

System Identification

GianAndrea Müller

October 31, 2018

CONTENTS

1 System Identification	2	5.3 Random Signal	4	8 Windowing and Input Signals	10
2 Definitions	2	5.3.1 Autocovariance (Random Signal)	4	8.1 Frequency domain smoothing in matlab	11
3 Frequency Domain Methods	2	5.3.2 Power Spectral Density (Random Signal)	5	8.2 Time domain windows	11
3.1 Sampling Operation	2	5.3.3 Cross-Covariance (Random Signal)	5	8.2.1 Time domain smoothing in matlab	11
3.2 Fourier Series of Periodic Signals	2	5.3.4 Cross Power Spectral Density (Random Signal)	5	8.2.2 Window characteristics	12
4 Spectral Estimation	2	5.4 Finite Length Signal	5	8.3 Input Signals	12
4.1 Sinusoidal correlation methods	3	5.4.1 Discrete-Fourier Transform (Finite Length Signal)	5	8.3.1 PRBS	12
5 Frequency Domain Methods	3	5.4.2 Periodogram (Finite Length Signal)	5	8.3.2 PRBS in matlab	13
5.1 Finite Energy Signal	3	6 ETFE	6	8.3.3 Multi-sinusoidal signals	13
5.1.1 Energy Spectral Density (Finite Energy Signal)	3	6.1 Input-output relationship	6	9 Residual Spectra, Coherency, Aperiodicity, Offsets and Drifts	13
5.1.2 Autocorrelation (Finite Energy Signal)	3	6.2 Periodic input case	6	9.1 Residual Spectrum	13
5.2 Discrete Periodic Signal	4	6.3 Spectral Transformations	7	9.1.1 Estimating $\phi_v(e^{j\omega_n})$	13
5.2.1 Discrete Fourier Series (Discrete Periodic Signal)	4	6.4 Approaches to spectral estimation	7	9.2 Time-domain data windowing	13
5.2.2 Autocorrelation (Discrete Periodic Signal)	4	6.4.1 Spectral estimation (Periodogram)	7	9.2.1 Welch's Method	14
5.2.3 Cross-Correlation (Discrete Periodic Signal)	4	6.4.2 Spectral estimation (via covariances)	8	9.2.2 Drifts and Offsets	14
		6.4.3 Spectral estimation (periodic signals)	8	10 Frequency Domain Subspace ID	14
		6.4.4 Spectral estimation (more general case)	8	10.1 Summarizing the subspace identification algorithm	15
		7 Averaging and Smoothing	8	10.1.1 Properties	16
		7.1 Bias-variance trade-offs in data record splitting	9	10.2 Nonuniformly spaced frequencies	16
		7.2 Smoothing the ETFE	9	11 Closed-Loop ID	17
		7.2.1 Assumptions on $\phi_v(e^{j\omega})$	9	12 Time-Domain Correlation Method	17
		7.2.2 Characteristic windows	10	13 Prediction Error Methods	17
		7.2.3 Asymptotic bias properties	10	14 Parameter estimation statistics	17
		7.2.4 Asymptotic variance properties	10	15 Nomenclature	17
		7.2.5 Asymptotic MSE properties	10	16 Overviews	18
				16.1 Transfer Functions	18
				16.2 Transfer Function Estimation	19

1 SYSTEM IDENTIFICATION

2 DEFINITIONS

Definition 1. A system is said to be **time invariant** if the response to a certain input is not depending on absolute time.

Definition 2. A system is said to be **linear** if its output response to a linear combination of inputs is the same as the linear combination of the output responses of the individual inputs.

Definition 3. A system is said to be **causal** if the output at a certain time depends on the input up to that time only.

Definition 4. A process is said to be **stationary** if it does not depend on time.

3 FREQUENCY DOMAIN METHODS

3.1 SAMPLING OPERATION

$$y(k) = y(t)|_{t=kT, k=0,1,2,\dots} \quad \text{Sampling with period } T$$

3.2 FOURIER SERIES OF PERIODIC SIGNALS

$$X(e^{j\omega_m}) = \sum_{k=0}^{M-1} x(k)e^{-j\omega_m k}$$

$$\omega_m = \frac{2\pi m}{M} = \omega_0$$

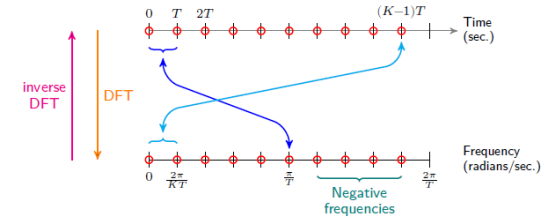
Non-negative frequencies are $m = 0$ to $m = M/s$.

They correspond to: $\omega_m = 0, \frac{2\pi}{M}, \frac{4\pi}{M}, \dots, \frac{2\pi(M/2-1)}{M}, \pi$.

M	number of samples	
$\omega_0 = \frac{2\pi}{M}$	fundamental frequency ($y(k)$)	[rad]
T	sampling time	[s]
$\tau_p = MT$	period	[s]
$\omega_0 = \frac{2\pi}{\tau_p}$	fundamental frequency ($y(t)$)	[rad s ⁻¹]

$$0, \underbrace{\frac{2\pi}{\tau_p}}_{\text{Fundamental frequency}}, \underbrace{2\left(\frac{2\pi}{\tau_p}\right), \dots, \frac{M}{2}\left(\frac{2\pi}{\tau_p}\right)}_{\text{Harmonics}}$$

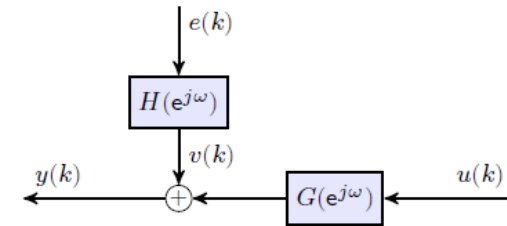
Definition 5. The highest frequency $\omega_u = \omega_{M/2} = \frac{\pi}{T}$ is called the **Nyquist frequency**.



```
1 %Non negative frequency vector
2 omega = omega_n(N, 'p0');
3 %Discrete Fourier Transform
4 fft(u); %Matlab
5 DFT(u); %Lecture
```

Note that the definition of the fft in the Matlab documentation does not describe the actual implementation perfectly. The index actually runs from 0 to $N - 1$ as in the definition made in class.

4 SPECTRAL ESTIMATION



$$Y(j\omega) = G(j\omega)U(j\omega) \quad \text{Transfer function}$$

$$Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) \quad \text{Discrete time TF}$$

$$\frac{\omega_u}{2\pi} = \frac{r}{NT}$$

ω_u	input frequency	[rad s ⁻¹]
N	calculation length	[]
T	experiment duration?	[s]
r	some integer	[]

$$u(k) = \alpha \cos(\omega_u k), \quad k = 0, 1, \dots, K-1 \quad \text{with } K \geq N \quad \text{Input}$$

$$y(k) = \alpha |G(e^{j\omega_u})| \cos(\omega_u k + \theta(\omega_u)) + v(k) + \text{transient} \quad \text{Output}$$

where $\theta(\omega_u) = \arg(G(e^{j\omega_u}))$

4.1 SINUSOIDAL CORRELATION METHODS

Correlation functions:

$$I_c(N) = \frac{1}{N} \sum_{k=0}^{N-1} y(k) \cos(\omega_u k)$$

$$I_s(N) = \frac{1}{N} \sum_{k=0}^{N-1} y(k) \sin(\omega_u k)$$

To calculate those from the data:

$$I_c(N) = \frac{\alpha}{2} \left| G(e^{j\omega_u}) \right| \cos(\theta(\omega_u)) + \frac{\alpha}{2} \left| G(e^{j\omega_u}) \right| \frac{1}{N} \sum_{k=0}^{N-1} \cos(2\omega_u k + \theta(\omega_u)) + \frac{1}{N} \sum_{k=0}^{N-1} v(k) \cos(\omega_u k)$$

If the noise, $v(k)$ is sufficiently uncorrelated then the variance satisfies,

$$\lim_{N \rightarrow \infty} \text{var} \left\{ \frac{1}{N} \sum_{k=0}^{N-1} v(k) \cos(\omega_u k) \right\} = 0$$

with a convergence rate of $1/N$.

