# Spectral Analysis

Spectral analysis is a means of investigating signal's spectral content.

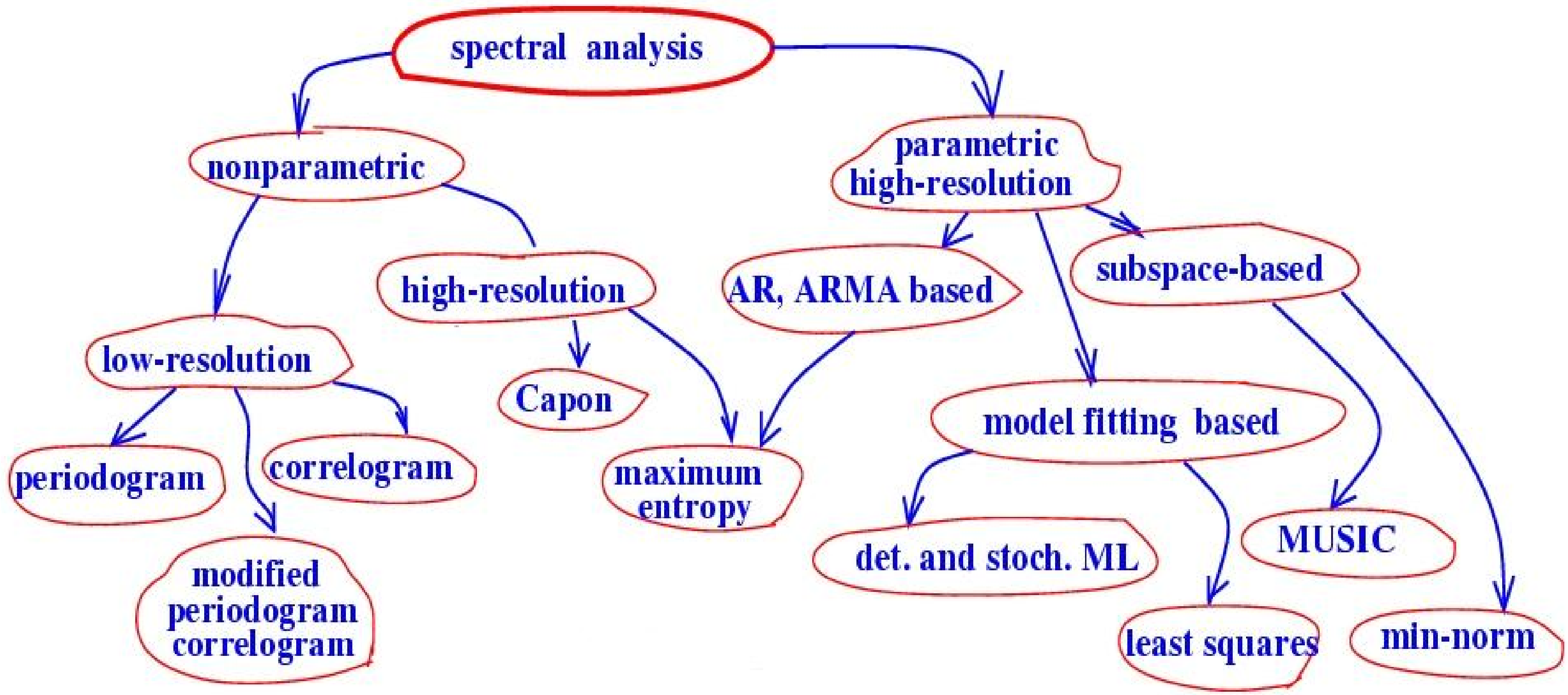
It is used in: optics, speech, sonar, radar, medicine, seizmology, chemistry, radioastronomy, etc.

There are

- *nonparametric* (classic) and
- *parametric* (modern)

methods.

# Spectral Analysis (cont.)



# Power Spectral Density (PSD) of Random Signals

Let $\{x(n)\}$ be a wide-sense stationary random signal:

$$\mathrm{E}\{x(n)\} = 0, \quad r(k) = \mathrm{E}\{x(n)x^*(n-k)\}.$$

First definition of PSD:

$$\begin{aligned} P(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r(k)e^{-j\omega k}, \\ r(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P(e^{j\omega})e^{j\omega k}d\omega. \end{aligned}$$

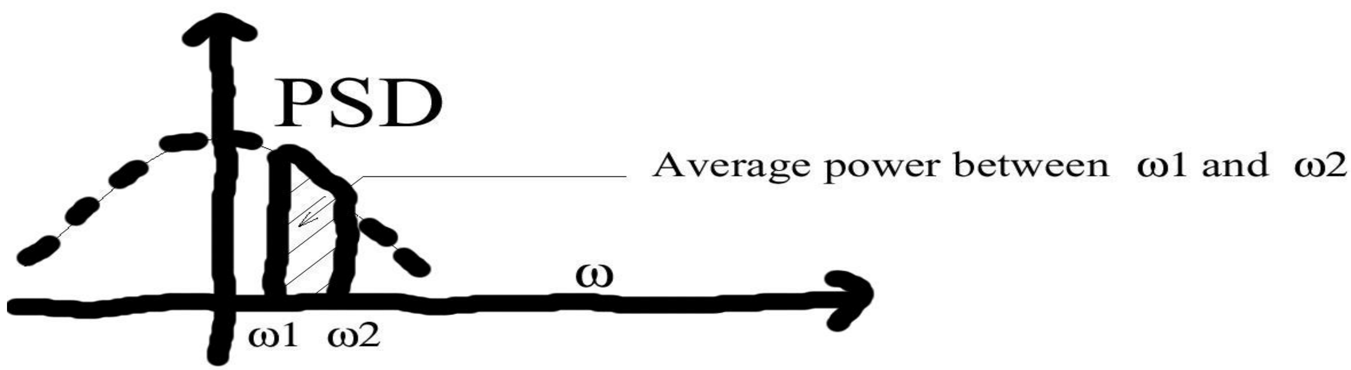
Second definition of PSD:

$$P(e^{j\omega}) = \lim_{N\to\infty} \mathrm{E}\left\{\frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-j\omega n}\right|^2\right\}.$$

Power averaged over frequency:

$$r(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{j\omega}) d\omega.$$

**Remark:** Since $r(k)$ is discrete, $P(e^{j\omega})$ is periodic, with period $2\pi$ $(\omega)$ or $1$ $(f)$.



# Power Spectral Density of Random Signals (cont.)

**Result (without proof):** First and second definitions of PSD are equivalent if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=-N+1}^{N-1} |k||r(k)| = 0$$

and also if

$$\sum_{k=-\infty}^{\infty} |r(k)| < \infty.$$

That is, $r(k)$ must decay sufficiently fast!

# Nonparametric Methods: Periodogram and Correlogram

*Periodogram* (from the second definition of PSD):

$$\widehat{P}_P(e^{j\omega}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-j\omega n}\right|^2.$$

*Correlogram* (from the first definition of PSD):

$$\widehat{P}_C(e^{j\omega}) = \sum_{k=-N+1}^{N-1} \widehat{r}(k)e^{-j\omega k}$$

where we can use either *unbiased* or *biased* estimates of $r(k)$:

**Unbiased estimate:**

$$\widehat{r}(k) = \begin{cases} \frac{1}{N-k}\sum_{i=k}^{N-1} x(i)x^*(i-k), & k \geq 0, \\ \widehat{r}^*(-k), & k < 0. \end{cases}$$

**Biased estimate:**

$$\widehat{r}(k) = \begin{cases} \frac{1}{N} \sum_{i=k}^{N-1} x(i) x^*(i-k), & k \geq 0, \\ \widehat{r}^*(-k), & k < 0. \end{cases}$$

The biased estimate is *more reliable* than the unbiased one, because it assigns lower weights to the poorer estimates of long correlation lags.

