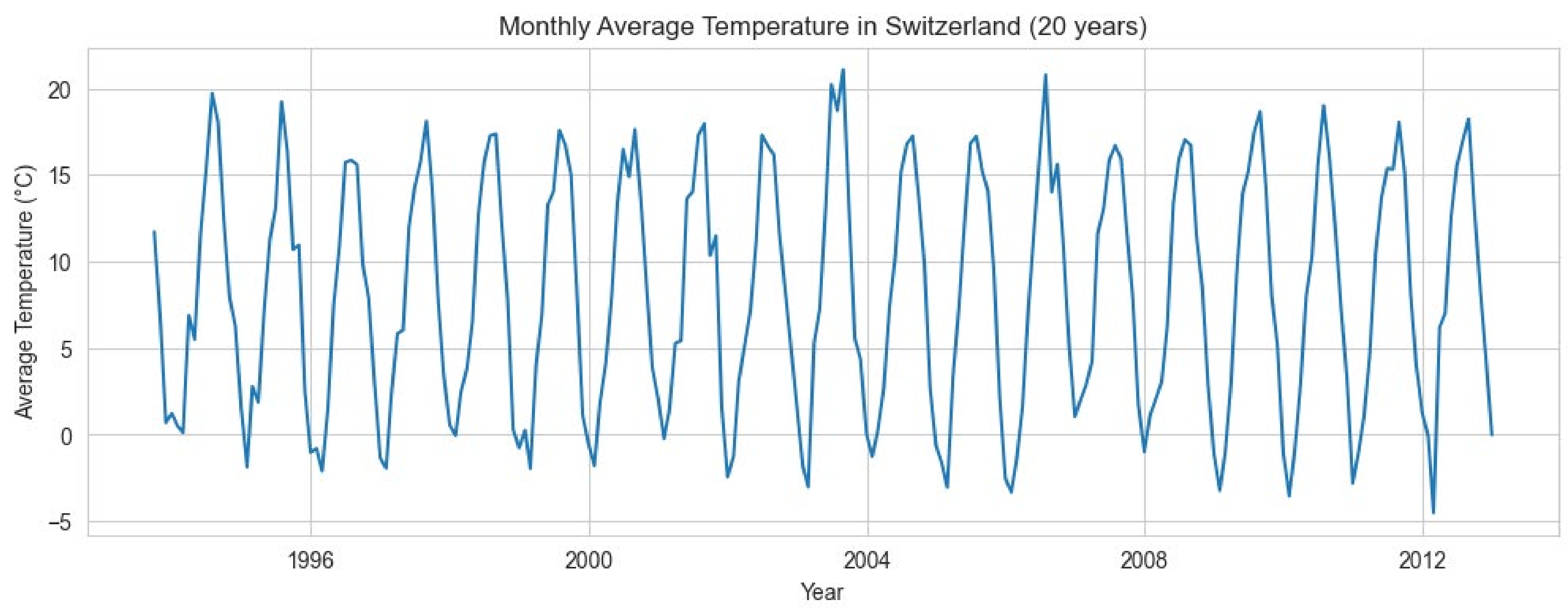
$(\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta))$

Variance:

$$Var(X_t) = Var(Ac_t + Bs_t) = c_t^2 Var(A) + s_t^2 Var(B) + 2cov(A, B) = \sigma^2\left(c_t^2 + s_t^2\right) = \sigma^2 = \gamma_0$$

Given a, b realization of A, B, then an **estimate** of σ^2 (i.e., the $\mathrm{E}\left[R^2\right]$) is $\hat{\sigma}^2 = a^2 + b^2$.

Note on deterministic seasonality



We can model the signal using **deterministic** components:

$$X_t = \mu + r \cdot \cos(2\pi\omega t) + \epsilon_t \text{ with } \epsilon_t \sim \mathcal{N}(0,1)$$

The **mean** is time dependent implying **non-stationarity**:

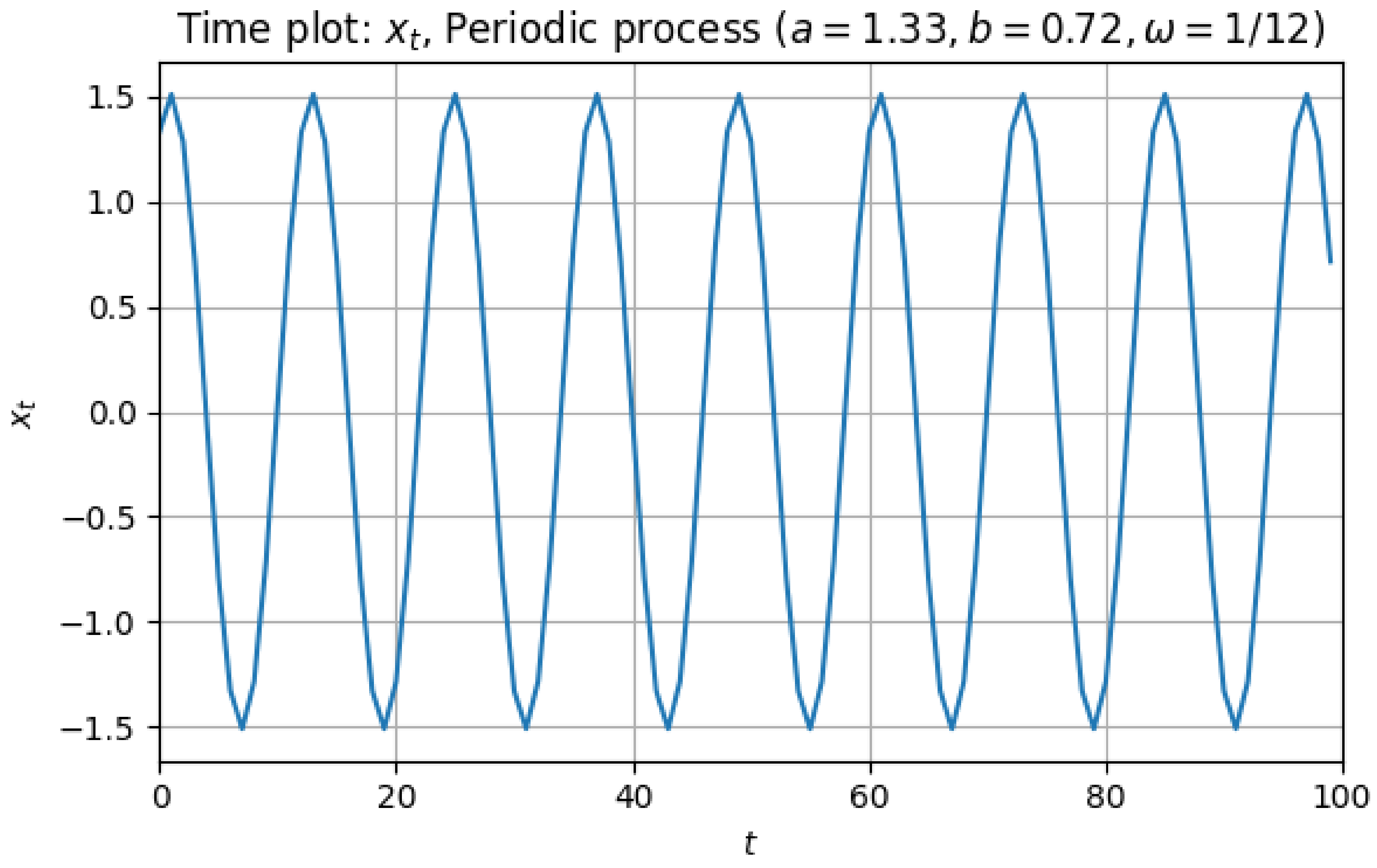
$$E[X_t] = E[\mu + r \cdot \cos(2\pi\omega t) + \epsilon_t] = \mu + r \cdot \cos(2\pi\omega t) + E[\epsilon_t] = \mu + r \cdot \cos(2\pi\omega t)$$