Thus in the limit $N \rightarrow \infty$,

$$E \{ I_c(N) \} \rightarrow \frac{\alpha}{2} \left| G(e^{j\omega_u}) \right| \cos(\theta(\omega_u))$$

$$E \{ I_s(N) \} \rightarrow -\frac{\alpha}{2} \left| G(e^{j\omega_u}) \right| \sin(\theta(\omega_u))$$

and since $\lim_{N \rightarrow \infty} \text{var} \{ I_c(N) \} = 0$, $\lim_{N \rightarrow \infty} \text{var} \{ I_s(N) \} = 0$

The transfer function can be estimated via:

$$\hat{G}_N(e^{j\omega_u}) = \frac{I_c(N) - jI_s(N)}{\alpha/2}$$

- Advantages

- Energy is concentrated at the frequencies of interest.
- Amplitude of $u(k)$ can easily be tuned as a function of frequency.
- Easy to avoid saturation and tune signal/noise (S/N) ratio.

- Disadvantages

- A large amount of data is required.
- Significant amount of time required for experiments.
- Some processes won't allow sinusoidal inputs.

5 FREQUENCY DOMAIN METHODS

$x(k)$, $k = -\infty, \dots, \infty$ Discrete-time domain signal

$$X(e^{j\omega}) = \sum_{k=-\infty}^{\infty} x(k) e^{-j\omega k} \quad \text{Fourier Transform}$$

- $X(e^{j\omega})$ is 2π periodic.
- If $\sum_{k=-\infty}^{\infty} |x(k)| < \infty$ then $X(e^{j\omega})$ converges.

$$x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega k} d\omega \quad \text{Inverse Fourier Transform}$$

where $k = -\infty, \dots, \infty$

5.1 FINITE ENERGY SIGNAL

5.1.1 ENERGY SPECTRAL DENSITY (FINITE ENERGY SIGNAL)

If $x(k)$ is a finite energy signal,

$$\|x(k)\|_2^2 = \sum_{k=-\infty}^{\infty} |x(k)|^2 < \infty$$

$$S_x(e^{j\omega}) = |X(e^{j\omega})|^2 \quad \text{Energy Spectral Density}$$

- For finite energy signals the energy spectral density is easily calculated, for signals of infinite length and thus infinite energy however, this is not possible. For that reason the **power** spectral density is calculated instead!

5.1.2 AUTOCORRELATION (FINITE ENERGY SIGNAL)

$$R_x(\tau) = \sum_{k=-\infty}^{\infty} x(k) x(k - \tau), \quad \tau = -\infty, \dots, 0, \dots, \infty$$

The spectral density is the Fourier Transform of the autocorrelation:

$$\sum_{\tau=-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} = S_x(e^{j\omega})$$

```
1 % autocorrelation for a non-periodic, finite energy signal
2 Correlation('finen',u); %Lecture
3 xcorr(u); %Matlab
4
5 % crosscorrelation for non-periodic, finite energy signals
6 Correlation('finen',u,y); %Lecture
7 xcorr(y,u); %Matlab
```

5.2 DISCRETE PERIODIC SIGNAL

$$x(k) = x(k + M), \quad \forall k \in \{-\infty, \infty\} \quad \text{Periodic signal}$$

$$\omega_0 = \frac{2\pi}{M} \quad \text{Fundamental frequency}$$

- There are only M unique harmonics of the sinusoid $e^{j\omega_0}$.

- The non-negative harmonic frequencies are,

$$e^{jn\omega_0}, \quad n = 0, 1, \dots, M/2$$

```
1 u_period = sqrt(variance)*randn(N,1)+bias;
2 u = repmat(u_period,periods,1);
```

5.2.1 DISCRETE FOURIER SERIES (DISCRETE PERIODIC SIGNAL)

$$X(e^{j\omega_n}) = \sum_{k=0}^{N-1} x(k)e^{-j\omega_n k}, \quad \text{where } \omega_n = \frac{2\pi n}{N} = n\omega_0$$

$$x(k) = \frac{1}{N} \sum_{k=0}^{N-1} X(e^{j\omega_n})e^{j\omega_n k} \quad \text{Inverse Transform}$$

5.2.2 AUTOCORRELATION (DISCRETE PERIODIC SIGNAL)

$$R_x(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} x(k)x(k-\tau)$$

The Fourier transform of $R_x(\tau)$ is now defined as the **power spectral density**, since it is normalized with the signal length.

$$\phi_x(e^{j\omega_n}) = \sum_{\tau=0}^{N-1} R_x(\tau)e^{-j\omega_n \tau} = \frac{1}{N} |X(e^{j\omega_n})|^2$$

The energy in a single period is:

$$\sum_{k=0}^{N-1} |x(k)|^2 = \sum_{n=0}^{N-1} \phi_x(e^{j\omega_n})$$

```
1 Correlation('periodic',u);
```

5.2.3 CROSS-CORRELATION (DISCRETE PERIODIC SIGNAL)

$$R_{yu}(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} y(k)u(k-\tau)$$

The Fourier transform of $R_{yu}(\tau)$ is now defined as the **cross-spectral density**.

$$\phi_{yu}(e^{j\omega_n}) = \sum_{\tau=0}^{N-1} R_{yu}(\tau)e^{-j\omega_n \tau} = \frac{1}{N} Y(e^{j\omega_n})U^*(e^{j\omega_n})$$

```
1 Correlation('periodic',u,y);
```

5.3 RANDOM SIGNAL

Normally distributed noise:

$$e(k) \in \mathcal{N}(0, \lambda) \Rightarrow \begin{cases} \mathbb{E}[e(k)] = 0 \text{ (zero mean)} \\ \mathbb{E}[|e(k)|^2] = \lambda \text{ (variance)} \end{cases}$$

The $e(k)$ are independent and identically distributed (i.i.d.).

```
1 standard_deviation = 2;
2 variance = standard_deviation^2;
3 bias = 0;
4 N = 1024;
5 u = randn(N,1)*standard_deviation + bias;
```

5.3.1 AUTOCOVARANCE (RANDOM SIGNAL)

$$\begin{aligned} R_x(\tau) &= \mathbb{E}[x(k)x(k-\tau)] \\ &= \mathbb{E}[x(k)x^*(k-\tau)] \quad (\text{in the complex case}) \\ &= \mathbb{E}[x(k)x^*(x-\tau)] \quad (\text{in the multivariable case}) \end{aligned}$$

General (non-stationary, non-zero mean) case:

$$\begin{aligned} R_x(s, t) &= \mathbb{E}[(x(s) - \mathbb{E}[x])(x(t) - \mathbb{E}[E])] \\ &= \mathbb{E}[x(s)x(t)] \quad (\text{if zero mean}) \\ &= R_x(s - t) \quad (\text{if stationary}) \end{aligned}$$

Further properties are

- $R_x(-\tau) = R_x^*(\tau)$
- $R_x(0) \geq |R_x(\tau)| \forall \tau > 0$

```
1 xcorr(u)/N; %Lecture 3.37
```

5.3.2 POWER SPECTRAL DENSITY (RANDOM SIGNAL)

$$\phi_x(e^{j\omega}) := \sum_{\tau=-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} \text{ where } \omega \in [-\pi, \pi)$$

For a zero-mean random signal:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} |x(k)|^2 = \text{Var}(x(k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_x(e^{j\omega}) d\omega$$

Further properties are

- $\phi_x(e^{j\omega}) \in \mathbb{R}$
- $\phi_x(e^{j\omega}) \geq 0 \forall \omega$
- $\phi_x(e^{j\omega}) = \phi_x(e^{-j\omega})$ for all real-valued $x(k)$

```
1 fft(xcorr(u)/N) %Lecture 3.37 (?)
```

5.3.3 CROSS-COVARIANCE (RANDOM SIGNAL)

$$R_{yu}(\tau) = \mathbb{E}[(y(k) - \mathbb{E}[y(k)])(u(k - \tau) - \mathbb{E}[u(k)])]$$

For zero mean signals:

$$R_{yu}(\tau) = \mathbb{E}[y(k)u(k - \tau)]$$

Joint stationarity is required to make the definition dependent on τ only.