# Correlogram

The biased estimate is *asymptotically unbiased*:

$$\begin{aligned}
\lim_{N\to\infty} \mathrm{E}\left\{\widehat{r}(k)\right\} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=k}^{N-1} \mathrm{E}\left\{x(i)x^*(i-k)\right\} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{i=k}^{N-1} r(k) \\
&= \lim_{N\to\infty} \frac{N-k}{N} r(k) = r(k).
\end{aligned}$$

**Proposition.** *Correlogram computed through the biased estimate of* $r(k)$ *coincides with periodogram.*

**Proof.** Consider the *auxiliary* signal

$$y(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x(k)\epsilon(m-k),$$

where $\{x(k)\}$ are considered to be fixed constants and $\{\epsilon(k)\}$ is a unit-variance white noise:

$$r_\epsilon(m-l) = \mathrm{E}\left\{\epsilon(m)\epsilon^*(l)\right\} = \delta(m-l).$$

$y(m)$ can be viewed as the output of the filter with transfer function

$$X(e^{j\omega}) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x(k)e^{-j\omega k}.$$

Relationship between filter input and output PSD's:

$$\begin{aligned} P_y(e^{j\omega}) &= |X(e^{j\omega})|^2 P_\epsilon(e^{j\omega}) = |X(e^{j\omega})|^2 \sum_{k=-\infty}^{\infty} r_\epsilon(k) e^{-j\omega k} \\ &= |X(e^{j\omega})|^2 \sum_{k=-\infty}^{\infty} \delta(k) e^{-j\omega k} = |X(e^{j\omega})|^2 \\ &= \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j\omega n} \right|^2 = \widehat{P}_P(e^{j\omega}). \end{aligned}$$

Now, we need to prove that $P_y(e^{j\omega}) = \widehat{P}_C(e^{j\omega})$.

Observe that

$$
\begin{aligned}
r_y(k) &= \mathrm{E}\left\{y(m)y^*(m-k)\right\} \\
&= \frac{1}{N}\mathrm{E}\left\{\left[\sum_{p=0}^{N-1} x(p)\epsilon(m-p)\right]\left[\sum_{s=0}^{N-1} x^*(s)\epsilon^*(m-k-s)\right]\right\} \\
&= \frac{1}{N}\sum_{p=0}^{N-1}\sum_{s=0}^{N-1} x(p)x^*(s)\mathrm{E}\left\{\epsilon(m-p)\epsilon^*(m-k-s)\right\} \\
&= \frac{1}{N}\sum_{p=0}^{N-1}\sum_{s=0}^{N-1} x(p)x^*(s)\delta(p-k-s)
\end{aligned}
$$

$$
= \frac{1}{N}\sum_{p=k}^{N-1} x(p)x^*(p-k) = \begin{cases} \widehat{r}_x(k), & 0 \le k \le N-1, \\ 0, & k \ge N. \end{cases} \quad \text{biased}
$$

Inserting the last result in the first definition of PSD, we obtain

$$\begin{aligned} P_y(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_y(k) e^{-j\omega k} \\ &= \sum_{k=-N+1}^{N-1} \widehat{r}_x(k) e^{-j\omega k} = \widehat{P}_C(e^{j\omega}). \end{aligned}$$