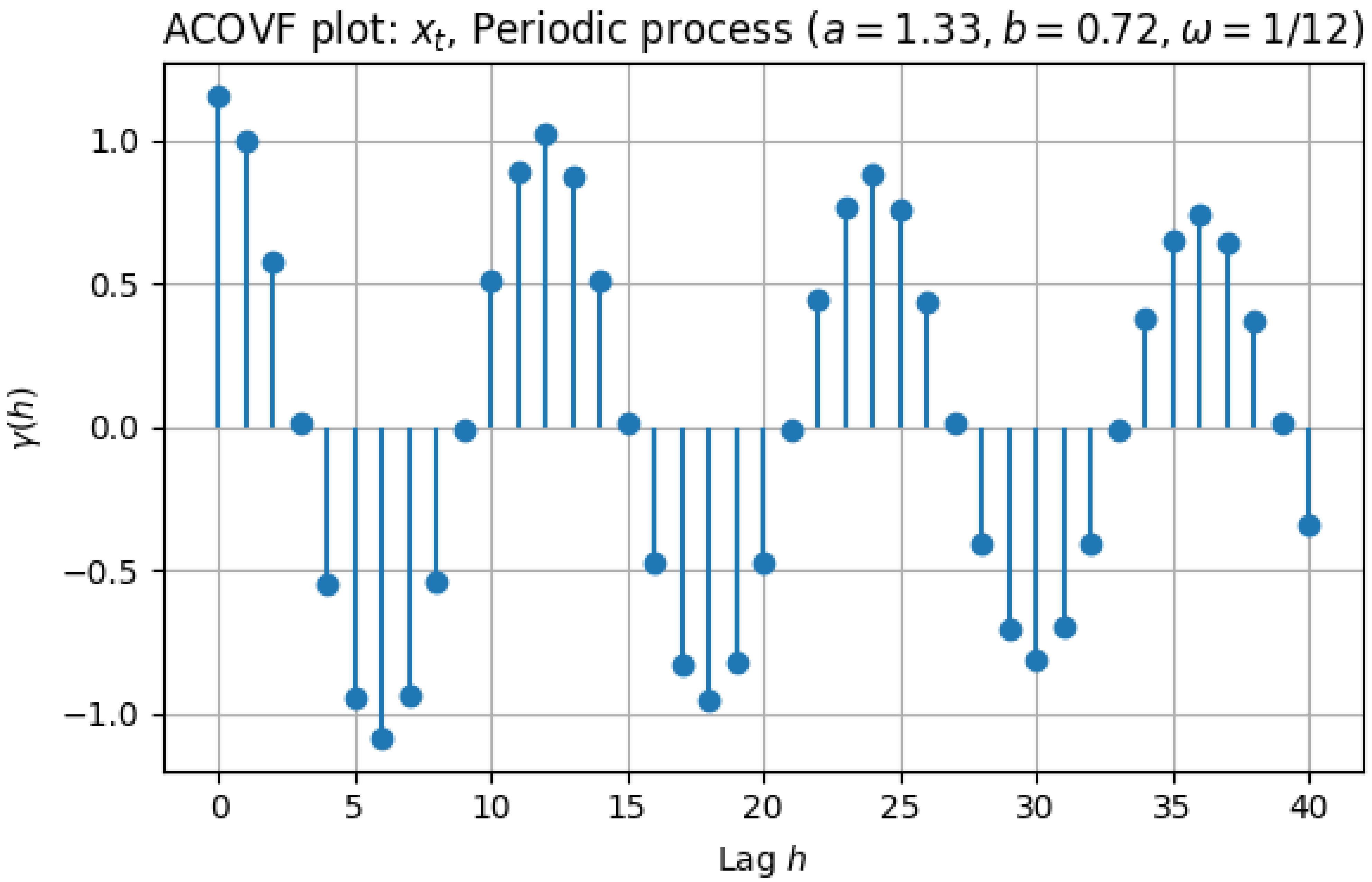
Intuition: if you know the month, you can predict the expected value.

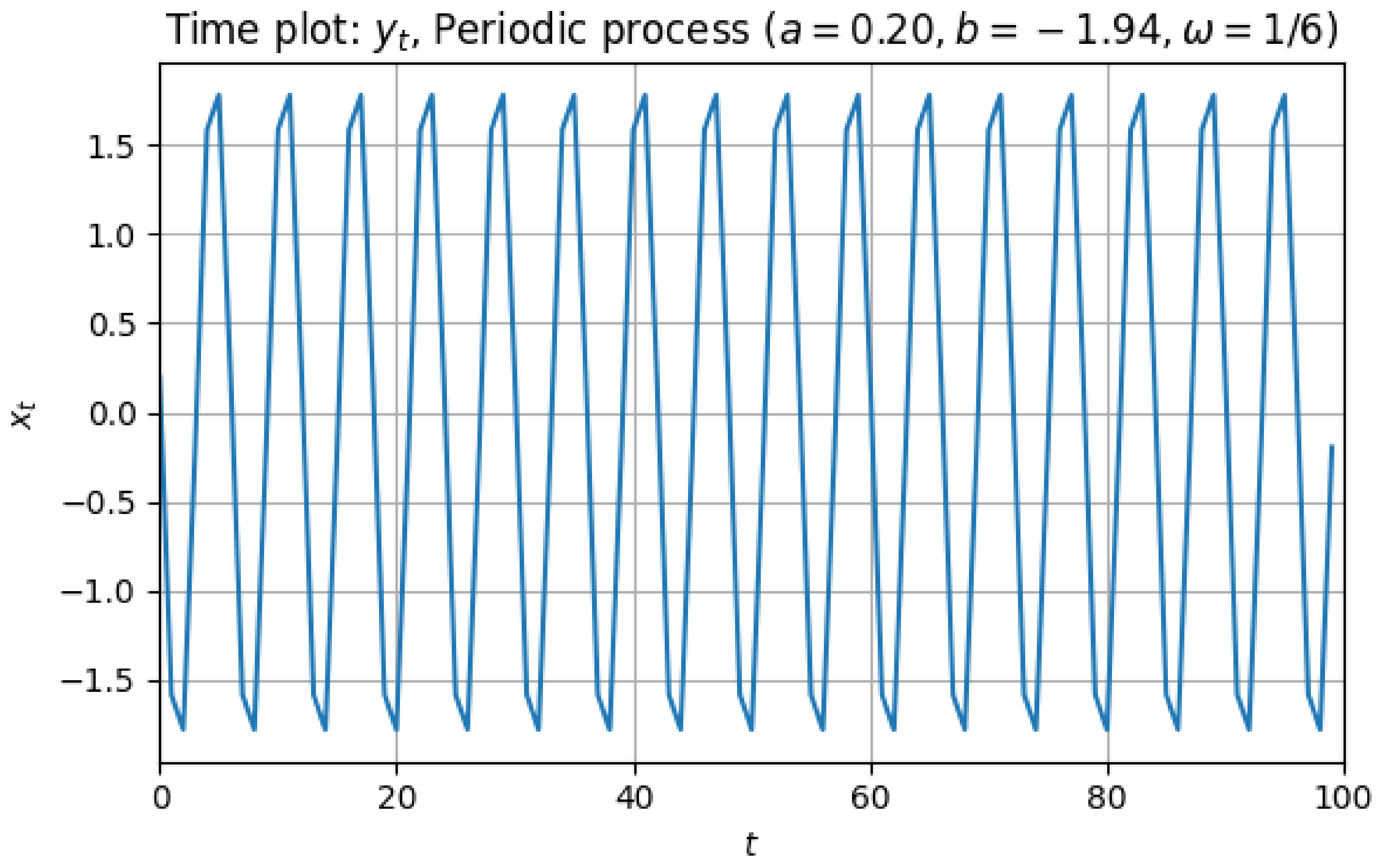
- The mean in January is different from the mean in July.

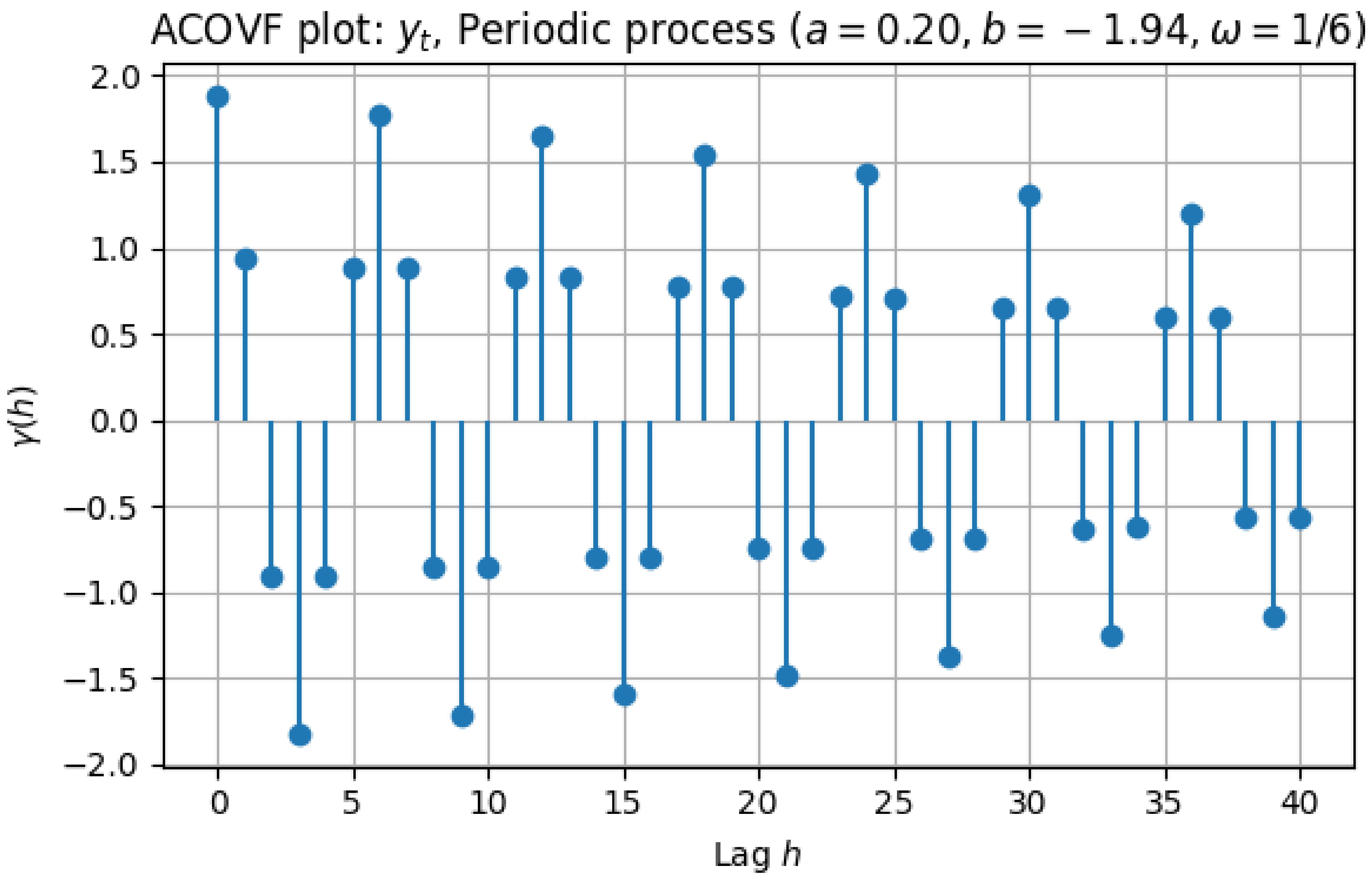
With stochastic periodic process the periodic behavior comes from the autocovariance, not the mean.