If $R_{yu}(\tau) = 0$ for all τ then $y(k)$ and $u(k)$ are uncorrelated.

```
1 xcorr(u,y)/N; %Lecture 3.37 (?)
```

5.3.4 CROSS POWER SPECTRAL DENSITY (RANDOM SIGNAL)

$$\phi_{yu}(e^{j\omega}) = \sum_{\tau=-\infty}^{\infty} R_{yu}(\tau) e^{-j\omega\tau}, \omega \in [-\pi, \pi)$$

The inverse is,

$$R_{yu}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{yu}(e^{j\omega}) e^{j\omega\tau} d\omega$$

```
1 fft(xcorr(y,u)/N); %Lecture 3.37 (?)
```

5.4 FINITE LENGTH SIGNAL

5.4.1 DISCRETE-FOURIER TRANSFORM (FINITE LENGTH SIGNAL)

$$X_N(e^{j\omega_n}) = \sum_{k=0}^{N-1} x(k) e^{-j\omega_n k}, \text{ where } \omega_n = \frac{2\pi n}{N}$$

The inverse DFT is

$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} X_N(e^{j\omega_n}) e^{j\omega_n k}, \quad k = 0, \dots, N-1$$

5.4.2 PERIODOGRAM (FINITE LENGTH SIGNAL)

$$\left[\frac{1}{N} |V_N(e^{j\omega})|^2 \right]$$

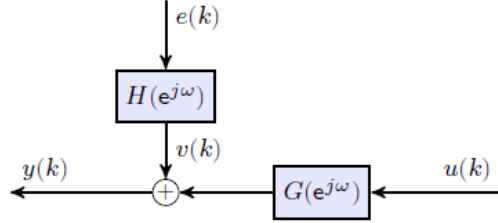
An asymptotically unbiased estimator of the spectrum is

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\frac{1}{N} |V_N(e^{j\omega})|^2 \right] = \phi_v(\omega)$$

This assumes that the autocorrelation decays quickly enough:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\tau=-N}^N |\tau R_v(\tau)| = 0$$

6 ETFE



Linear, time-invariant system, $g(l)$:

$$y(k) = \sum_{l=0}^{\infty} g(l)u(k-l) + v(k), \quad k = 0, 1, \dots$$

Assumptions:

1. causal system: $g(l) = 0, \forall l < 0$
2. noise: $E\{v(k)\} = 0$, zero mean, stationary

Given $\{u(k), y(k)\}$ find an estimate $\hat{G}(e^{j\omega})$ such that it fits the $G(e^{j\omega})$.

$$\boxed{\text{Bias}(\hat{G}))G - E\{\hat{G}\}} \quad \text{Bias}$$

$$\boxed{\text{var}((\hat{G})) = E\left\{|\hat{G} - E\{\hat{G}\}|^2\right\}} \quad \text{Variance}$$

$$\boxed{\text{MSE}(\hat{G}) = E\left\{|G - \hat{G}|^2\right\}} \quad \text{Mean-square error}$$

Note that $\text{MSE}(\hat{G}) = \text{var}(\hat{G}) + \text{Bias}^2(\hat{G})$.

6.1 INPUT-OUTPUT RELATIONSHIP

For finite energy signals:

$$y(k) = \sum_{l=0}^{\infty} g(l)u(k-l) + v(k)$$

$$Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) + V(e^{j\omega})$$

which in the idealized case leads to:

$$\boxed{\frac{Y(e^{j\omega})}{U(e^{j\omega})} = G(e^{j\omega}) + \frac{V(e^{j\omega})}{U(e^{j\omega})} \approx G(e^{j\omega})}$$

In reality we only have N samples:

$$\underbrace{Y_N(e^{j\omega_n})}_{\text{length-}N \text{ DFT}} = \sum_{k=0}^{N-1} y(k)e^{-j\omega_n k} \approx \sum_{k=-\infty}^{\infty} y(k)e^{-j\omega_n k} = Y(e^{j\omega_n})$$

$$\underbrace{U_N(e^{j\omega_n})}_{\text{length-}N \text{ DFT}} = \sum_{k=0}^{N-1} u(k)e^{-j\omega_n k} \approx \sum_{k=-\infty}^{\infty} u(k)e^{-j\omega_n k} = U(e^{j\omega_n})$$

$$\boxed{\hat{G}_N(e^{j\omega_n}) := \frac{Y_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}} \quad \text{ETFE}$$

6.2 PERIODIC INPUT CASE

Period M inputs: $u(k) = u(k+M)$

If $sM = N$ for an integer s , the fourier series over N samples is equal to the real fourier series!

$$U_N(e^{j\omega_n}) = U(e^{j\omega_n}) \forall \omega_n = \frac{2\pi n}{N}, \quad n = 0, \dots, N-1$$

Then

$$Y_N(e^{j\omega_n}) = G(e^{j\omega_n})U_N(e^{j\omega_n}) + V_N(e^{j\omega_n})$$

$$\hat{G}_N(e^{j\omega_n}) = G(e^{j\omega_n}) + \frac{V_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}$$

Bias:

$$E\{\hat{G}_N(e^{j\omega_n})\} = G(e^{j\omega_n}) + E\left\{\frac{V_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}\right\} = G(e^{j\omega_n})$$

when assuming zero mean noise. Thus for periodic inputs with N being an integer number of periods, **the ETFE is unbiased**.

Variance:

For the unbiased case:

$$E\left\{|\hat{G}_N(e^{j\omega_n}) - G(e^{j\omega_n})|^2\right\} = \frac{\phi_v(e^{j\omega_n}) + \frac{2}{N}c}{\frac{1}{N}|U_N(e^{j\omega_n})|^2}$$

where $|c| \leq C = \sum_{\tau=1}^{\infty} |\tau R_v(\tau)|$ is assumed to be finite.

For estimates at different frequencies ($\omega_n \neq \omega_i$):

$$E\left\{(\hat{G}_N(e^{j\omega_n}) - G(e^{j\omega_n}))(\hat{G}_N(e^{-j\omega_i}) - G(e^{-j\omega_i}))\right\} = 0$$

Transient responses:

Initial transient corrupts the measurement

$$y(k) = G(u_{\text{periodic}}(k)W_{[0, N-1]}(k)) + v(k)$$

with the window function:

$$W_{[0, N-1]}(k) = \begin{cases} 1 & \text{if } 0 \leq k < N \\ 0 & \text{otherwise} \end{cases}$$

For all outputs up to time $k = N - 1$

$$y(k) = G_{\text{periodic}}(k) - \underbrace{G(u_{\text{periodic}} W_{(-\infty, -1)})}_{r(k)} + v(k)$$

$$Y_N(e^{j\omega_n}) = G(e^{j\omega_n})U_N(e^{j\omega_n}) + R_N(e^{j\omega_n}) + V_N(e^{j\omega_n})$$

The input in negative time, which is present in a ideal periodic input, and missing in a real periodic input, has an influence on positive time, which is described by $r(k)$.

When using a periodic signal multiple times the resulting DFT does not contain more information, since in a periodic signal there are only a certain number of frequencies contained, but the energy in those frequencies increases!

Transient bias error:

$$\hat{G}(e^{j\omega_n}) = \frac{Y_N(e^{j\omega_n})}{U_N(e^{j\omega_n})} = G(e^{j\omega_n}) + \frac{R_N(e^{j\omega_n})}{U_N(e^{j\omega_n})} + \frac{V_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}$$

For periodic $u(k)$

As $N = mM, m \rightarrow \infty$

$$|U_N(e^{j\omega_n})| = m|U_M(e^{j\omega_n})|$$

For random $u(k)$

As $N \rightarrow \infty$

$$E\{|U_N(e^{j\omega_n})|\} \rightarrow \sqrt{N} \sqrt{\phi_u(e^{j\omega_n})}$$

Thus

$$\left| \frac{R_N(e^{j\omega_n})}{U_N(e^{j\omega_n})} \right| \rightarrow 0 \text{ with rate } \begin{cases} \frac{1}{N} & \text{for periodic input} \\ \frac{1}{\sqrt{N}} & \text{for random inputs} \end{cases}$$

A fix for getting rid of the influence of the transient response: Get rid of the first period.

```

1 r = 4;
2 N_split = N/r;
3 u_re = reshape(u(N_split+1:end), N_split, []);
4 y_re = reshape(y(N_split+1:end), N_split, []);
5
6 U_re = fft(u_re);
7 Y_re = fft(y_re);
8
9 G_re = Y_re ./ U_re;
10
11 G_avg = mean(G_re, 2);

```

6.3 SPECTRAL TRANSFORMATIONS

If $v(k) = 0$

$$\phi_y(e^{j\omega_n}) = G(e^{j\omega_n})\phi_u(e^{j\omega_n})G^T(e^{j\omega_n})$$

where $G^T(e^{j\omega_n})$ is the complex conjugate of $G(e^{j\omega_n})$.