□

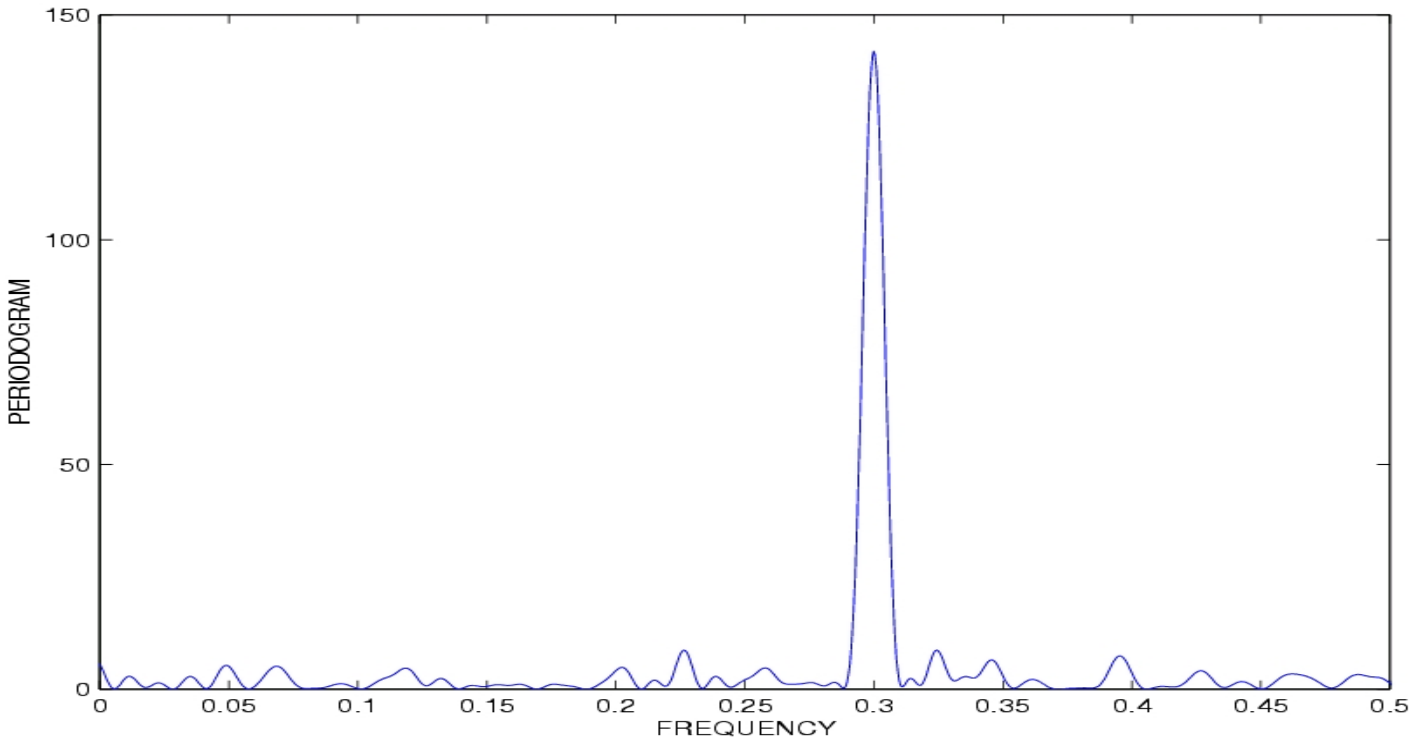
# Matlab Example

$$x(n) = A\exp(j2\pi f_s n + \phi) + \epsilon(n)$$

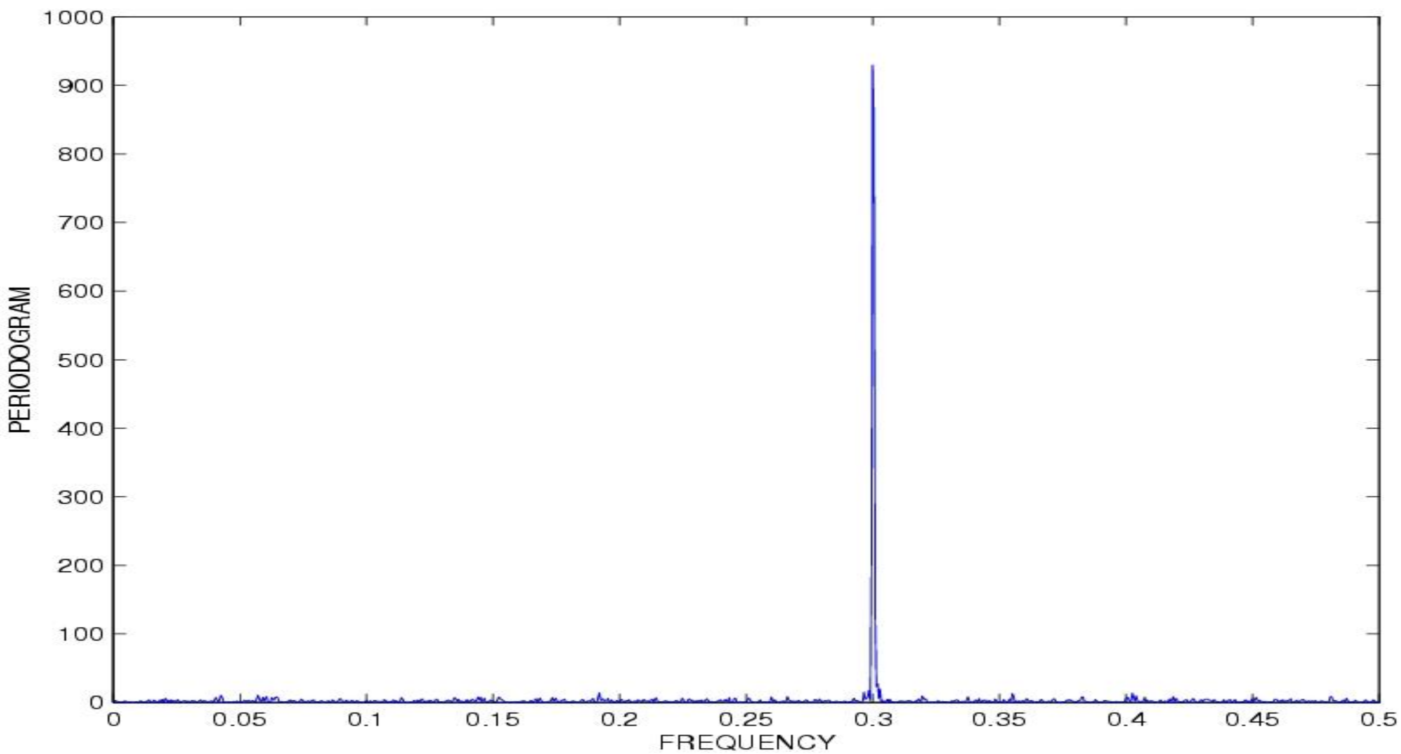
where

- $f_s = 0.3$ - discrete-time signal frequency
- $\epsilon$ - zero-mean unit-variance complex Gaussian noise
- $\phi$ - random phase uniformly distributed in $[0, 2\pi]$.