- The mean remain constant, while the autocovariance depends only on lag h, implying **stationarity**.


Time plot: x_t, Periodic process ($a = 1.33, b = 0.72, \omega = 1/12$)
x_t
t


ACOVF plot: x_t, Periodic process ($a = 1.33, b = 0.72, \omega = 1/12$)
$\gamma(h)$
Lag h


Time plot: y_t, Periodic process ($a = 0.20, b = -1.94, \omega = 1/6$)
x_t
t


ACOVF plot: y_t, Periodic process ($a = 0.20, b = -1.94, \omega = 1/6$)
$\gamma(h)$
Lag h

Mixtures of periodic series

A generalization of the periodic process is

$$X_t = \sum_{j=1}^{k} \left(A_j c_{t,j} + \mathrm{B}_j s_{t,j} \right)$$

with $A_j, B_j \sim \mathcal{N}\left(0, \sigma_j^2\right)$ **uncorrelated** random variables $\forall j$, and $c_{t,j} = cos\left(2\pi\omega_j t\right)$, $s_{t,j} = sin\left(2\pi\omega_j t\right)$.

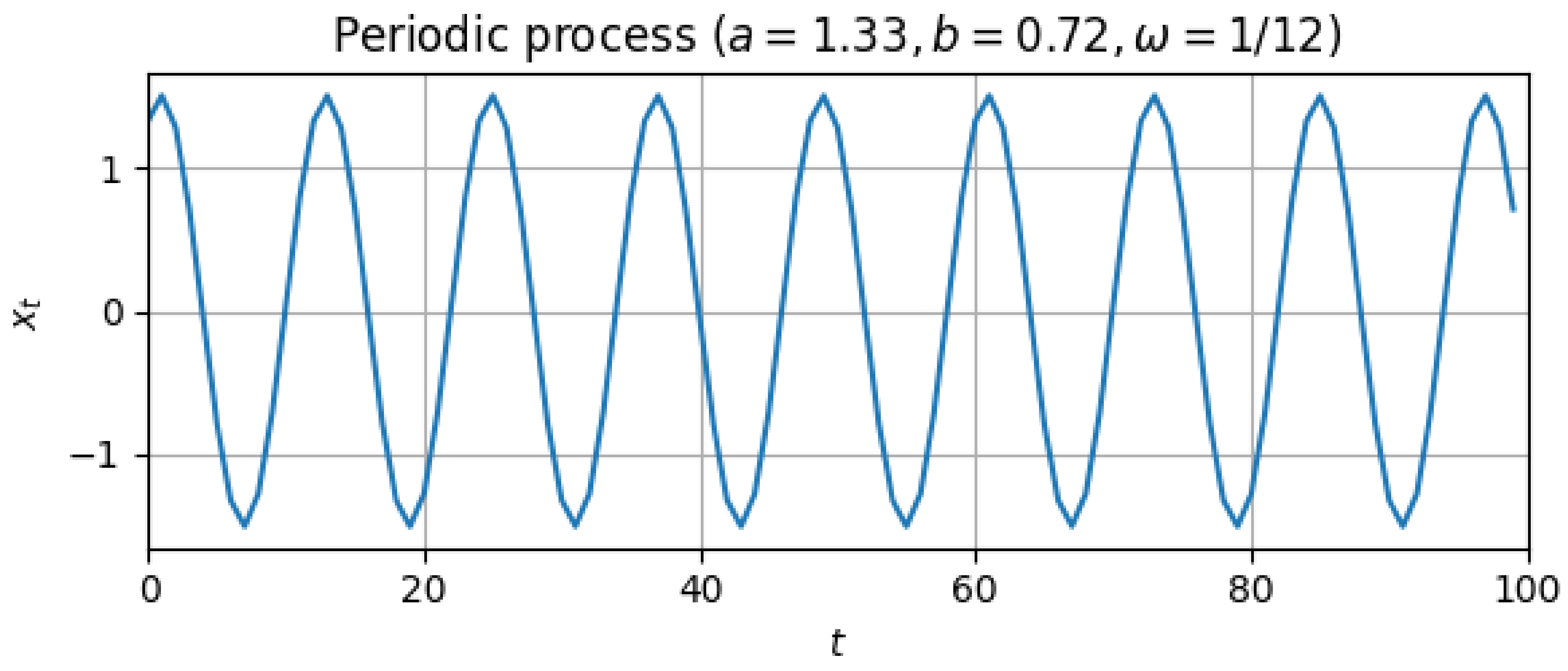
X_t is a **zero-mean stationary periodic** process with autocovariance

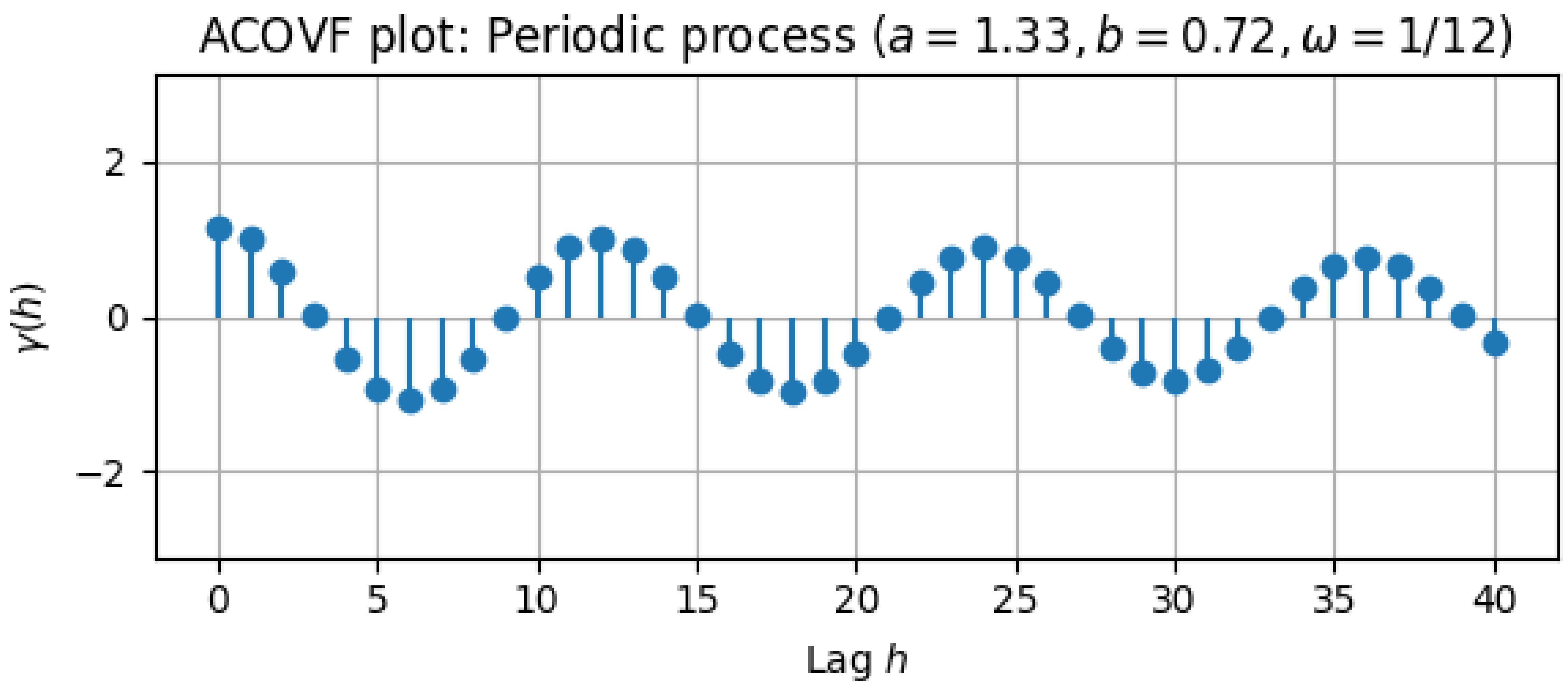
$$\gamma_h = \sum_{j=1}^{k} \sigma_j^2 c_{h,j}$$