If $v(k) \neq 0$ and uncorrelated

$$\phi_y(e^{j\omega_n}) = |G(e^{j\omega_n})|^2 \phi_u(e^{j\omega_n}) + |H(e^{j\omega_n})|^2$$

But this approach has no more phase information. For that reason use the cross spectrum:

$$\phi_{yu}(e^{j\omega_n}) = G(e^{j\omega_n})\phi_u(e^{j\omega_n}) + \phi_{uv}(e^{j\omega_n}) = G(e^{j\omega_n})\phi_u(e^{j\omega_n})$$

if $u(k)$ and $v(k)$ are uncorrelated.

$$\boxed{\hat{G}(e^{j\omega}) = \frac{\hat{\phi}_{yu}(e^{j\omega_n})}{\hat{\phi}_u(e^{j\omega_n})}} \text{ Spectral estimation methods}$$

where

$$\phi_y(e^{j\omega_n}) = |G(e^{j\omega_n})|^2 \phi_u(e^{j\omega_n}) + \phi_v(e^{j\omega_n})$$

$$\phi_{yu}(e^{j\omega_n}) = G(e^{j\omega_n})\phi_u(e^{j\omega_n})$$

6.4 APPROACHES TO SPECTRAL ESTIMATION

6.4.1 SPECTRAL ESTIMATION (PERIODOGRAM)

The periodogram is an asymptotically unbiased estimator of the spectrum given

$$\lim_{n \rightarrow \infty} \frac{1}{N} \sum_{\tau=-N}^N |\tau R_v(\tau)| = 0$$

$$\frac{1}{N} |V_N(e^{j\omega_n})|^2 \text{ Periodogram}$$

$$\lim_{N \rightarrow \infty} E \left\{ \frac{1}{N} |V_N(e^{j\omega_n})|^2 \right\} = \phi_v(e^{j\omega_n})$$

which is under the assumption

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\tau=-N}^N |\tau R_v(\tau)| = 0$$

6.4.2 SPECTRAL ESTIMATION (VIA COVARIANCES)

The autocovariance of the noise for stochastic $v(k)$ is described as:

$$\hat{R}_v(\tau) = \begin{cases} \frac{1}{N-|\tau|} \sum_{k=\tau}^{N_1} v(k)v(k-\tau), & \text{for } \tau \geq 0 \\ \frac{1}{N-|\tau|} \sum_{k=0}^{N+\tau-1} v(k)v(k-\tau), & \text{for } \tau < 0 \end{cases}$$

This is an unbiased estimator of $R_v(\tau)$: $E\{\hat{R}_v(\tau)\} = R_v(\tau)$

$$\hat{\phi}_v(e^{j\omega_n}) = \sum_{\tau=-N+1}^{N-1} \hat{R}_v(\tau) e^{-j\omega\tau}$$

```
1 xcorr(a-mean(a))-xcov(a) == 0 %true
```

The functions above both calculate the zero-mean shifted autocorrelation of a signal which is **not** equivalent to the autocovariance of the signal! They lack the normalization by $N - |\tau|$!

```
1 Covariance('zero-mean',a);
2 Covariance('zero-mean',a,b);
```

6.4.3 SPECTRAL ESTIMATION (PERIODIC SIGNALS)

Periodic signal $x(k)$ with period M , $N = mM$ for some integer m

$$R_x(\tau) = \frac{1}{M} \sum_{k=0}^{M-1} x(k)x(k-\tau)$$

The power spectral density can be calculated and is equal to the periodogram:

$$\phi_x(e^{j\omega_n}) = \sum_{\tau=0}^{M-1} R_x(\tau) e^{-j\omega_n\tau} = \frac{1}{M} |X_M(e^{j\omega_n})|^2$$

```
1 Ru = Correlation('periodic',u);
```

6.4.4 SPECTRAL ESTIMATION (MORE GENERAL CASE)

Alternative autocorrelation estimate:

$$\hat{R}_x(\tau) = \begin{cases} \frac{1}{N} \sum_{k=\tau}^{N-1} x(k)x(k-\tau), & \text{for } \tau \geq 0 \\ \frac{1}{N} \sum_{k=0}^{N+\tau-1} x(k)x(k-\tau), & \text{for } \tau < 0 \end{cases}$$

Periodic $x(k)$: unbiased (exact) if $N = mM$

Random $x(k)$ biased $E\{\hat{R}_x(\tau)\} = \frac{N-|\tau|}{N} R_x(\tau)$.

asymptotically biased as $N \rightarrow \infty, \tau/N \rightarrow 0$

```
1 Ru = xcorr(u)/N;
```

- The bias when estimating R_x for a random signal has the form a triangular weighting across τ . A scaling of the signal would remove the bias, which is effectively done by the spectral estimation for via covariances 6.4.2.

7 AVERAGING AND SMOOTHING

Multiple experiments $u_r(k), y_r(k), r = 1 \dots, R, k = 0, \dots, K-1$

$$\hat{G}(e^{j\omega_n}) = \sum_{r=1}^R \alpha_r \hat{G}_r(e^{j\omega_n})$$

where $\sum_{r=1}^G \alpha_r = 1$ and for calculating the average $\alpha_r = \frac{1}{R}$.

The averaging can be optimized by selecting α_r such that the variance $\sigma_r^2(e^{j\omega_n})$ is minimized.

$$\text{Var}(\hat{G}(e^{j\omega_n})) = \text{Var}\left(\sum_{r=1}^R \alpha_r(e^{j\omega_n}) \hat{G}_r(e^{j\omega_n})\right) = \sum_{r=1}^R \alpha_r^2 \sigma_r^2(e^{j\omega_n})$$

This is minimized by

$$\alpha_r(e^{j\omega_n}) = \frac{1/\sigma_r^2(e^{j\omega_n})}{\sum_{r=1}^T 1/\sigma_r^2(e^{j\omega_n})}$$

Thus the signal is weighted inversely proportional to the variance.

Thus if $\text{Var}(\hat{G}_r(e^{j\omega_n})) = \frac{\phi_v(e^{j\omega_n})}{\frac{1}{N}|U_r(e^{j\omega_n})|^2}$ then $\alpha_r(e^{j\omega_n}) = \frac{|U_r(e^{j\omega_n})|^2}{\sum_{r=1}^R |U_r(e^{j\omega_n})|^2}$.

The best result is obtained if the input is the same for all r , which will lead to a reduction of the variance as follows:

$$\text{Var}(\hat{G}(e^{j\omega_n})) = \frac{\text{Var}(\hat{G}_r(e^{j\omega_n}))}{R}$$

Biased estimates will reduce the improvement in variance.

- Since we are adding complex numbers the magnitude of the average is not equal to the average of the magnitudes r_i .

7.1 BIAS-VARIANCE TRADE-OFFS IN DATA RECORD SPLITTING

Divide a data record into smaller parts for averaging:

$$\{u(k), y(k)\}, k = 0, \dots, K-1$$

Choose R records and calculation length N , such that $NR \leq K$:

$$u_r(n) = u(rN + n)$$

And average the resulting estimates:

$$\hat{G}(e^{j\omega_n}) = \frac{1}{R} \sum_{r=0}^{R-1} \hat{G}_r(e^{j\omega_n}) = \frac{1}{R} \sum_{r=0}^{R-1} \frac{\hat{Y}_r(e^{j\omega_n})}{\hat{U}_r(e^{j\omega_n})}$$

As R increases:

- The number of points calculated, N decreases.
- The variance decreases (by up to $1/R$).
- The bias increases (due to non-periodicity transients).

Mean-square error

- Transient bias grows linearly with the number of data splits.
- Variance decays with a rate of up to $1/(\text{number of averages})$.

What if there is no option of running periodic input experiments? \rightarrow exploit the assumed smoothness of the underlying system.

7.2 SMOOTHING THE ETFE

Assume the true system to be close to constant for a range of frequencies: $G(e^{j\omega_{n+r}}) \approx G(e^{j\omega_n})$ for $r = 0, \pm 1, \dots, \pm R$.

The minimum variance smoothed estimate is:

$$\tilde{G}_N(e^{j\omega_n}) = \frac{\sum_{r=-R}^R \alpha_r \hat{G}_N(e^{j\omega_{n+r}})}{\sum_{r=-R}^R \alpha_r}, \quad \alpha_r = \frac{\frac{1}{N}|U_N(e^{j\omega_{n+r}})|^2}{\phi_v(e^{j\omega_{n+r}})}$$

The summation above can then be approximated by an integral:

$$\approx \frac{\int_{\omega_{n-R}}^{\omega_{n+R}} \alpha(e^{j\zeta}) \hat{G}_N(e^{j\zeta}) d\zeta}{\int_{\omega_{n-R}}^{\omega_{n+R}} \alpha(e^{j\zeta}) d\zeta}, \quad \text{with } \alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v(e^{j\zeta})}$$

Which can be reformulated using a smoothing window:

$$\tilde{G}_N(e^{j\omega_n}) = \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \alpha(e^{j\zeta}) \hat{G}_N(e^{j\zeta}) d\zeta}{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \alpha(e^{j\zeta}) d\zeta} \quad \text{with } \alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v(e^{j\zeta})}$$

7.2.1 ASSUMPTIONS ON $\phi_v(e^{j\omega})$

Assume $\phi_v(e^{j\omega})$ is also a smooth function of frequency.