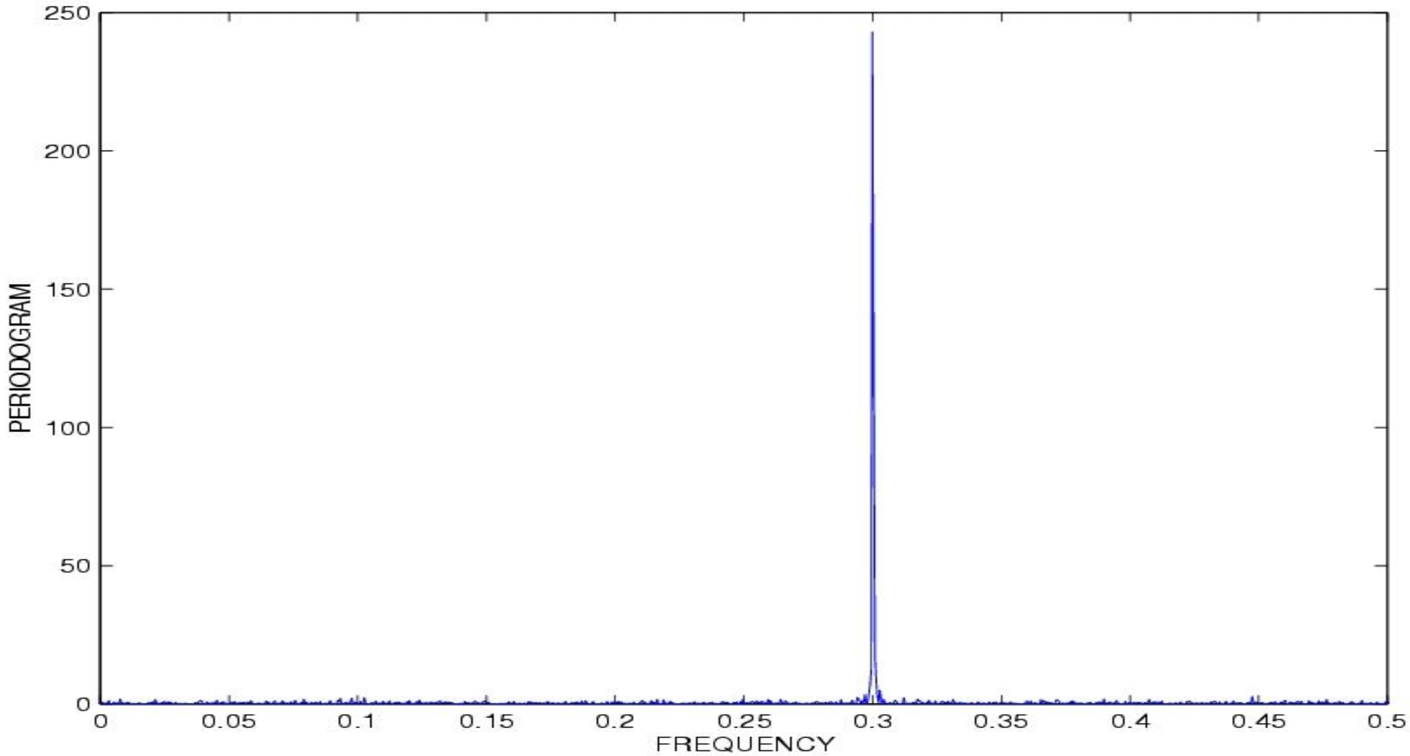
# Periodogram: $A = 1, N = 100$



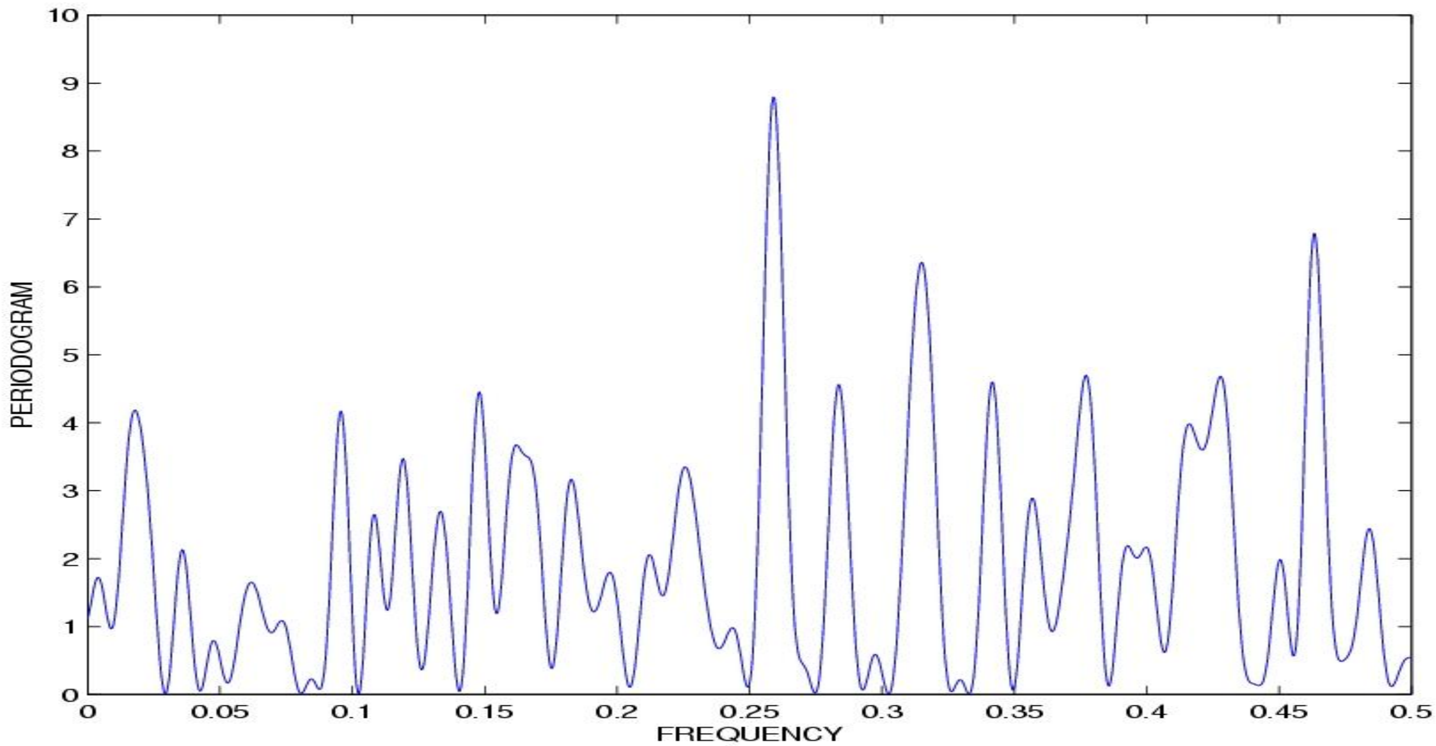
# **Periodogram:** $A = 1, N = 1000$



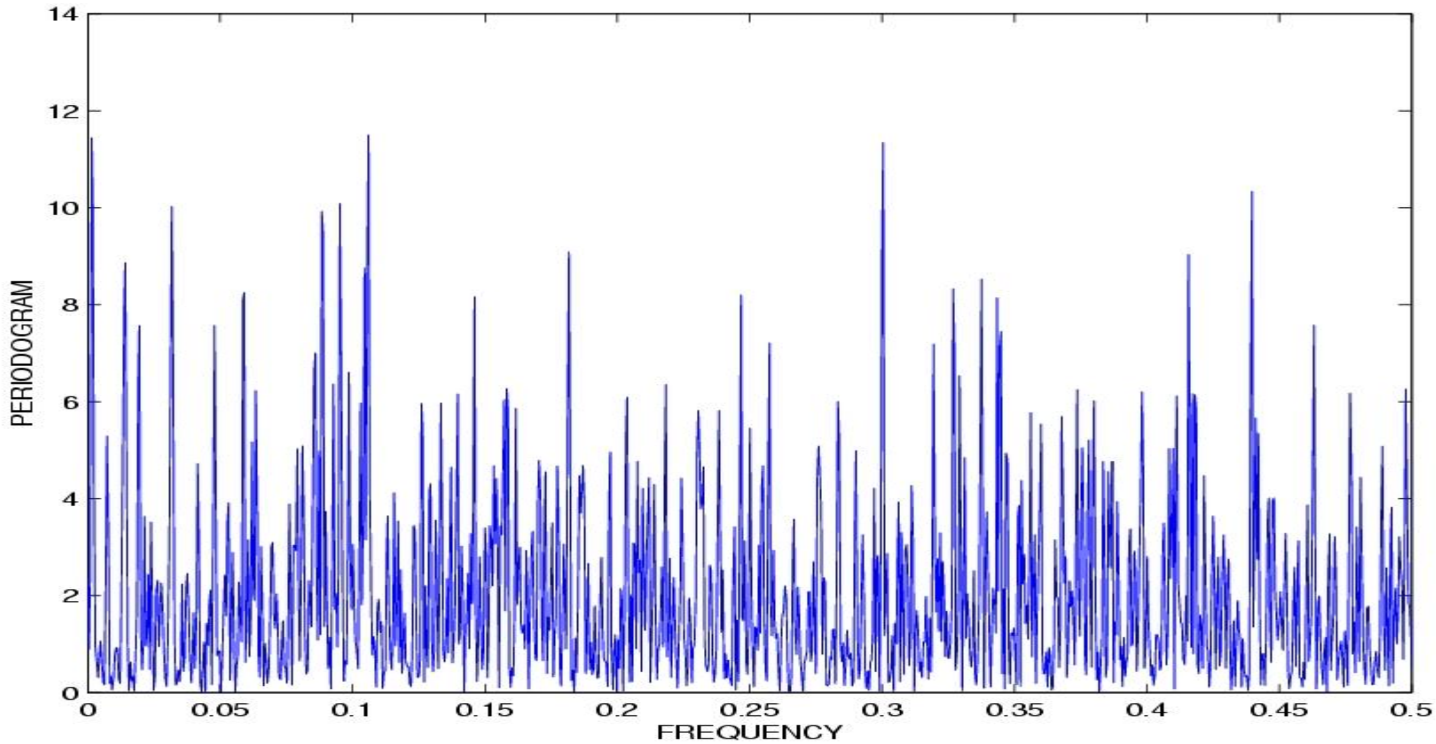
# Periodogram: $A = 1, N = 10000$



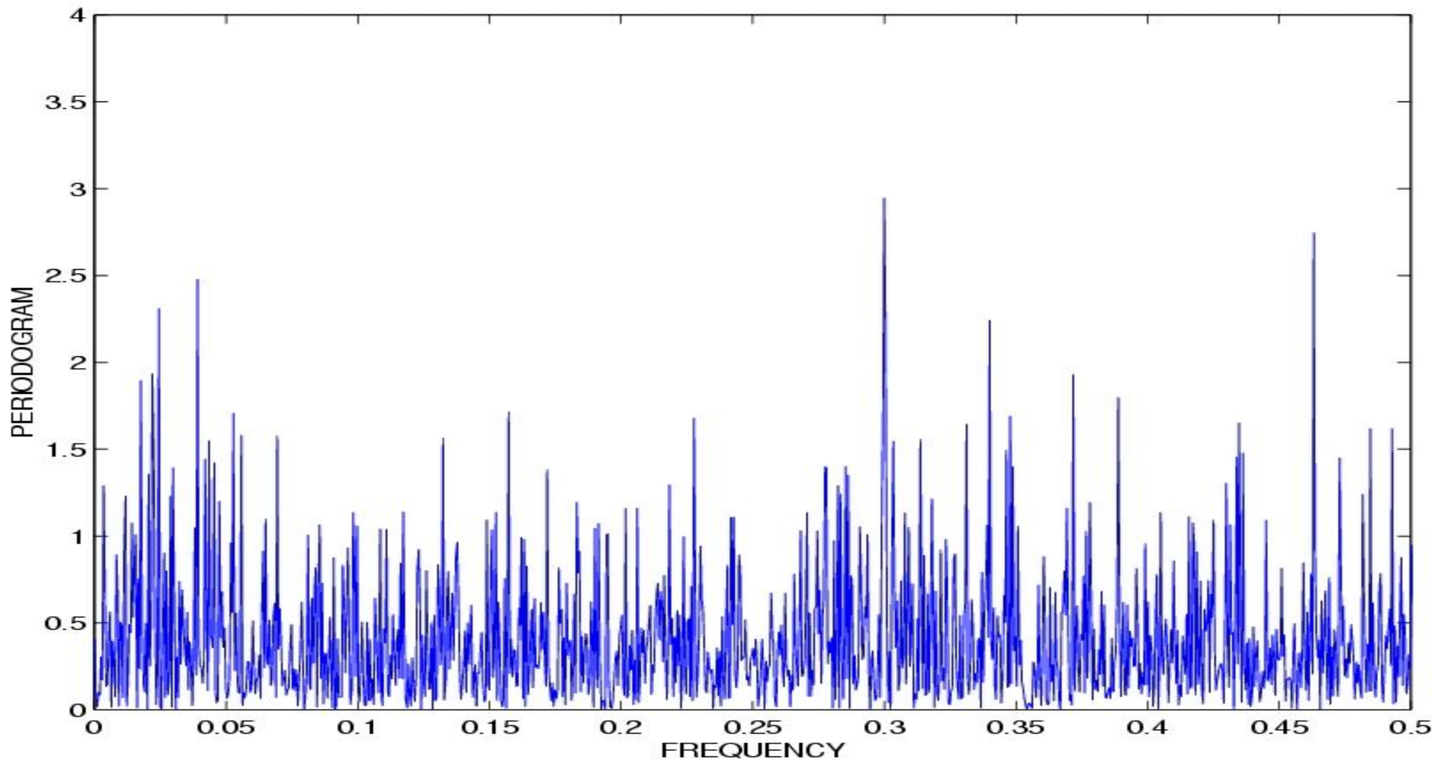
# Periodogram: $A = 0.1, N = 100$



# **Periodogram:** $A = 0.1, N = 1000$



# **Periodogram:** $A = 0.1, N = 10000$



# Statistical Analysis of Periodogram

First, consider periodogram's *bias*:

$$\mathrm{E}\{\widehat{P}_P(e^{j\omega})\} = \mathrm{E}\{\widehat{P}_C(e^{j\omega})\} = \sum_{k=-N+1}^{N-1} \mathrm{E}\{\widehat{r}(k)\}e^{-j\omega k}.$$