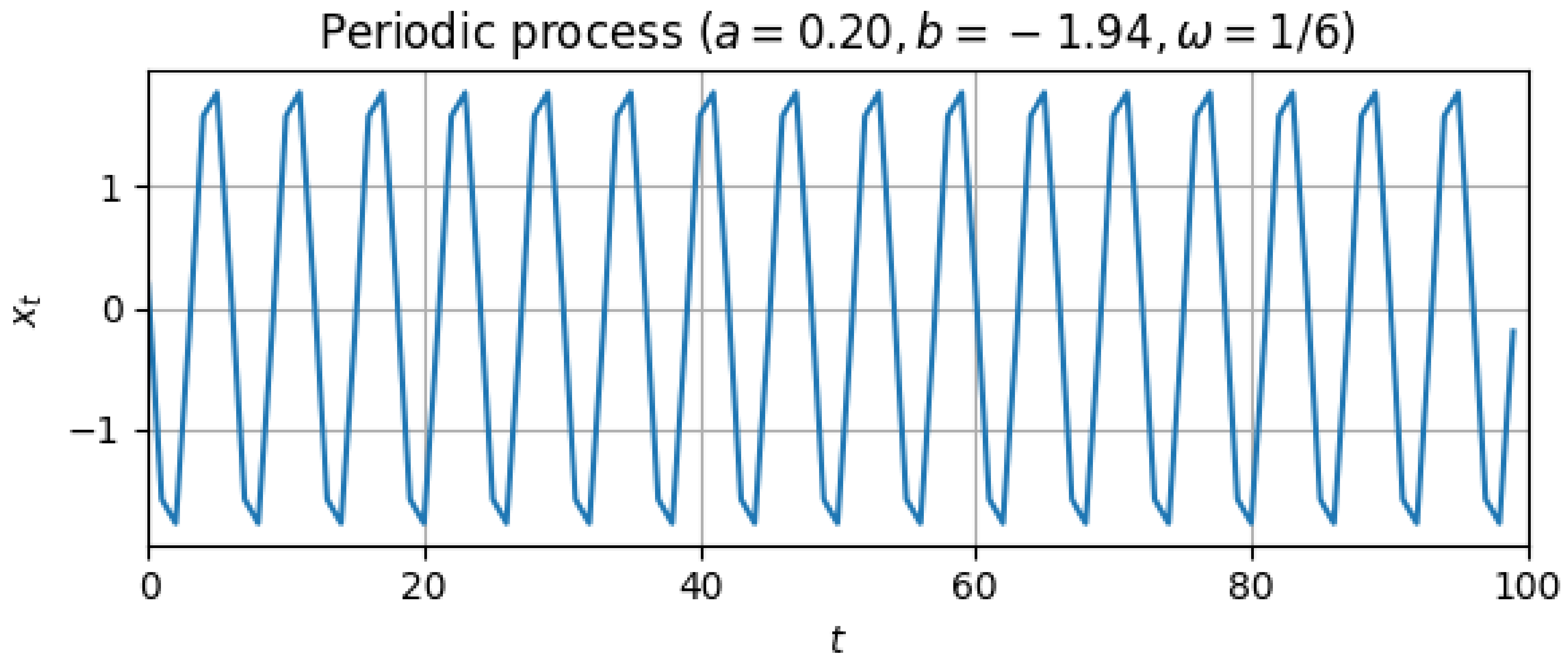
Given a_j, b_j realization of A_j, B_j then an **estimate** of σ_j^2 (i.e., $\mathrm{E}\left[R_j^2\right]$) is $\hat{\sigma}_j{}^2 = a_j^2 + b_j^2$.

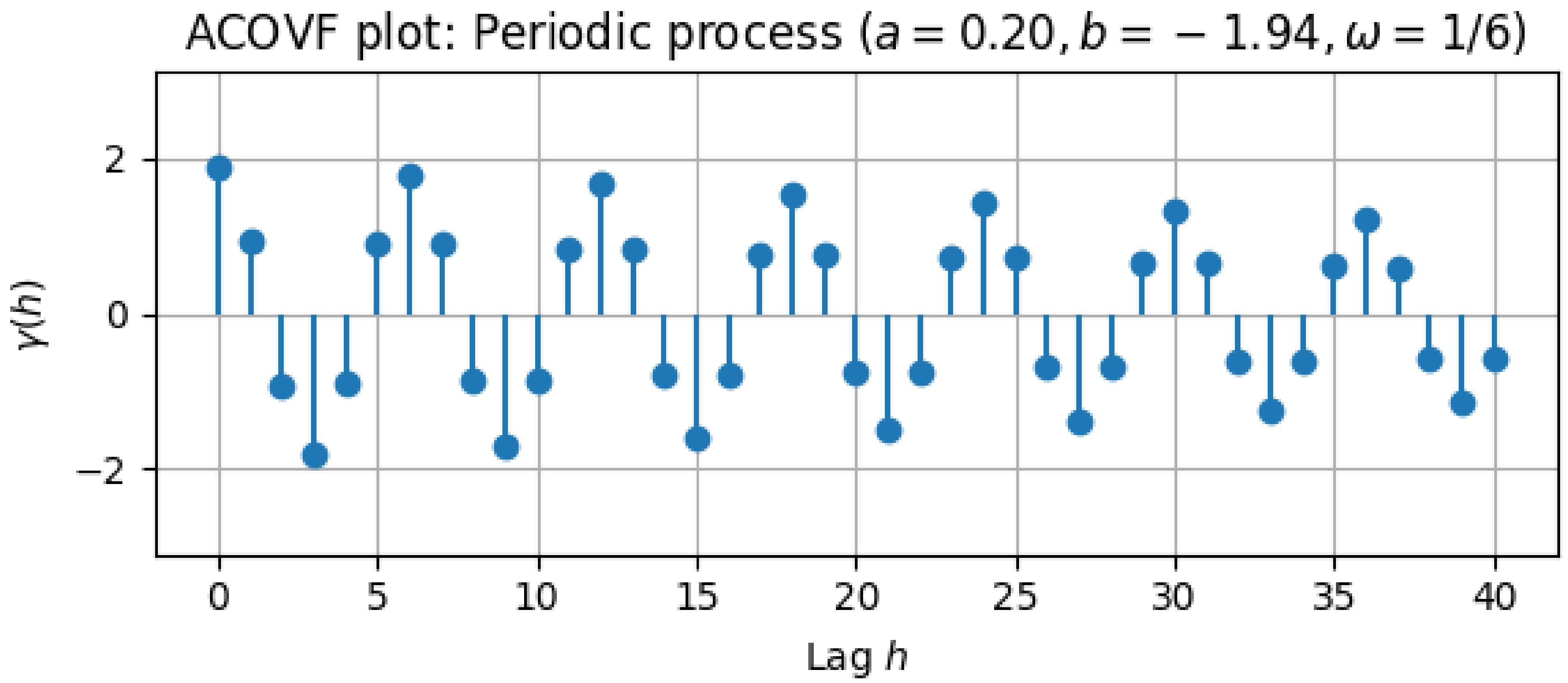
The estimate of the **total variance** is

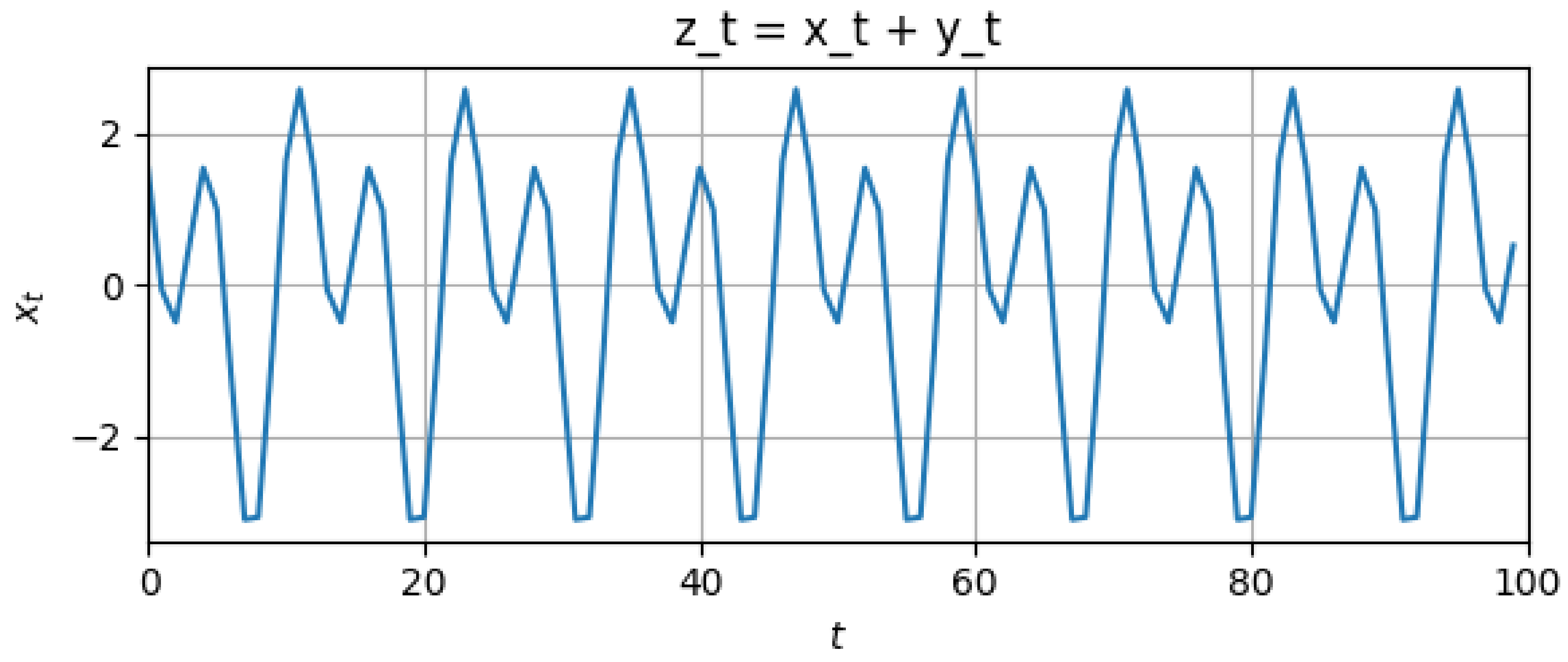
$$\hat{\gamma}_0 = \sum_{j=1}^{k} \left(a_j^2 + b_j^2 \right)$$