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \left| \frac{1}{\phi_v(e^{j\zeta})} - \frac{1}{\phi_v(e^{j\omega_n})} \right| d\zeta \approx 0$$

Then use,

$$\alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v(e^{j\omega_n})}$$

to get

$$\tilde{G}_N(e^{j\omega_n}) = \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{-j(\zeta-\omega_n)}) \frac{1}{N}|U_N(e^{j\zeta})|^2 \hat{G}_N(e^{j\zeta}) d\zeta}{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \frac{1}{N}|U_N(e^{j\zeta})|^2 d\zeta}$$

The wider the frequency window (decreasing γ)

- the more adjacent frequencies included in the smoothness estimate.
- the smoother the result.
- the lower the noise induced variance.
- the higher the bias.

7.2.2 CHARACTERISTIC WINDOWS

$$W_\gamma(e^{j\omega}) = \frac{1}{\gamma} \left(\frac{\sin \gamma\omega/2}{\sin \omega/2} \right)^2 \quad \text{Bartlett}$$

$$W_\gamma(e^{j\omega}) = \frac{1}{2}D_\gamma(\omega) + \frac{1}{4}D_\gamma(\omega - \pi/\gamma) + \frac{1}{4}D_\gamma(\omega + \pi/\gamma) \quad \text{Hann}$$

where

$$D_\gamma(\omega) = \frac{\sin \omega(\gamma+0.5)}{\sin \omega/2}$$

Properties of window functions:

- $\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j\zeta}) d\zeta = 1$
- $\int_{-\pi}^{\pi} \zeta W_\gamma(e^{j\zeta}) d\zeta = 0$
- $M(\gamma) := \int_{-\pi}^{\pi} \zeta^2 W_\gamma(e^{j\zeta}) d\zeta$
- $\bar{W}(\gamma) := 2\pi \int_{-\pi}^{\pi} W_\gamma^2(e^{j\zeta}) d\zeta$

$$\begin{array}{ll} \text{Bartlett} & M(\gamma) = \frac{2.78}{\gamma^2}, \quad \bar{W}(\gamma) \approx 0.67\gamma \quad (\text{for } \gamma > 5) \\ \text{Hamming} & M(\gamma) = \frac{\pi^2}{2\gamma^2}, \quad \bar{W}(\gamma) \approx 0.75\gamma \quad (\text{for } \gamma > 5) \end{array}$$

- $M(\gamma)$ gives an idea of the bias effect.
- $\bar{W}(\gamma)$ gives an idea of the variance effect.

7.2.3 ASYMPTOTIC BIAS PROPERTIES

$$\begin{aligned} \mathbb{E} \left\{ \tilde{G}(e^{j\omega_n}) - \mathbb{E} \left\{ G(e^{j\omega_n}) \right\} \right\} &= \mathbb{E} \left\{ \tilde{G}(e^{j\omega_n}) - G(e^{j\omega_n}) \right\} = \\ M(\gamma) \left(\underbrace{\frac{1}{2} G''(e^{j\omega_n})}_{\text{curvature}} + \underbrace{G'(e^{j\omega_n}) \frac{\phi'_u(e^{j\omega_n})}{\phi_u(e^{j\omega_n})}}_{\text{slope}} \right) &+ H.O.T. \end{aligned}$$

Increasing γ

- makes the frequency window smaller.
- averages over fewer frequency values.
- makes $M(\gamma)$ smaller
- reduces the bias of the smoothed estimate $\tilde{G}(e^{j\omega_n})$

7.2.4 ASYMPTOTIC VARIANCE PROPERTIES

$$\mathbb{E} \left\{ (\tilde{G}(e^{j\omega_n}) - \mathbb{E} \left\{ \tilde{G}(e^{j\omega_n}) \right\})^2 \right\} = \frac{1}{N} \bar{W}(\gamma) \frac{\phi_v(e^{j\omega_n})}{\phi_u(e^{j\omega_n})} + H.O.T.$$

Increasing γ

- makes the frequency window narrower.
- averages over fewer frequency values.
- makes \bar{W}_γ larger.
- increases the variance of the smoothed estimate $\tilde{G}(e^{j\omega_n})$.

7.2.5 ASYMPTOTIC MSE PROPERTIES

$$\mathbb{E} \left\{ |\tilde{G}(e^{j\omega_n}) - G(e^{j\omega_n})|^2 \right\} \approx M^2(\gamma) |F(e^{j\omega_n})|^2 + \frac{1}{N} \bar{W}(\gamma) \frac{\phi_v(e^{j\omega_n})}{\phi_u(e^{j\omega_n})}$$

where

$$F(e^{j\omega_n}) = \frac{1}{2} G''(e^{j\omega_n}) + G'(e^{j\omega_n}) \frac{\phi'_u(e^{j\omega_n})}{\phi_u(e^{j\omega_n})}$$

If $M(\gamma) = M/\gamma^2$ and $\bar{W}(\gamma) = \bar{W}\gamma$ then MSE is minimised by:

$$\gamma_{\text{optimal}} = \left(\frac{4M^2 |F(e^{j\omega_n})|^2 \phi_u(e^{j\omega_n})}{\bar{W} \phi_v(e^{j\omega_n})} \right)^{1/5} N^{1/5}$$

and

$$\text{MSE at } \gamma_{\text{optimal}} \approx CN^{-4/5}$$

8 WINDOWING AND INPUT SIGNALS

$$\phi_{yu}(e^{j\omega}) = G(e^{j\omega}) \phi_u(e^{j\omega})$$

$$\hat{G}(e^{j\omega_n}) = \frac{\hat{\phi}_{yu}(e^{j\omega_n})}{\hat{\phi}_u(e^{j\omega_n})}$$

Recall that the smoothed ETFE is:

$$\tilde{G}_N(e^{j\omega_n}) = \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\zeta})|^2 \hat{G}_N(e^{j\zeta}) d\zeta}{\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\zeta})|^2 d\zeta}$$

The denominator term approaches $\frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)}) \phi(e^{j\omega_n}) d\zeta$ as $N \rightarrow \infty$.

If in addition $W_\gamma(e^{j\omega})$ is concentrated around $\zeta = 0$ (i.e. $\gamma/N \rightarrow 0$) then the denominator term approaches $\phi_u(e^{j\omega_n})$ as $N \rightarrow \infty$.

This motivates the smoothed spectral estimate:

$$\tilde{\phi}_u(e^{j\omega_n}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\omega})|^2 d\zeta$$

Similarly the numerator approaches ϕ_{yu} as $N \rightarrow \infty$:

$$\tilde{\phi}_{yu}(e^{j\omega_n}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\omega})|^2 \hat{G}_N(e^{j\omega}) d\zeta$$

For this reason the smoothed ETFE is equal to the smoothed spectral estimate for $N \rightarrow \infty$.

8.1 FREQUENCY DOMAIN SMOOTHING IN MATLAB

```
1 gamma = 80;
2 U = fft(u); Y = fft(y);
3 G_est = Y./U;
4 G_est_smooth = G_est*0;
5
6 [om,Wg] = WfHann(g,N);
7 %shift to start at zero
8 zidx = find(om == 0);
9 omega = [om(zidx:N); om(1:zidx-1)];
10 Wg = [Wg(zidx:N) Wg(1:zidx-1)];
11 %variance weighting
12 a = U.*conj(U);
```

```
1 for wn = 1:N
2     %reset normalization
3     Wnorm = 0;
4     for xi = 1:N
5         %wrap window index
6         widx = mod(xi-wn,N)+1;
7         G_est_smooth(wn) = G_est_smooth(wn) +...
8             Wg(widx)*G_est(xi)*a(xi);
9         Wnorm = Wnorm + Wg(widx)*a(xi);
10    end
11    %weigh normalisation
12    G_est_smooth(wn) = G_est_smooth(wn)/Wnorm;
13 end
```

8.2 TIME DOMAIN WINDOWS

Define, via the inverse Fourier transform a time domain window:

$$\omega_{\gamma}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j\omega}) e^{j\zeta\tau} d\zeta$$