For the biased $\widehat{r}(k)$, we obtain

$$\mathrm{E}\{\widehat{r}(k)\} = \left(1 - \frac{k}{N}\right) r(k), \quad k \geq 0$$

and

$$\mathrm{E}\{\widehat{r}(k)\} = \mathrm{E}\{\widehat{r}^*(-k)\} = \left(1 + \frac{k}{N}\right) r(k), \quad k < 0.$$

Hence

$$
\begin{aligned}
\mathrm{E}\,\{\widehat{P}_P(e^{j\omega})\} &= \sum_{k=-N+1}^{N-1} \mathrm{E}\,\{\widehat{r}(k)\}e^{-j\omega k} \\
&= \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k)e^{-j\omega k} \\
&= \sum_{k=-\infty}^{\infty} w_{\mathrm{B}}(k)r(k)e^{-j\omega k}.
\end{aligned}
$$

where $w_{\mathrm{B}}(k)$ is a Bartlett (triangular) window:

$$
w_{\mathrm{B}}(k) = \begin{cases} 1 - \frac{|k|}{N}, & -N+1 \le k \le N-1, \\ 0, & \text{otherwise.} \end{cases}
$$

## Statistical Analysis of Periodogram (cont.)

The last equations mean

$$\begin{aligned}
\lim_{N\to\infty} \mathrm{E}\{\widehat{P}_P(e^{j\omega})\} &= \lim_{N\to\infty} \sum_{k=-N+1}^{N-1} \mathrm{E}\{\widehat{r}(k)\} e^{-j\omega k} \\
&= \sum_{k=-\infty}^{\infty} r(k) e^{-j\omega k} = P(e^{j\omega}) \Longrightarrow
\end{aligned}$$

periodogram is asymptotically unbiased estimator of PSD. For finite $N$, notice that

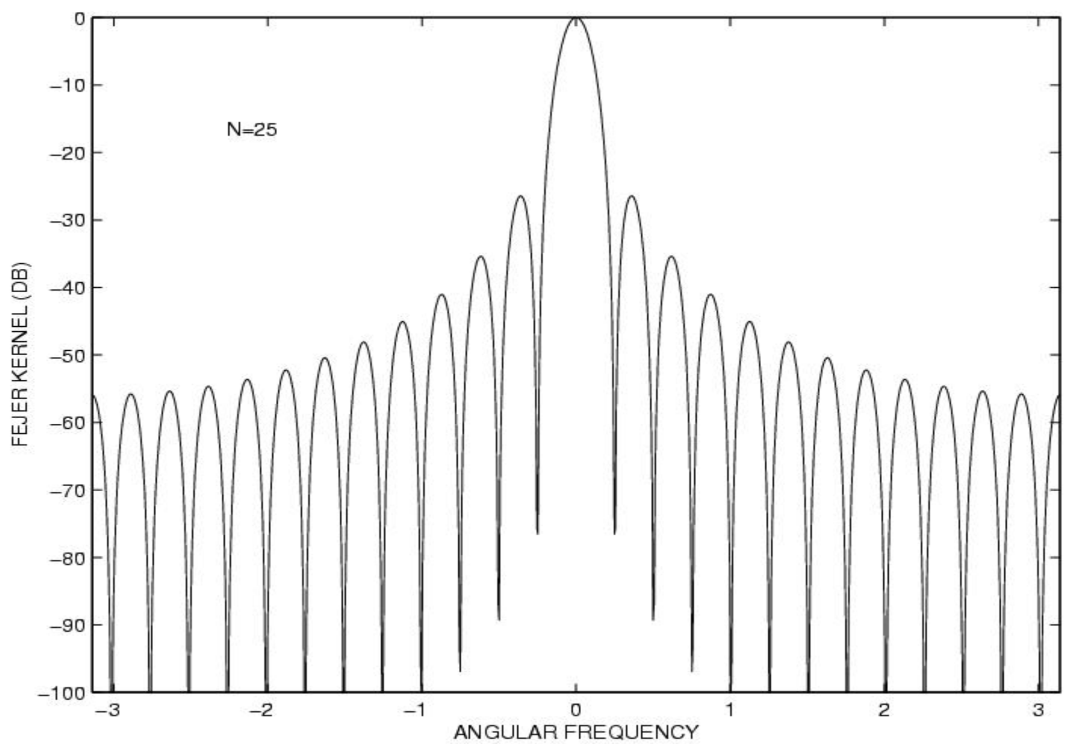
$$\mathrm{E}\{\widehat{P}_P(e^{j\omega})\} = \mathrm{DTFT}\{w_{\mathrm{B}}(k) r(k)\} \quad \Longrightarrow$$

and, hence

$$\mathrm{E}\{\widehat{P}_P(e^{j\omega})\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(e^{j\nu}) W_{\mathrm{B}}(e^{j\omega-\nu}) d\nu,$$

$$P(e^{j\omega}) = \mathrm{DTFT}\{r(k)\}, \quad W_{\mathrm{B}}(e^{j\omega}) = \mathrm{DTFT}\{w_{\mathrm{B}}(k)\}.$$

$$W_{\mathrm{B}}(e^{j\omega}) = \frac{1}{N}\Big[\frac{\sin(\omega N/2)}{\sin(\omega/2)}\Big]^2.$$



# Statistical Analysis of Periodogram (cont.)

**Remarks:**

- Frequency resolution of periodogram is approximately equal to $1/N$, because the $-3$ dB mainlobe width $W_{\mathrm{B}}$ in frequency $f$ is $\approx 1/N$.

- The mainlobe smears or smoothes the estimated spectrum,

- Sidelobes *transfer* power from the frequency bands that concentrate most of the power to bands that contain less or no power. This effect is called *leakage*.

# Statistical Analysis of Periodogram (cont.)

Now, consider periodogram variance.

*Assumption:* $x(n)$ is zero-mean circular complex Gaussian white noise:

$$\begin{aligned}
\mathrm{E}\left\{\mathrm{Re}\left[x(n)\right]\mathrm{Re}\left[x(k)\right]\right\} &= \frac{\sigma^2}{2}\delta(n-k), \\
\mathrm{E}\left\{\mathrm{Im}\left[x(n)\right]\mathrm{Im}\left[x(k)\right]\right\} &= \frac{\sigma^2}{2}\delta(n-k), \\
\mathrm{E}\left\{\mathrm{Re}\left[x(n)\right]\mathrm{Im}\left[x(k)\right]\right\} &= 0,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
\mathrm{E}\left\{x(n)x^*(k)\right\} &= \sigma^2\delta(n-k), \\
\mathrm{E}\left\{x(n)x(k)\right\} &= 0.
\end{aligned}$$

$$
\begin{aligned}
\mathrm{E}\left\{\widehat{P}_P(e^{j\omega})\right\} &= \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N}\right) r(k) e^{-j\omega k} \\
&= \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N}\right) \sigma^2 \delta(k) e^{-j\omega k} \\
&= \sigma^2 = P(e^{j\omega}).
\end{aligned}
$$