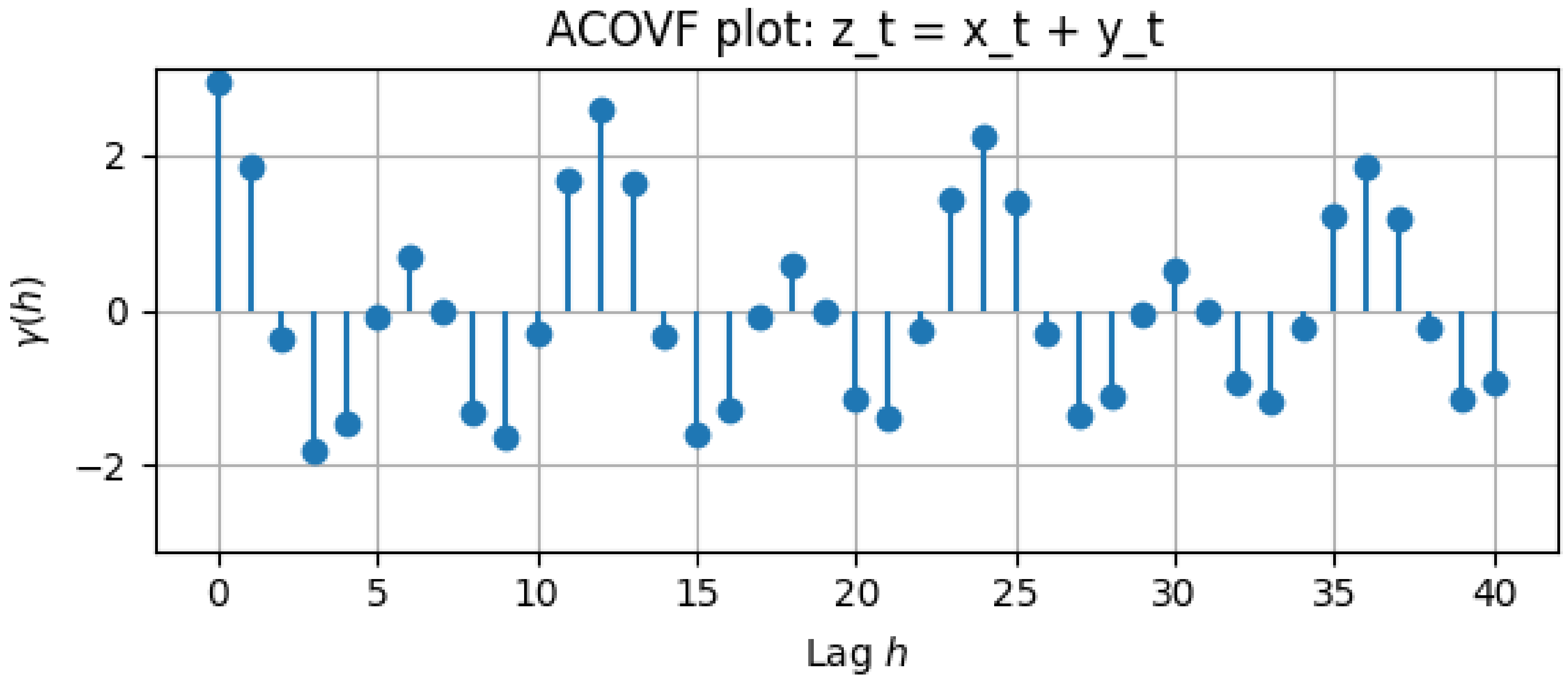

Periodic process ($a = 1.33, b = 0.72, \omega = 1/12$)
x_t
t


ACOVF plot: Periodic process ($a = 1.33, b = 0.72, \omega = 1/12$)
$\gamma(h)$
Lag h

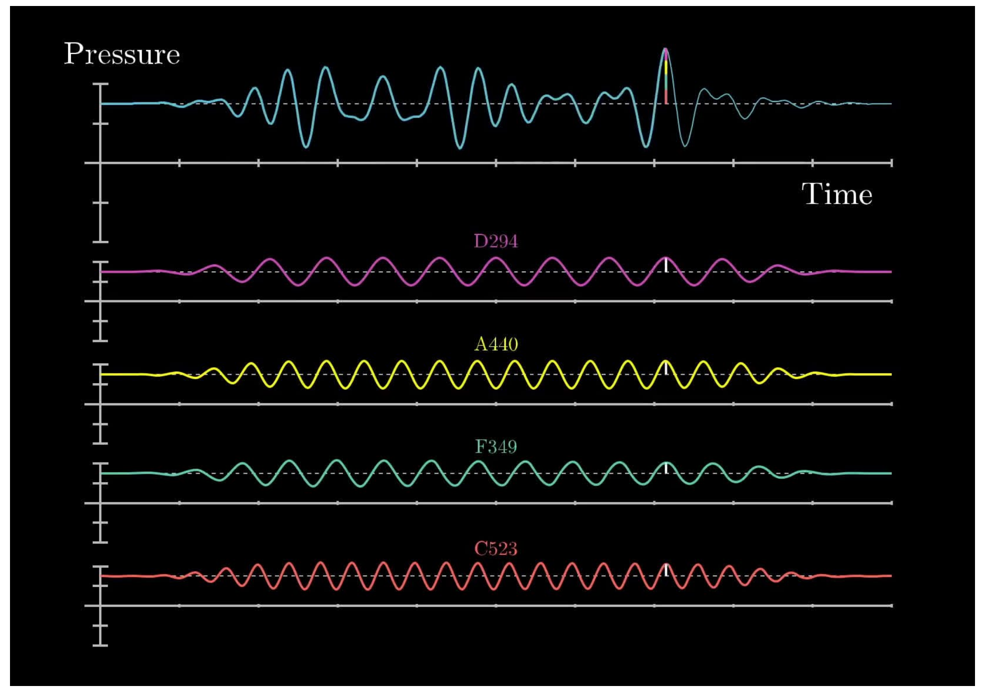

Periodic process ($a = 0.20, b = -1.94, \omega = 1/6$)
x_t
t


ACOVF plot: Periodic process ($a = 0.20, b = -1.94, \omega = 1/6$)
$\gamma(h)$
Lag h


z_t = x_t + y_t
x_t
t


ACOVF plot: z_t = x_t + y_t
$\gamma(h)$
Lag h

Recap – Fourier analysis



3Blue1Brown

A TS realization $\boldsymbol{x} = \{x_0, \dots x_{n-1}\}$ can be **exactly** represented as a sum of k sinusoids at the Fourier frequencies

$$x_t = \mu + \sum_{j=1}^{k} \left(a_j c_{t,j} + \mathrm{b}_j s_{t,j}\right), \qquad \text{with } \omega_j = j/n \text{ for } j = \mathbf{0}, 1, \dots, k = \begin{cases} n/2\,, & n \text{ is even} \\ (n-1)/2\,, & n \text{ is odd} \end{cases}$$

$$\mu = \frac{1}{n}\sum_{t=0}^{n-1} x_t = \frac{a_0}{2}, \qquad a_j = \frac{2}{n}\sum_{t=0}^{n-1} x_t c_{t,j}\,, \qquad b_j = \frac{2}{n}\sum_{t=0}^{n-1} x_t s_{t,j}$$

The **Discrete Fourier Transform** $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ are defined as

$$\mathcal{F}(\boldsymbol{x})_j = d(\omega_j) = \frac{1}{n}\sum_{t=0}^{n-1} x_t e^{-2\pi i \omega_j t}\,, \qquad \mathcal{F}^{-1}\big(\mathcal{F}(\boldsymbol{x})\big)_t = \sum_{j=0}^{n-1} d(\omega_j) e^{2\pi i \omega_j t} = x_t$$

Denoting $\mathcal{F}(\boldsymbol{x})_j$ with $d(\omega_j)$, we have shown that $d(\omega_j) = (a_j - i\, b_j)/2$.

Periodogram

Given a time series realization $\boldsymbol{x} = \{x_0, \dots x_{n-1}\}$ and its DFT $d(\omega_j)$,

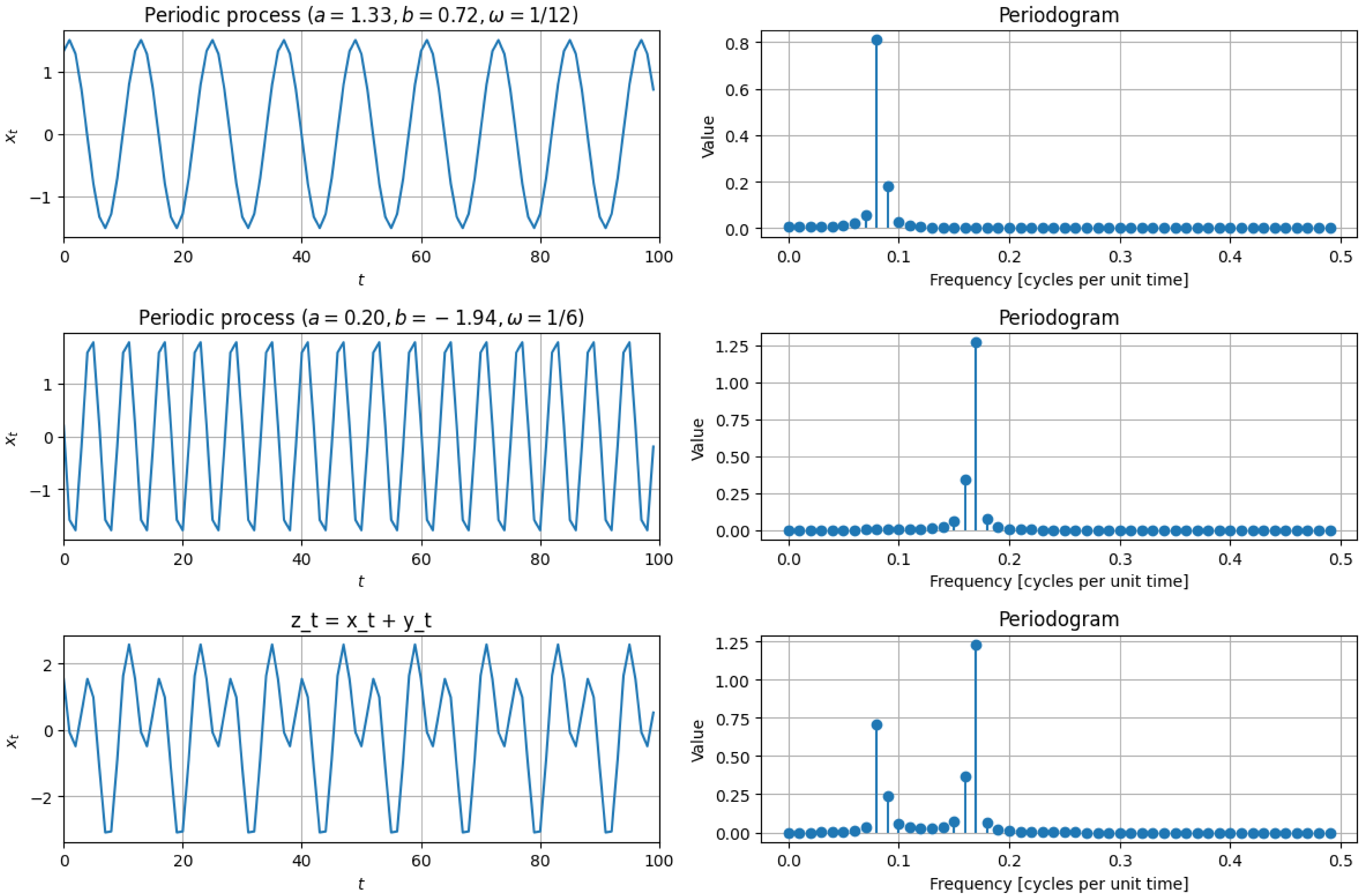
The **periodogram** is defined as the squared modulus of the DFT: $P(\omega_j) = |d(\omega_j)|^2$.