Then the smoothed input spectral estimate $\tilde{\phi}_u(e^{j\omega_n})$ is:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\omega})|^2 d\zeta \approx \sum_{\tau=-\infty}^{\infty} \omega_{\gamma}(\tau) \hat{R}_u(\tau) e^{-j\tau\omega_n}$$

where

$$\omega_{\gamma} = \begin{cases} 0 & \text{for } \tau < -\gamma \\ > 0 & \text{for } -\gamma \leq \tau \leq \gamma \\ 0 & \text{for } \tau > \gamma \end{cases}$$

where often $\gamma \ll N$, which enables the faster calculated redefinition:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j(\zeta-\omega_n)}) \frac{1}{N} |U_N(e^{j\omega})|^2 d\zeta \approx \sum_{\tau=-\gamma}^{\gamma} \omega_{\gamma}(\tau) \hat{R}_u(\tau) e^{-j\tau\omega_n}$$

The cross spectral estimate can also be formulated as a convolution in the frequency domain which leads to the analogous formulation to the spectral estimate of u :

$$\tilde{\phi}_u(e^{j\omega_n}) = \sum_{\tau=-\gamma}^{\gamma} \omega_{\gamma}(\tau) \hat{R}_u(\tau) e^{-j\tau\omega_n}$$

$$\tilde{\phi}_{yu}(e^{j\omega_n}) = \sum_{\tau=-\gamma}^{\gamma} \omega_{\gamma}(\tau) \hat{R}_{yu}(\tau) e^{-j\tau\omega_n}$$

8.2.1 TIME DOMAIN SMOOTHING IN MATLAB

```

1 gamma = 80;
2 [~,Wg] = WtHann(gamma,N);
3 R_u = xcorr(u,N/2)/N; %Lecture 3.37
4 R_yu = xcorr(y,u,N/2)/N;
5
6 omega = Omega_n(N);
7 phi_u = zeros(size(omega));
8 phi_yu = zeros(size(omega));
9
10 tau = -g:g;
11 ind = tau+N/2;
12 for i = 1:N
13 %Lecture 5.9 and 5.11
14     phi_u(i) = sum(Wg(tau+N/2).*R_u(ind).*exp(-1j.*tau.*
15     omega(i)));
16     phi_yu(i) = sum(Wg(tau+N/2).*R_yu(ind).*exp(-1j.*tau.*
17     omega(i)));
18 end
19 G_est_smooth = phi_yu./phi_u;

```

8.2.2 WINDOW CHARACTERISTICS

Decreasing γ : narrower $\omega_\gamma(\tau)$, wider $W_\gamma(e^{j\omega})$

- the more frequencies, $\hat{G}(e^{j\omega_n})$ included in the smoothing.
- the fewer $\hat{R}(\tau)$ estimates included in the smoothing.
- the smoother the result.
- the lower the noise induced variance.
- the **higher the bias**.

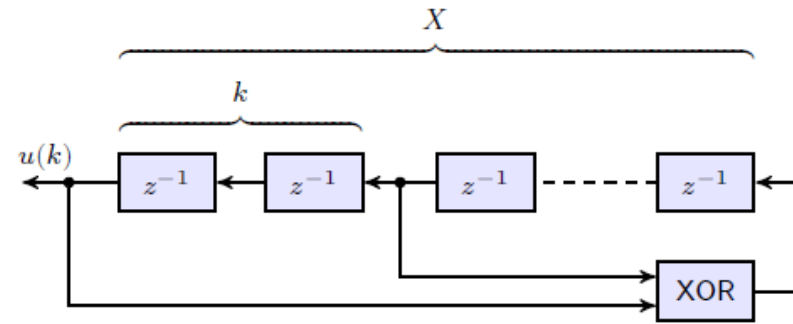
8.3 INPUT SIGNALS

- Steps
- Doublet
- Sinusiods, Chirpts, Multi-Sines
- Filtered white noise
- Pseudo-Random Binary Signals (PRBS)

8.3.1 PRBS

$$u(k) = a \text{ or } -a$$

Shift-register generation



X	k
3	1 or 2
4	1 or 3
5	2 or 3
6	1 or 5
7	1 or 6
8	—
9	4 or 5
10	3 or 7

Periodicity

Periodic with period equal to at most $M = 2^X - 1$.

$$R_u(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} u(k)u(k-\tau) = \begin{cases} a^2 & \text{if } \tau = 0 \\ \frac{-a^2}{2^X-1} & \text{if } \tau \neq 0 \end{cases}$$

Definition 6. *Run length* defines how long the signal stays high.

The run length distribution of $u(k)$ is then:

1/2 runs of length 1

1/4 runs of length 2

1/8 runs of length 3

⋮

Other properties:

- Equal energy at all frequencies.
- The maximum period of a PRBS signal can be found based on the discrete time dynamic system that generates the signal:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 2\gamma \end{bmatrix}$$

$$z = Ax(k)$$

$$x(k+1) = \text{mod}_2(z(k))$$

$$y = Cx(k) - \gamma$$

where γ is the amplitude of the signal.

Now the maximum length period result from the idea that $x(k)$ and therefore the output $y(k)$ can only change so many times until it reaches a state it has already held before. From that point on, since the system is first order, the signal repeats. Thus in the longest case, all possible values contained in x are taken exactly once. Thus the maximum periodic length is equal to the number of possible binary numbers with 6 bits, where 6 is the order of the PRBS signal.

$$\boxed{2^6 - 1} \text{ Maximum Length Period}$$

8.3.2 PRBS IN MATLAB

```
1 signal_order = 8;
2 singal_length = 2.^signal_order-1;
3 u = idinput(signal_length, 'PRBS');
```

- Note that `idinput` only allows the choice of the total length of the signal and derives the fitting signal order such that the signal is at least of that chosen length. A warning is made if the chosen signal length is not equivalent to the run length.

8.3.3 MULTI-SINUSOIDAL SIGNALS

$$u(k) = \sum_{s=1}^S \sqrt{2\alpha_s} \cos(\omega_s kT + \phi_s)$$

where T is the sampling period, $\omega_s = \frac{2\pi}{T_P}$, $\frac{T_P}{T} = N$, $S \leq \frac{N}{2}$.

Choose N to be a power of 2 for efficient FFT calculations.

$$\sum_{s=1}^S \alpha_s = 1 \quad \text{Total signal power}$$

Schroeder phasing

Select the phases ϕ_s such that the minimize the peak amplitude:

$$\phi_s = 2\pi \sum_{j=1}^s j\alpha_s.$$

for equal power in each sinusoids:

$$\alpha_s = 1/S \text{ and } \phi_s = \frac{\pi(s^2+s)}{S}$$

9 RESIDUAL SPECTRA, COHERENCY, APERIODICTY, OFFSETS AND DRIFTS

9.1 RESIDUAL SPECTRUM

9.1.1 ESTIMATING $\phi_v(e^{j\omega_n})$

$$v(k) = y(k) - G(e^{j\omega})u(k)$$

$$\tilde{\phi}_v(e^{j\omega_n}) \approx \frac{1}{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{\gamma}(e^{j(\zeta-\omega_n)}) \left| Y_N(e^{j\omega}) - \tilde{G}(e^{j\omega})U_N(e^{j\omega}) \right|^2 d\zeta \approx \tilde{\phi}_y(e^{j\omega_n}) - \frac{|\tilde{\phi}_{yu}(e^{j\omega_n})|^2}{\tilde{\phi}_u(e^{j\omega_n})}$$

How much energy is accounted for by the model? How much by noise?

$$\phi_v(e^{j\omega_n}) = \phi_y(e^{j\omega_n}) \left(1 - \frac{|\phi_{yu}(e^{j\omega_n})|^2}{\phi_y(e^{j\omega_n})\phi_u(e^{j\omega_n})} \right)$$

$$\hat{\kappa}_{yu}(e^{j\omega_n}) = \sqrt{\frac{|\hat{\phi}_{yu}(e^{j\omega_n})|^2}{\hat{\phi}_y(e^{j\omega_n})\hat{\phi}_u(e^{j\omega_n})}} \quad \text{Coherency Spectrum}$$

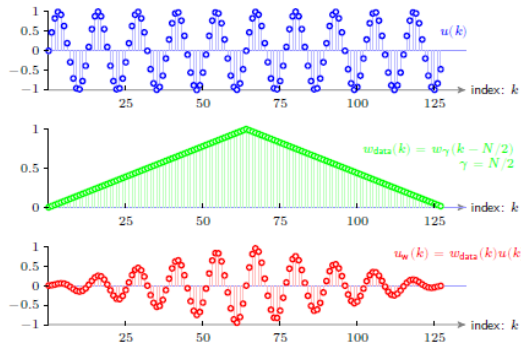
- If all of the energy in the output is due to the model for a frequency ω_n then $\hat{\kappa}_{yu}(e^{j\omega_n}) = 1$.
- This can be used as a measure of effectiveness of the modelling at a particular frequency.
- Theoretically, $0 \leq \hat{\kappa}_{yu}(e^{j\omega_n}) \leq 1$. One should aim to keep the coherency spectrum as high as possible. It can be adjusted by adjusting the smoothing.