For our zero-mean circular white $x(n)$:

$$
\begin{aligned}
\mathrm{E}\left\{x(k)x^*(l)x(m)x^*(n)\right\} &= \mathrm{E}\left\{x(k)x^*(l)\right\}\mathrm{E}\left\{x(m)x^*(n)\right\} \\
&\quad +\mathrm{E}\left\{x(k)x^*(n)\right\}\mathrm{E}\left\{x(m)x^*(l)\right\} \\
&= \sigma^4[\delta(k-l)\delta(m-n) + \delta(k-n)\delta(m-l)].
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}\,\{\widehat{P}_P(e^{j\omega_1})\widehat{P}_P(e^{j\omega_2})\} \\
= \quad & \mathrm{E}\left\{ \underbrace{\frac{1}{N}\Big(\sum_{k=0}^{N-1} x(k)e^{-j\omega_1 k}\Big)\Big(\sum_{l=0}^{N-1} x^*(l)e^{j\omega_1 l}\Big)}_{\widehat{P}_P(e^{j\omega_1})} \right. \\
\times \quad & \left. \underbrace{\frac{1}{N}\Big(\sum_{m=0}^{N-1} x(m)e^{-j\omega_2 m}\Big)\Big(\sum_{n=0}^{N-1} x^*(n)e^{j\omega_2 n}\Big)}_{\widehat{P}_P(e^{j\omega_2})} \right\}.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\{\widehat{P}_P(e^{j\omega_1})\widehat{P}_P(e^{j\omega_2})\} \quad &= \quad \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} \\
&\qquad \mathrm{E}\{x(k)x^*(l)x(m)x^*(n)\}e^{-j\omega_1(k-l)}e^{-j\omega_2(m-n)} \\
&= \quad \frac{\sigma^4}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}[\delta(k-l)\delta(m-n) \\
&+ \quad \delta(k-n)\delta(m-l)]e^{-j\omega_1(k-l)}e^{-j\omega_2(m-n)} \\
&= \quad \sigma^4 + \frac{\sigma^4}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} e^{-j(\omega_1-\omega_2)(k-l)} \\
&= \quad \sigma^4 + \frac{\sigma^4}{N^2}\sum_{k=0}^{N-1} e^{-j(\omega_1-\omega_2)k}\sum_{l=0}^{N-1} e^{j(\omega_1-\omega_2)l} \\
&= \sigma^4 + \frac{\sigma^4}{N^2}\Big[\frac{1-e^{-jN(\omega_1-\omega_2)}}{1-e^{-j(\omega_1-\omega_2)}}\Big]\Big[\frac{1-e^{jN(\omega_1-\omega_2)}}{1-e^{j(\omega_1-\omega_2)}}\Big] \\
&= \quad \sigma^4 + \frac{\sigma^4}{N^2}\Big\{\frac{\sin[(\omega_1-\omega_2)N/2]}{\sin[(\omega_1-\omega_2)/2]}\Big\}^2.
\end{aligned}
$$

# Statistical Analysis of Periodogram (cont.)

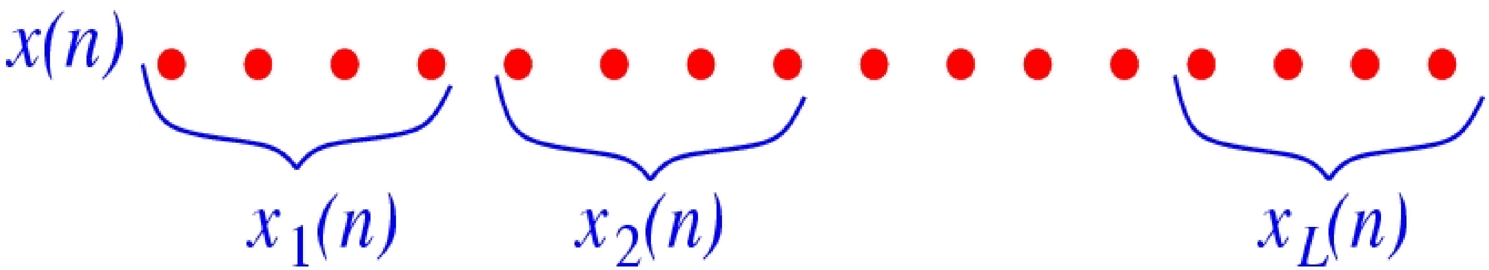
$$\lim_{N\to\infty} \mathrm{E}\{\widehat{P}_P(e^{j\omega_1})\widehat{P}_P(e^{j\omega_2})\}=P(e^{j\omega_1})P(e^{j\omega_2}) + P^2(e^{j\omega_1})\delta(\omega_1-\omega_2) \Rightarrow$$

$$\begin{aligned} &\lim_{N\to\infty} \mathrm{E}\{[\widehat{P}_P(e^{j\omega_1}) - P(e^{j\omega_1})][\widehat{P}_P(e^{j\omega_2}) - P(e^{j\omega_2})]\} \\ &= \begin{cases} P^2(e^{j\omega_1}), & \omega_1 = \omega_2, \\ 0, & \omega_1 \neq \omega_2. \end{cases} \end{aligned}$$

The variance of periodogram *cannot* be reduced by taking longer observation interval ($N \to \infty$). Thus, periodogram is a poor estimate of the PSD $P(e^{j\omega})$!

# Refined Periodogram- and Correlogram-based Methods

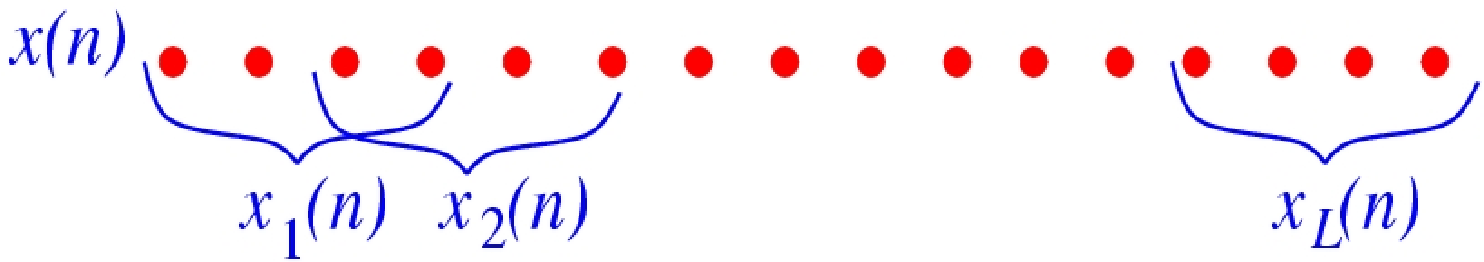
*Refined periodogram* Bartlett's method (**8.2.4 in Hayes)**:




Based on dividing the original sequence into $L = N/M$ nonoverlapping sequences of length $M$, computing periodogram for each *subsequence,* and averaging the result:

$$\widehat{P}_{\mathrm{B}}(e^{j\omega}) = \frac{1}{L}\sum_{l=1}^{L}\widehat{P}_l(e^{j\omega}), \quad \widehat{P}_l(e^{j\omega}) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x_l(n)e^{-j\omega n}\right|^2.$$