- Check your DFT implementation normalization!

The **scaled periodogram** is defined as $P^*(\omega_j) = a_j^2 + b_j^2 = \hat{\sigma}_j{}^2 = 4P(\omega_j)$

- Sample variance at each frequency component i.e., an estimate of σ_j^2.
- Squared magnitude r_j^2 of the j-th sinusoid with r_j a realization of R_j.
- Measures how strongly the oscillation with frequency ω_j is represented in the data.

For **real** x_t, we have $P(\omega_j) = P(1 - \omega_j)$ due to the **conjugate symmetry** of the DFT.


Periodic process ($a = 1.33, b = 0.72, \omega = 1/12$)
Periodogram
Periodic process ($a = 0.20, b = -1.94, \omega = 1/6$)
Periodogram
z_t = x_t + y_t
Periodogram
x_t
t
Value
Frequency [cycles per unit time]

Spectral distribution

Given a **stationary process** X_t with auto-covariance function $\gamma(h)$, there exists an unique monotonically increasing function $F(\omega)$, called the **spectral distribution function**, satisfying

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} \, dF(\omega) \quad \text{with } F(\omega) = \begin{cases} 0, & \omega \leq -\frac{1}{2} \\ \gamma(0), & \omega \geq +\frac{1}{2} \end{cases}$$

Note that the integral is a Riemann–Stieltjes integration:

- *Regular integration integrate with respect to an uniform increment dx.*
- *Here, integrate with respect to $dF(\omega)$, which allows to weight contributions non-uniformly based on changes in $F(\omega)$.*
- *Especially useful in cases where $F(\omega)$ is not smooth e.g., **discrete** case.*

$F(\omega)$ is a **cumulative distribution function**, not of probabilities, but of **variances**.

- Captures how the *total variance* of the process accumulates as you sweep through frequencies.
- $F(\infty) = \gamma(0)$ is the **total variance** of the process X_t.
- Discontinuous in the presence of deterministic seasonality.

Spectral density

If $\gamma(h)$ satisfies $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, then $F(\omega)$ is **continuous** and we can take its **derivative**

$$dF(\omega) = f(\omega)d\omega \text{ and } \gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i\omega h} f(\omega)d\omega$$

$f(\omega)$ is called the **spectral density** or **power spectrum**:

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h)e^{-2\pi i\omega h} \text{ with } -1/2 \leq \omega \leq 1/2$$

The spectral density is non-negative and symmetric $f(\omega) = f(-\omega)$.

The **total variance** (also referred as the **total power**) of the process X_t is $\gamma(0) = \int_{-1/2}^{1/2} f(\omega)d\omega$.

- The **normalized** spectral density is $f^*(\omega) = \sum_{h=-\infty}^{\infty} \rho(h)e^{-2\pi i\omega h}$.

Spectral density and autocovariance

Observe that $\gamma(h)$ and $f(\omega)$ are **Fourier transform pairs**.

Time domain: the auto-covariance expresses information in terms of **lags**.

$\gamma(h)$ quantifies the co-variance i.e., the **similarity** of TS values at different **lags**.

Frequency domain: the spectral density expresses the same information in terms of **cycles**.

$f(\omega)$ quantifies the TS variance i.e., the **power** at different **frequencies**.

Spectral density – White noise process

Stationary process composed of uncorrelated zero-mean RVs with constant finite variance σ^2.

The auto-covariance is $\gamma_h = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases}$

The spectral density is

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} = \sigma^2$$

White noise contains an equal mix of all frequencies.

- Analogy to white light, which contains all frequencies in the color spectrum at the same level of intensity.

Spectral density – MA(1) process

Consider a zero-mean stationary MA(1) process $X_t = W_t + \theta W_{t-1}$ with $W_t \sim WN\left(0, \sigma^2\right)$.

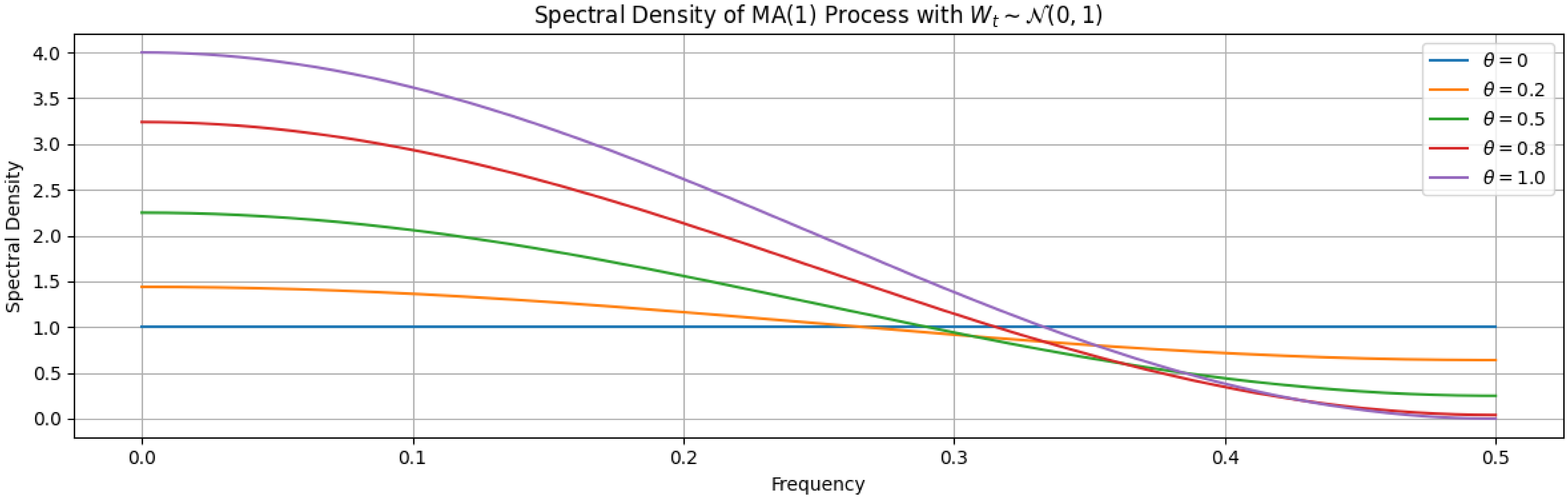
The auto-covariance is $\gamma_h = \begin{cases} \sigma^2\left(1+\theta^2\right), & h = 0 \\ \sigma^2\theta, & |h| = 1 \\ 0, & |h| > 1 \end{cases}$

The spectral density is

$$
\begin{aligned}
f(\omega) &= \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h} \\
&= \sigma^2\theta e^{2\pi i\omega} + \sigma^2\left(1+\theta^2\right)e^0 + \sigma^2\theta e^{-2\pi i\omega} \\
&= \sigma^2\left(1+\theta^2\right) + \sigma^2\theta\left(e^{2\pi i\omega} + e^{-2\pi i\omega}\right) \\
&= \sigma^2\left(1+\theta^2 + 2\theta cos(2\pi\omega)\right)
\end{aligned}
$$

(recall that $\cos(\varphi) = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\varphi} + \mathrm{e}^{-\mathrm{i}\varphi}\right)$ *with* $\mathrm{e}^{\mathrm{i}\varphi} = \cos(\varphi) + i\sin(\varphi)$*)*

Spectral density – MA(1) process



This process has a **decaying spectral density**. The **larger** the value of θ, the **steeper** the decay.

Spectral density – AR(1) process (1)

Consider a zero-mean stationary AR(1) process $X_t = \phi X_{t-1} + W_t$ with $W_t \sim WN\left(0, \sigma^2\right)$ and $|\phi| < 1$.

The auto-covariance is $\gamma_h = \sigma^2 \frac{\phi^{|h|}}{1-\phi^2}$.