9.2 TIME-DOMAIN DATA WINDOWING

Putting a time domain window directly on the data.

$$U_w(e^{j\omega_n}) = \sum_{k=0}^{N-1} w_{data}(k)u(k)e^{-jk\omega_n}$$

often with $w_{data}(k) = w_{\gamma}(k - N/2)$ (shifted to middle). Typically $\gamma = N/2$ such that all of the data is used.

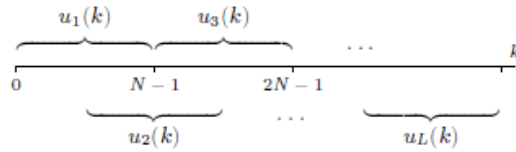


For an estimation of the periodogram **scaling is necessary!**

$$\frac{1}{E_{scl}} \frac{1}{N} |U_w(e^{j\omega_n})|^2 \quad \text{where } E_{scl} = \frac{\sum_{k=0}^{N-1} |w_{data}(k)u(k)|^2}{\sum_{k=0}^{N-1} |u(k)|^2} \approx \frac{1}{N} \sum_{k=0}^{N-1} |w_{data}(k)|^2 \quad \text{Periodogram}$$

- Transients influence the periodogram of a sequence, since by the application of the DFT, the signal is assumed to be periodic. Thus even if a sinusoid is sampled - if the periodic extension does not represent the same sinusoid - the DFT will show a range of frequencies, instead of a single distinct one.
- Time-domain data windowing can help reducing the influence of transients.

9.2.1 WELCH'S METHOD



1. Split the data record into L overlapping segments of length N .

$$2. U_l(e^{j\omega_n}) = \sum_{k=0}^{N-1} w_{data}(k)u_l(k)e^{j\omega_n k}$$

$$3. \tilde{\phi}_u(e^{j\omega_n}) = \frac{1}{NLE_{scl}} \sum_{l=1}^L |U_l(e^{j\omega_n})|^2$$

- + Windowing can reduce transient response effects.
- + Noise reduction from averaging and windowing.
- + Variance error can be reduced.

- Windowing can cause energy leakage to adjacent frequencies.
- Frequency resolution deteriorates.
- Bias error can be increased.
- Noise on $u_l(k)$ and u_{l+1} is not uncorrelated.
- Do not use `welch()`, since it does not fit the definition here.

9.2.2 DRIFTS AND OFFSETS

- Time domain windowing when an offset is present can lead to the introduction of additional frequencies close to zero, which is undesirable.
- Time domain windowing when a drift is present can lead to the introduction of additional frequencies close to the peak of the sampled sinusoid, which is undesirable as well.

- A possible solution is **preprocessing** the data via the assumption

$$u_d(k) = u(k) - (\alpha k + \beta)$$

where $\alpha k + \beta$ is the best linear fit to $u(k)$.

```
1 detrend(u);
```

+ **detrend** can completely remove the effect of drifts and offsets.

- It must be applied carefully, since if no drift/offset is present or if it is detected wrongly, large errors can be introduced.

10 FREQUENCY DOMAIN SUBSPACE ID

$$x(k+1) = Ax(k) + Bu(k) \quad A \in \mathbb{R}^{n_x \times n_x}$$

$$y(k) = Cx(k) + Du(k) \quad D \in \mathbb{R}^{n_y \times n_u}$$

$$g(k) = \begin{cases} 0 & k = 0 \\ D & k = 0 \\ CA^{k-1} & k > 0 \end{cases} \quad \text{Pulse response coefficients}$$

$$G(e^{j\omega_n}) = \sum_{k=0}^{\infty} g(k)e^{-j\omega_n k} \quad 0 \leq \omega \leq \pi$$

Given

$$G(n) = G(e^{j\omega_n}) + V(e^{j\omega_n}), \quad n = 0, \dots, N/2$$

Find

$$\hat{G}(e^{j\omega_n}) = \hat{C} \left(e^{j\omega} I - \hat{A} \right)^{-1} \hat{B} + \hat{D}$$

such that

$$\lim_{N \rightarrow \infty} \|\hat{G}(e^{j\omega_n}) - G(e^{j\omega_n})\|_{\infty} = 0$$

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} \in \mathbb{R}^{n_y q \times n_x} \quad \text{Extended observability}$$

$\text{rank}(\mathcal{O}) = n_x$ for all $n_y q \geq n_x$

$$\mathcal{C} = \begin{bmatrix} B & AB & \cdots & A^{r-1}B \end{bmatrix} \in \mathbb{R}^{n_x \times n_u r} \quad \text{Extended controllability}$$

$\text{rank}(\mathcal{C}) = n_x$ for all $n_u r \geq n_x$

The ETFE matches the true transfer function only if there is no noise. In that case the inverse Fourier transform results in the time-aliased impulse response of the system h_k :

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} G(e^{j\omega_n}) e^{j2\pi kn/N} = CA^{k-1} \left(\sum_{l=0}^{\infty} A^{Nl} \right) B = CA^{k-1} (I - A^N)^{-1} B$$

- We do not get the exact impulse response $g(k)$ since we do not have all the data until $N \rightarrow \infty$.

Now the **Hankel matrix** is formulated with the ultimate goal to extract estimates of the state space matrices.

$$H = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots & h_r \\ h_2 & h_3 & \ddots & \ddots & h_{r+1} \\ h_3 & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & & \vdots \\ h_q & h_{q+1} & & \cdots & h_{q+r-1} \end{bmatrix} \quad \text{choose } q > n_x, r > n_x \text{ and } q + r - 1 \leq N - 1$$

which leads to:

$$H = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} (I - A^N)^{-1} \begin{bmatrix} B & AB & \cdots & A^{r-1}B \end{bmatrix}$$

Next singular value decomposition can be used to decompose the Hankel matrix, which allows to calculate the **rank** of the state space, as the dimension of Σ_1 .

$$H = U \Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \quad \Sigma_1 \in \mathbb{R}^{n_x \times n_x}$$

Since U_2 is multiplied with $\mathbf{0}$ the $\text{range}(H) = \text{range}(U_1) = \mathcal{O}$. Which directly leads to $\text{span}(\{U_1\}) = \text{span}(\mathcal{O})$, which means that the observability matrix is some linear transform of U_1 . Or differently said, for some choice of A they are equal. **probably not correct!?**

Next an estimate of \hat{A} is made using the observability matrix, for which the selection matrix J is introduced:

$$J_1 \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-2} \end{bmatrix} \quad \text{and} \quad J_2 \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} = \begin{bmatrix} CA \\ \vdots \\ CA^{q-1} \end{bmatrix}$$

Then $J_1 \mathcal{O} A = J_2 \mathcal{O}$

So \hat{A} is the least squares solution to $J_1 \hat{U}_1 \hat{A} = J_2 \hat{U}_1$.

And \hat{C} can be found as: $J_e \mathcal{O} = C \implies J_3 \hat{U}_1 = \hat{C}$.