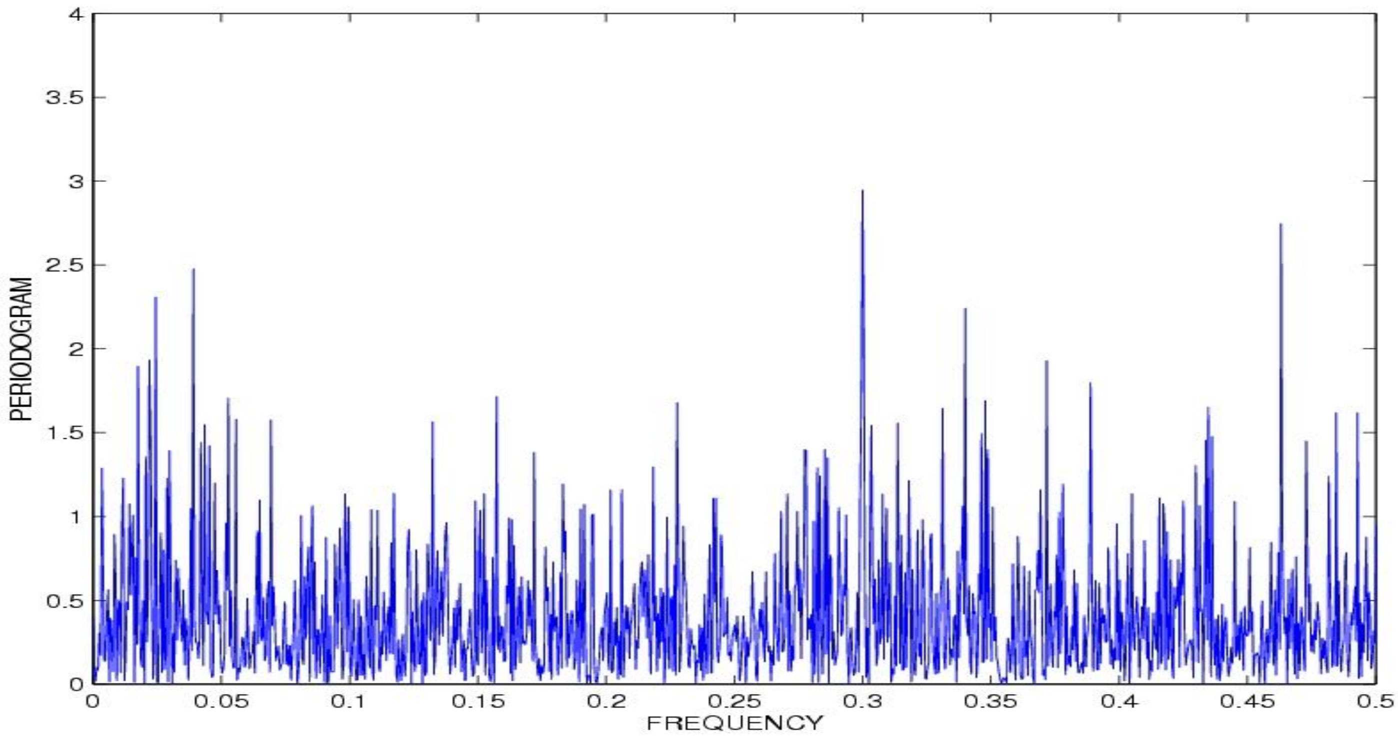
*Further Refinements of periodogram* **(Welch's method, 8.2.5 in Hayes)**:



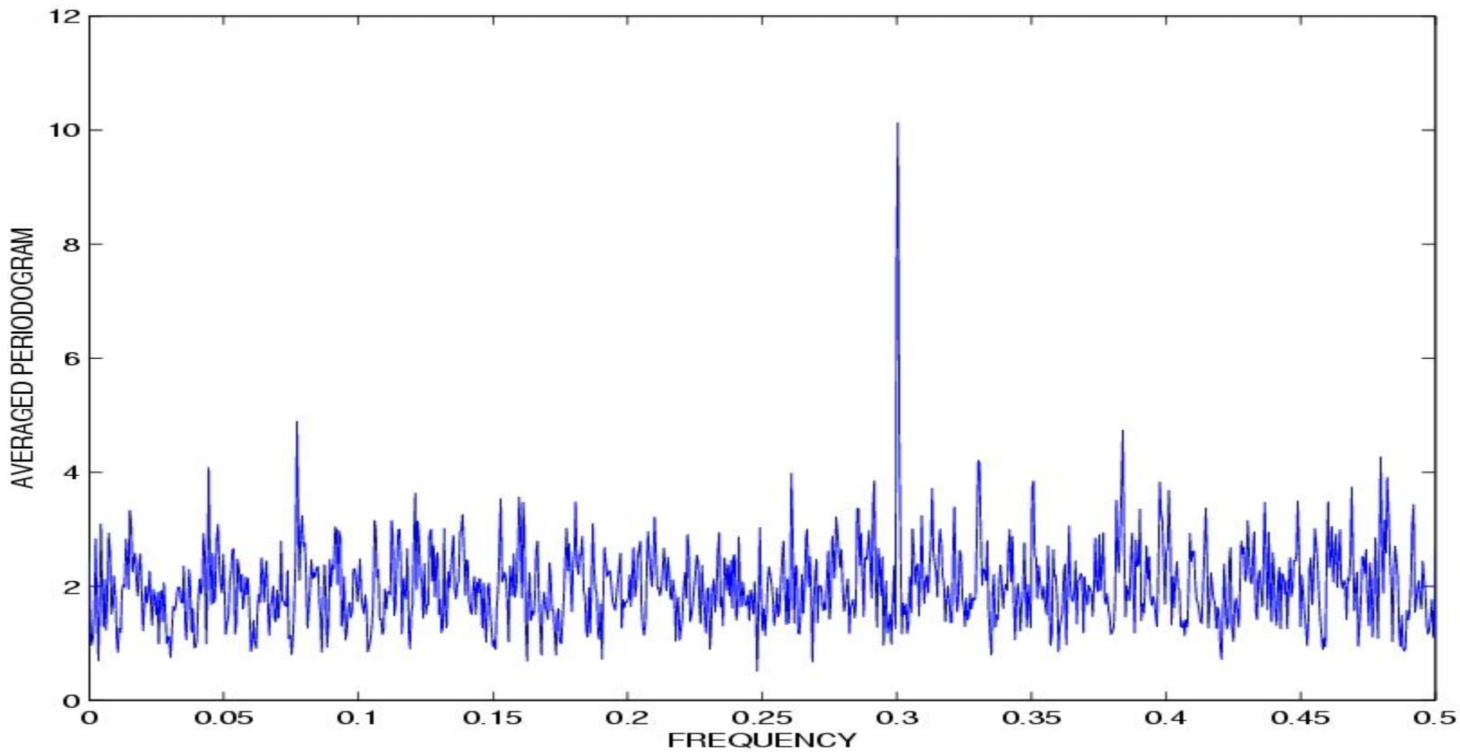

Welch's method refines the Bartlett's periodogram by:

- using *overlapping* subsequences,
- *windowing* of each subsequence.

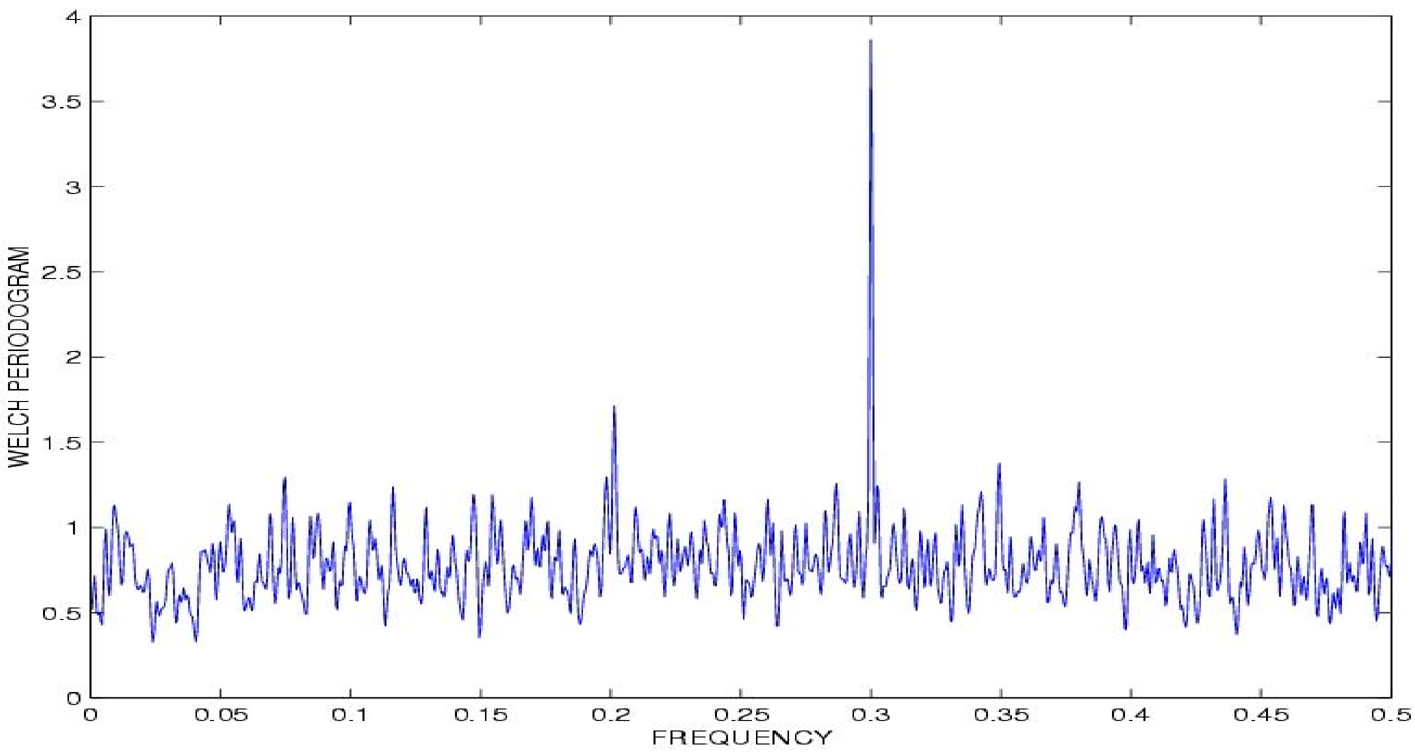
# Continue Matlab Example
# Convemtional Periodogram: $A = 0.1, N = 10000$