The spectral density is

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h}$$

(Symmetry $\gamma_h = \gamma_{-h}$ for stationary process)

$$= \gamma(0) + \sum_{h=1}^{\infty} \gamma(h)\left(e^{2\pi i \omega h} + e^{-2\pi i \omega h}\right)$$

(recall that $\cos(\varphi) = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\varphi} + \mathrm{e}^{-\mathrm{i}\varphi}\right)$)

$$= \gamma(0) + \sum_{h=1}^{\infty} \gamma(h) 2cos(2\pi \omega h)$$

$$= \frac{\sigma^2}{1-\phi^2}\left(1 + 2\underbrace{\sum_{h=1}^{\infty} \phi^{|h|} cos(2\pi\omega h)}_{S}\right)$$

Spectral density – AR(1) process (2)

The infinite sum S can be rewritten as a **geometric series** *(recall that when $|r| < 1$, $\sum_{j=1}^{\infty} r^j = \sum_{j=0}^{\infty} r^j - 1 = \frac{1}{1-r} - 1 = \frac{r}{1-r}$)*

$$S = \sum_{h=1}^{\infty} \phi^{|h|} cos(2\pi\omega h) = \sum_{h=1}^{\infty} \phi^{|h|} Re\left(e^{-2\pi i\omega h}\right) = Re\left(\sum_{h=1}^{\infty} \left(\phi e^{-2\pi i\omega}\right)^h\right)$$

$$= Re\left(\frac{\phi e^{-2\pi i\omega}}{1-\phi e^{-2\pi i\omega}}\right) = Re\left(\frac{\phi e^{-2\pi i\omega}}{1-\phi e^{-2\pi i\omega}} \frac{1-\phi e^{2\pi i\omega}}{1-\phi e^{2\pi i\omega}}\right)$$

(Multiply with conjugate)

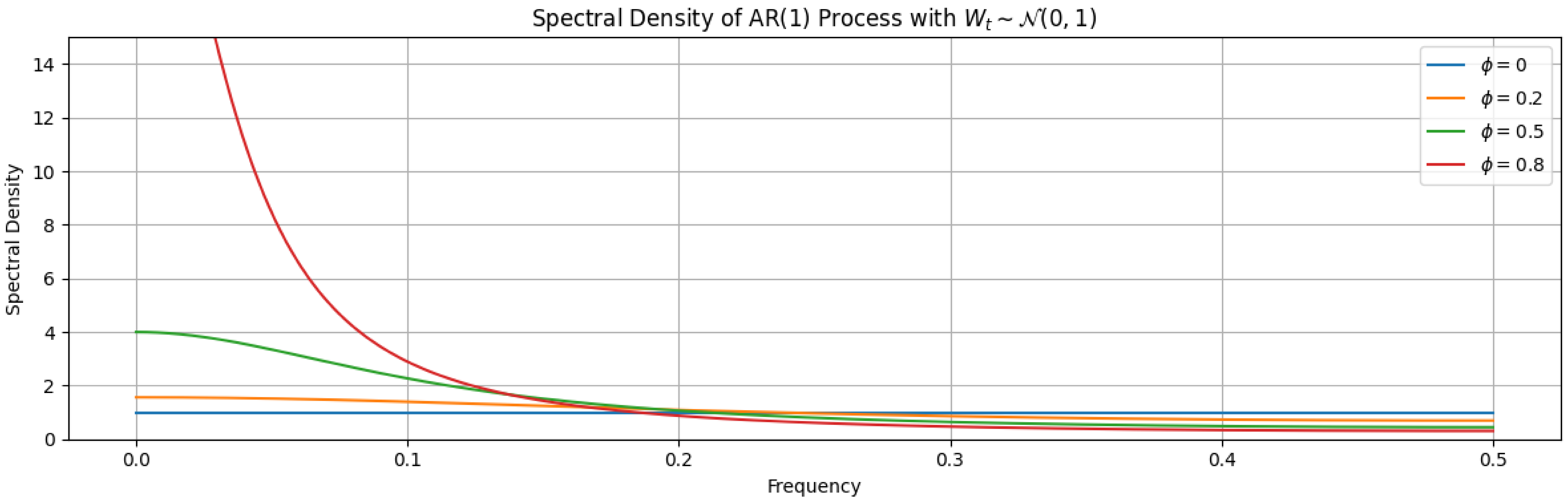
$(e^{i\varphi} + e^{-i\varphi} = 2\cos(\varphi))$

$$= Re\left(\frac{\phi(cos(2\pi\omega) - \phi) - i\,\phi sin(2\pi\omega)}{1 - 2\phi cos(2\pi\omega) + \phi^2}\right) = \frac{\phi(cos(2\pi\omega) - \phi)}{1 - 2\phi cos(2\pi\omega) + \phi^2}$$

Substituting back,

$$f(\omega) = \frac{\sigma^2}{1-\phi^2}\left(1 + 2\frac{\phi(cos(2\pi\omega) - \phi)}{1 - 2\phi cos(2\pi\omega) + \phi^2}\right) = \frac{\sigma^2}{1 - 2\phi cos(2\pi\omega) + \phi^2}$$

Spectral density – AR(1) process (2)



Spectral density is concentrated at low frequencies. The **larger** the value of ϕ, the more the density **flattens**.

Spectral density – ARMA(p,q) process

Considering a zero-mean stationary ARMA(p,q),

$$\Phi(B)X_t = \left(1 - \sum_{i=1}^{p} \phi_i B^i\right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) W_t = \Theta(B) W_t$$

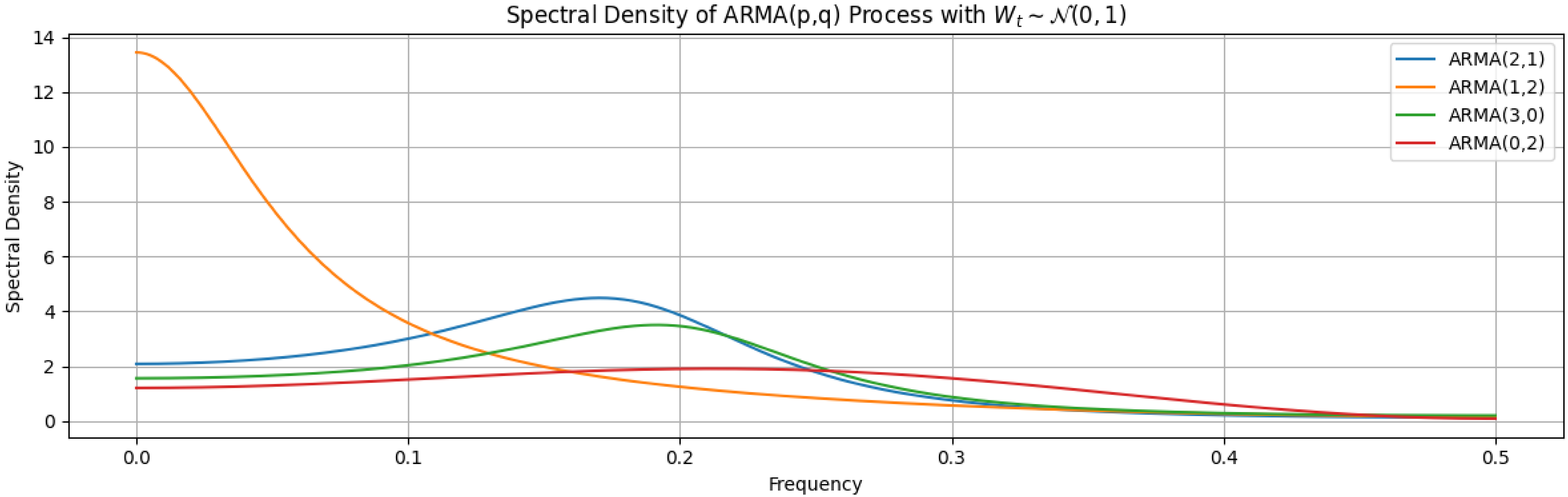
with $W_t \sim WN\left(0, \sigma^2\right)$.

The spectral density of X_t is given by

$$f(\omega) = \sigma^2 \frac{\left|\Theta\left(e^{-2\pi i\omega}\right)\right|^2}{\left|\Phi\left(e^{-2\pi i\omega}\right)\right|^2}$$

with $\Phi(z) = \left(1 - \sum_{i=1}^{p} \phi_i z^i\right)$ and $\Theta(z) = \left(1 + \sum_{i=1}^{q} \theta_i z^i\right)$

Spectral density – ARMA(p,q) process



Periodogram and spectral density (1)

The DFT can be rewritten in terms of the mean-subtracted data by decomposing x_t into its mean μ and its deviation from the mean $x_t - \mu$.

$$
\begin{aligned}
d(\omega_j) &= \frac{1}{n}\sum_{t=0}^{n-1} x_t e^{-2\pi i \omega_j t} = \frac{1}{n}\sum_{t=0}^{n-1}\big(\mu + (x_t - \mu)\big)e^{-2\pi i \omega_j t} \\
&= \frac{1}{n}\left(\mu \sum_{t=0}^{n-1} e^{-2\pi i \omega_j t} + \sum_{t=0}^{n-1}(x_t - \mu)e^{-2\pi i \omega_j t}\right) \\
&= \frac{1}{n}\sum_{t=0}^{n-1}(x_t - \mu)e^{-2\pi i \omega_j t}
\end{aligned}
$$

(orthogonality of complex exponentials for $j \neq 0$)

We can then compute the periodogram using this formulation of the DFT.