# **Averaged Periodogram:** $A = 0.1, N = 10000, M = 1000$



# **Welch Periodogram:** $A = 0.1, N = 10000, M = 1000$ **with 2/3 Overlap and Hamming Window**



# Refined Correlogram (Blackman-Tukey method, 8.2.6 in Hayes):

- $\widehat{r}(k)$ is a poor estimate of higher lags $k$. Hence, truncate it (use $M \ll N$ points).

- Use some *lag window*:

$$\widehat{P}_{\mathrm{BT}}(e^{j\omega}) = \sum_{k=-M+1}^{M-1} w(k)\widehat{r}(k)e^{-j\omega k}.$$

  Hence

$$\widehat{P}_{\mathrm{BT}}(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(e^{j(\omega-\nu)})\widehat{P}_{\mathrm{P}}(e^{j(\omega-\nu)})d\nu,$$

  i.e. *frequency smoothing* of the periodogram.

# High-resolution Nonparametric Methods (8.3 in Hayes)

Consider FIR filter with the impulse response $h^*(0), \ldots, h^*(N-1)$ and the output is

$$y(k) = \sum_{n=0}^{N-1} h^*(n)x(k-n) = \boldsymbol{h}^H \boldsymbol{x}(k).$$

The output power:

$$\begin{aligned} \mathrm{E}\{|y(k)|^2\} &= \mathrm{E}\{|\boldsymbol{h}^H \boldsymbol{x}(k)|^2\} \\ &= \boldsymbol{h}^H \mathrm{E}\{\boldsymbol{x}(k)\boldsymbol{x}^H(k)\}\boldsymbol{h} \\ &= \boldsymbol{h}^H R \boldsymbol{h}. \end{aligned}$$

## Filter frequency response

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h^*(n)e^{-j\omega n} = \boldsymbol{h}^H \boldsymbol{a}(\omega),$$

where

$$\boldsymbol{a}(\omega) = \begin{bmatrix} 1 \\ e^{-j\omega} \\ \vdots \\ e^{-j(N-1)\omega} \end{bmatrix}.$$

# High-resolution Nonparametric Methods: Capon

*The key idea of the Capon method:* let us "steer" our filter towards a particular frequency $\omega$ and try to reject the signals at all remaining frequencies:

$$\min_{\boldsymbol{h}} \mathrm{E}\{|y(k)|^2\} \quad \text{subject to } H(e^{j\omega}) = 1 \qquad \Longrightarrow$$

$$\min_{\boldsymbol{h}} \boldsymbol{h}^H R \boldsymbol{h} \quad \text{subject to} \quad \boldsymbol{h}^H \boldsymbol{a}(\omega) = 1.$$

$$Q(\boldsymbol{h}) = \boldsymbol{h}^H R \boldsymbol{h} + \lambda[1 - \boldsymbol{h}^H \boldsymbol{a}(\omega)] + \lambda^*[1 - \boldsymbol{a}(\omega)^H \boldsymbol{h}] \quad \Longrightarrow$$

$$\nabla Q = R\boldsymbol{h} - \lambda \boldsymbol{a}(\omega) = 0 \quad \Longrightarrow \boldsymbol{h}_{\mathrm{opt}} = \lambda R^{-1} \boldsymbol{a}(\omega)$$

note *similarity* with the Yule-Walker equations!

Substituting back into the constraint equation $\boldsymbol{h}^H \boldsymbol{a}(\omega) = 1$, we obtain

$$\boldsymbol{h}^H \boldsymbol{a}(\omega) = \lambda^* \boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{a}(\omega) = 1 \Longrightarrow \lambda = \frac{1}{\boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{a}(\omega)}.$$

Hence, the analytic solution is given by

$$\boldsymbol{h}_{\mathrm{opt}} = \frac{1}{\boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{a}(\omega)} R^{-1} \boldsymbol{a}(\omega).$$

# High-resolution Nonparametric Methods: Capon (cont.)

$$\begin{aligned}
P_{\text{CAPON}}(e^{j\omega}) &= \mathrm{E}\,\{|y(k)|^2\}|_{\boldsymbol{h}=\boldsymbol{h}_{\text{opt}}} \\
&= \boldsymbol{h}_{\text{opt}}^H R \boldsymbol{h}_{\text{opt}} \\
&= \frac{\boldsymbol{a}^H(\omega) R^{-1} R R^{-1} \boldsymbol{a}(\omega)}{[\boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{a}(\omega)]^2} \\
&= \frac{1}{\boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{a}(\omega)}.
\end{aligned}$$

This spectrum is still *impractical* because it includes the true covariance matrix $R$. Take its sample estimate

$$\widehat{P}_{\text{CAPON}}(e^{j\omega}) = \frac{1}{\boldsymbol{a}^H(\omega) \widehat{R}^{-1} \boldsymbol{a}(\omega)}.$$

# AR Spectral Estimation

**Idea:** Find the complex AR coefficients of the process and substitute them to the AR spectrum:

$$P_{\mathrm{AR}} = \frac{\sigma^2}{|A(e^{j\omega})|^2} = \frac{\sigma^2}{|\boldsymbol{c}^H \boldsymbol{a}(\omega)|^2}$$

where $\boldsymbol{c} = [1, a_1, \ldots, a_{N-1}]^H$. Recall that, according to the Yule-Walker equations:

$$\boldsymbol{c} = \sigma^2 R^{-1} \boldsymbol{e}_1$$

where $\boldsymbol{e}_1 = [1, 0, 0, \ldots, 0]^T$. Hence, omitting $\sigma^2$:

$$P_{\mathrm{AR}}(\omega) = \frac{1}{|\boldsymbol{a}^H(\omega) R^{-1} \boldsymbol{e}_1|^2}.$$

**Maximum entropy spectral estimation:** given covariance function measured at $N$ lags, extrapolate it out of the measurement interval by maximizing the entropy of the random process. Entropy of a Gaussian process can be written as (Burg):

$$\mathcal{H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln P(e^{j\omega})d\omega.$$

*Burg's method:* $\max \mathcal{H}$ subject to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P(e^{j\omega})e^{j\omega n}d\omega = \widehat{r}(n), \quad n = 0, 1, \ldots, N-1.$$

This was shown to give the AR spectral estimate!

# Digression: Entropy

Let the sample space for a dicrete RV $x$ be $x_1, \ldots, x_n$. The entropy $H(x)$ is proportional to

$$H(x) \sim -\sum_{i=1}^{n} p(x_i) \ \ln p(x_i).$$

where $p(x_i) = \mathrm{Prob}(x = x_i)$: For continuous RV

$$H(x) \sim -\int_{-\infty}^{\infty} f_x(x) \ln f_x(x) dx,$$

where $f_x(x)$ is the pdf of $x$.