Periodogram and spectral density (2)

$$P(\omega_j) = \left|d(\omega_j)\right|^2 = d(\omega_j)\overline{d(\omega_j)}$$

$$= \frac{1}{n}\sum_{s=0}^{n-1}(x_s-\mu)e^{-2\pi i\omega_j s}\frac{1}{n}\sum_{t=0}^{n-1}(x_t-\mu)e^{2\pi i\omega_j t}$$

$$= \frac{1}{n^2}\sum_{s=0}^{n-1}\sum_{t=0}^{n-1}(x_s-\mu)(x_t-\mu)e^{-2\pi i\omega_j(t-s)}$$

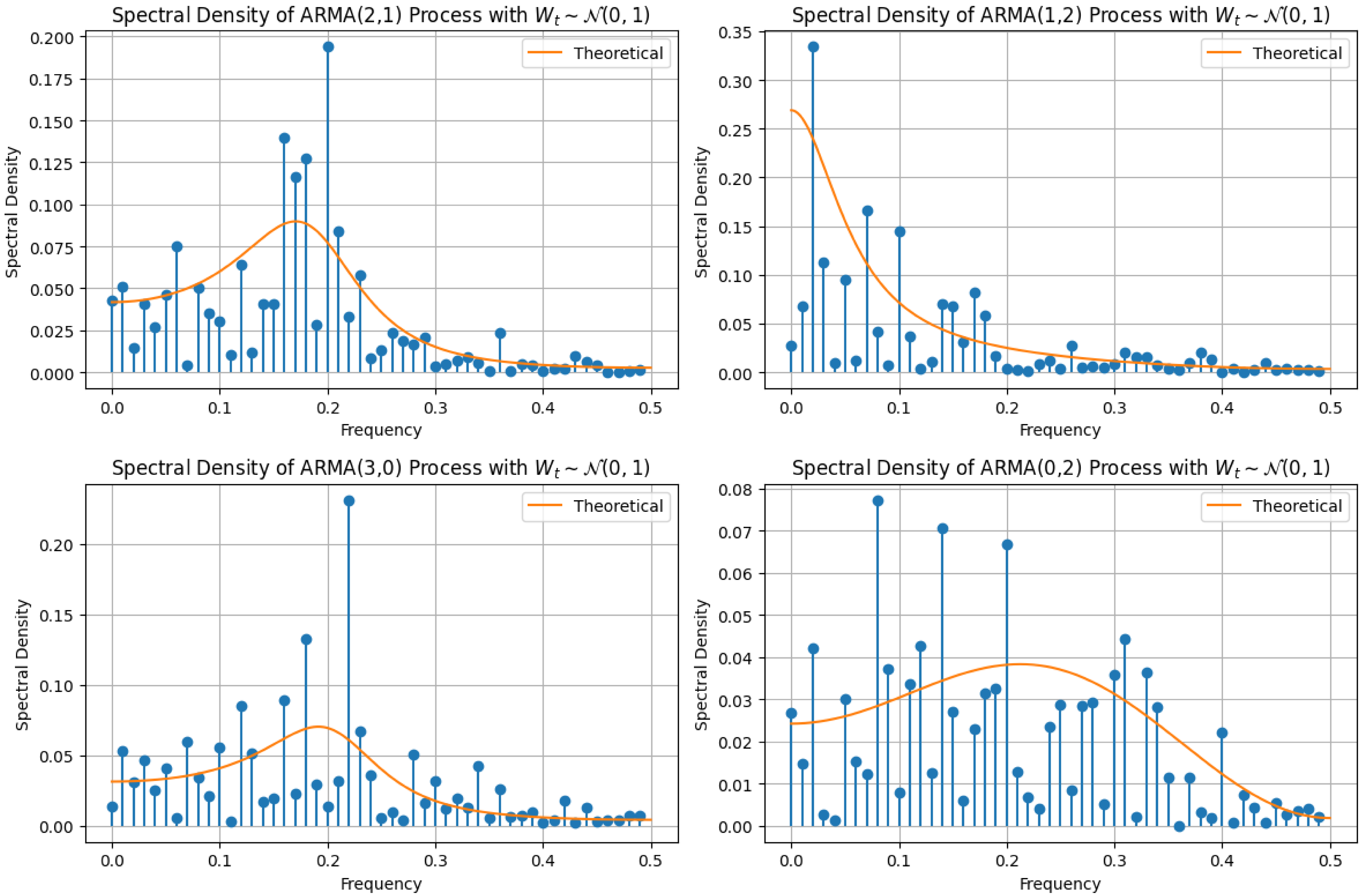
$(h = t - s)$

$$= \frac{1}{n}\sum_{h=-(n-1)}^{n-1}\frac{1}{n}\sum_{t=0}^{n-|h|}\left(x_{t+|h|}-\mu\right)(x_t-\mu)e^{-2\pi i\omega_j h}$$

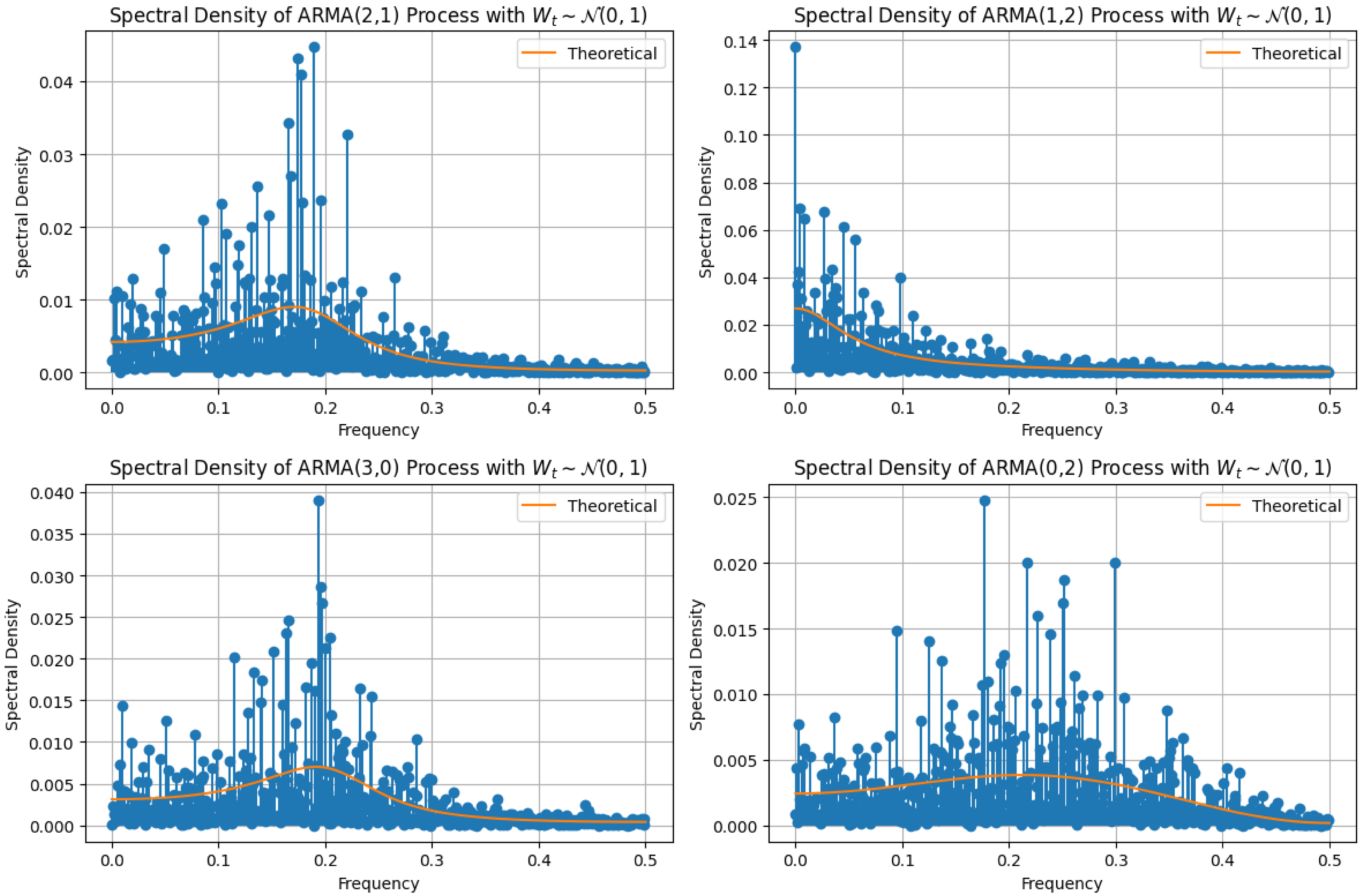
(recall that $\hat{\gamma}(h) = \frac{1}{n}\sum_{i=0}^{n-|h|}\left(x_{i+|h|}-\hat{\mu}\right)(x_i-\hat{\mu})$*)*

$$= \frac{1}{n}\sum_{h=-(n-1)}^{n-1}\hat{\gamma}(h)e^{-2\pi i\omega_j h} = \frac{1}{n}\hat{f}(\omega_j)$$

Thus, the periodogram is a rough **sample spectral density** estimator at **discrete** frequencies ω_j, $j = 0, \dots n-1$.



Spectral Density of ARMA(2,1) Process with $W_t \sim \mathcal{N}(0, 1)$
Spectral Density of ARMA(1,2) Process with $W_t \sim \mathcal{N}(0, 1)$
Spectral Density of ARMA(3,0) Process with $W_t \sim \mathcal{N}(0, 1)$
Spectral Density of ARMA(0,2) Process with $W_t \sim \mathcal{N}(0, 1)$
Theoretical
Spectral Density
Frequency



Spectral Density of ARMA(2,1) Process with $W_t \sim \mathcal{N}(0, 1)$
Theoretical
Spectral Density
Frequency
Spectral Density of ARMA(1,2) Process with $W_t \sim \mathcal{N}(0, 1)$
Theoretical
Spectral Density
Frequency
Spectral Density of ARMA(3,0) Process with $W_t \sim \mathcal{N}(0, 1)$
Theoretical
Spectral Density
Frequency
Spectral Density of ARMA(0,2) Process with $W_t \sim \mathcal{N}(0, 1)$
Theoretical
Spectral Density
Frequency

Exercise

Simulate different synthetic periodic/aperiodic signals.

- Review time-plot, ACF and auto-covariance plots.
- Compute periodogram and review spectra.
- Repeat exercise with noisy synthetic signals.

Analyze the spectral density of known processes.

- Review how parameters affect the spectral density.
- Integrate the spectral density to confirm it matches the total variance.
- Generate signals and compare periodogram with spectral density.

Study real-world time series.

- Compute periodogram and review spectra.