Next, using the initial formulation:

$$\hat{G}(e^{j\omega_n}) = \hat{C} \left(e^{j\omega} I - \hat{A} \right)^{-1} \hat{B} + \hat{D}$$

\hat{B} and \hat{D} can be found as the solution to the minimization:

$$\hat{B}, \hat{D} = \underset{B, D}{\text{argmin}} \sum_{n=0}^N \|G(e^{j\omega_n}) - D - \hat{C}(e^{j\omega_n} I - \hat{A})^{-1} B\|_F^2$$

10.1 SUMMARIZING THE SUBSPACE IDENTIFICATION ALGORITHM

Uniform data spacing case:

$$G(n), \quad \omega_n = \frac{\pi n}{N}, \quad n = 0, \dots, N/2$$

1. Extend data to negative frequencies

$$G(n) = \bar{G}(N - n), \quad n = N/2 + 1, \dots, N - 1$$

2. Calculate inverse DFT to obtain the time-aliased pulse response

$$\hat{h}_k = \frac{1}{N} \sum_{n=0}^{N-1} G(n) e^{j2\pi kn/N}, \quad k = 0, \dots, N - 1$$

3. Form a block-Hankel matrix:

$$\hat{H} = \begin{bmatrix} \hat{h}_1 & \hat{h}_2 & \hat{h}_3 & \cdots & \hat{h}_r \\ \hat{h}_2 & \hat{h}_3 & \ddots & \ddots & \hat{h}_{r+1} \\ \hat{h}_3 & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & & \vdots \\ \hat{h}_q & \hat{h}_{q+1} & & \cdots & \hat{h}_{q+r+1} \end{bmatrix} \in \mathbb{R}^{n_y q \times \nu_r}$$

4. Calculate a singular value decomposition:

$$\hat{H} = \hat{U} \hat{\Sigma} \hat{V}^T$$

5. Select a model order, \hat{n}_x and partition the SVD:

$$\hat{U} \hat{\Sigma} \hat{V}^T = \begin{bmatrix} \hat{U}_1 & \hat{U}_2 \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_1 & 0 \\ 0 & \hat{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \hat{V}_1^T \\ \hat{V}_2^T \end{bmatrix}, \quad \hat{\Sigma} \in \mathbb{R}^{\hat{n}_x \times \hat{n}_x}$$

6. Estimate \hat{A} via:

$$\begin{aligned} J_1 &= \begin{bmatrix} I_{n_y(q-1)} & O_{n_y(q-1) \times n_y} \end{bmatrix} \\ J_2 &= \begin{bmatrix} 0_{n_y(q-1) \times n_y} & I_{n_y(q-1)} \end{bmatrix} \end{aligned}$$

Solve for \hat{A} via LS:

$$\boxed{J_1 \hat{U}_1 \hat{A} = J_2 \hat{U}_1}$$

7. Estimate \hat{C} via

$$\begin{aligned} J_3 &= \begin{bmatrix} I_{n_y} & 0_{n_y \times n_y(q_1)} \end{bmatrix} \\ \boxed{\hat{C} &= J_3 \hat{U}_1} \end{aligned}$$

8. Find \hat{B} and \hat{D} via least squares:

$$\boxed{\hat{B}, \hat{D} = \underset{B, D}{\operatorname{argmin}} \sum_{n=0}^N \|G(e^{j\omega_n}) - D - \hat{C}(e^{j\omega_n} I - \hat{A})^{-1} \hat{B}\|_F^2}$$

9. Form the estimate

$$\boxed{\hat{G}(z) = \hat{D} + \hat{C}(zI - \hat{A})^{-1} \hat{B}}$$

- To get real valued B and D :

$$\begin{bmatrix} \operatorname{real} \left((\hat{C}((e^{j\omega_n})I - \hat{A})^{-1}) \right) & I \\ \operatorname{imag} \left((\hat{C}((e^{j\omega_n})I - \hat{A})^{-1}) \right) & 0 \end{bmatrix} \begin{bmatrix} B \\ D \end{bmatrix} = \begin{bmatrix} \operatorname{real}(G(n)) \\ \operatorname{imag}(G(n)) \end{bmatrix}$$

10.1.1 PROPERTIES

- Asymptotic convergence

$$\lim_{N \rightarrow \infty} \|\hat{G}(e^{j\omega_n}) - G(e^{j\omega_n})\|_\infty = 0, \quad n = 0, \dots, N-1 \text{ w.p. } 1$$

- The algorithm is „correct“. If $V(e^{j\omega_n}) = 0$ then there exists a data length $N_0 < \infty$ such that:

$$\|\hat{G}_N(e^{j\omega}) - G(e^{j\omega})\|_\infty = 0 \text{ for all } N > N_0$$

+ Time- and frequency-domain versions available. *N4SID*, etc

+ Many variants which depend on weighting for noise.

+ gives a state-space model directly.

+ Can be effective in determining system order.

+ Works equally well for MIMO systems.

- Unusual noise weighting in frequency-domain case.

- Truncated SVD reconstructions are not Hankel.

- \hat{U}_1 does not have the „shift“structure.

- Least-squares noise assumptions are not correct.

- Can give unstable models for stable systems.

10.2 NONUNIFORMLY SPACED FREQUENCIES

Time domain

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

Frequency domain

$$\begin{aligned} e^{j\omega} X(\omega) &= AX(\omega) + BU(\omega) \\ Y(\omega) &= CX(\omega) + BU(\omega) \end{aligned}$$

For a specific frequency ω on each channel

$$U_i(\omega) = e_i \quad i = 1, \dots, n_u$$

Resulting system equations

$$\begin{aligned} e^{j\omega} X_i(\omega) &= AX_i(\omega) + BU_i(\omega) \\ Y_i(\omega) &= CX_i(\omega) + DU_i(\omega) \end{aligned}$$

Defining

$$X_c(\omega) = [X_1(\omega) \quad \cdots \quad X_{n_u}(\omega)]$$

and stacking the equations column-wise gives:

$$\begin{aligned} e^{j\omega} X_c(\omega) &= A X_c(\omega) + B \\ G(\omega) &= C X_c(\omega) + D \end{aligned}$$

Multiplying by $e^{j\omega}$ and substituting, repeating and stacking row-wise:

$$\begin{bmatrix} G(e^{j\omega}) \\ e^{j\omega} G(e^{j\omega}) \\ \vdots \\ e^{j(q-1)\omega} G(e^{j\omega}) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix} X_c(\omega) + \Gamma \begin{bmatrix} I_{n_u} \\ e^{j\omega} I_{n_u} \\ \vdots \\ e^{j(q-1)\omega} I_{n_u} \end{bmatrix}$$

where

$$\Gamma = \begin{bmatrix} D & & & \\ CB & D & 0 & 0 \\ CAB & \ddots & \ddots & \\ \vdots & & \ddots & \\ CA^{q-2}B & \cdots & CB & D \end{bmatrix}$$

Repeat for all frequencies ω_i :

$$\begin{aligned} \mathcal{G} &= \frac{1}{\sqrt{N}} \begin{bmatrix} G(e^{j\omega_1}) & \cdots & G(e^{j\omega_N}) \\ e^{j\omega_1} G(e^{j\omega_1}) & & e^{j\omega_N} G(e^{j\omega_N}) \\ \vdots & & \vdots \\ e^{j(q-1)\omega_1} G(e^{j\omega_1}) & \cdots & e^{j(q-1)\omega_N} G(e^{j\omega_N}) \end{bmatrix} \\ \mathcal{W} &= \frac{1}{\sqrt{N}} \begin{bmatrix} I & \cdots & I \\ e^{j\omega_1} & & e^{j\omega_N} I \\ \vdots & & \vdots \\ e^{j(q-1)\omega_1} & \cdots & e^{j(q-1)\omega_N} I \end{bmatrix} \\ \mathcal{X}_c &= \frac{1}{\sqrt{N}} [X_c(\omega_1) \cdots X_c(\omega_N)] \end{aligned}$$

$$\boxed{g = \mathcal{O} \mathcal{X}_c + \Gamma \mathcal{W}}$$

Now as \mathcal{O} and Γ are real valued

$$\underbrace{[\text{real}(\mathcal{G}) \quad \text{imag}(\mathcal{G})]}_{\mathcal{G}_r} = \mathcal{O} \underbrace{[\text{real}(\mathcal{X}_c) \quad \text{imag}(\mathcal{X}_c)]}_{\mathcal{X}_{cr}} + \Gamma \underbrace{[\text{real}(\mathcal{W}) \quad \text{imag}(\mathcal{W})]}_{\mathcal{W}_r}$$

If $n_y q < n_u r$ then $\exists \mathcal{W}_r^\perp$ such that $\mathcal{W}_r \mathcal{W}_r^\perp = 0$.

$$\mathcal{G}_r \mathcal{W}_r^\perp = (\mathcal{O} \mathcal{X}_{rc} + \Gamma \mathcal{W}_r) \mathcal{W}_r^\perp = \mathcal{O} \mathcal{X}_{cr} \mathcal{W}_r^\perp$$

$$\text{range}(\mathcal{G}_r \mathcal{W}_r^\perp) = \text{range}(\mathcal{O})$$

```
1 %To solve linear least squares problem: Ax = b
2 x = A\b;
```

11 CLOSED-LOOP ID

12 TIME-DOMAIN CORRELATION METHOD

13 PREDICTION ERROR METHODS

14 PARAMETER ESTIMATION STATISTICS

15 NOMENCLATURE

$y(k) = Gu(k)$	output signal	$[\quad]$
$u(k)$	input signal	$[\quad]$
G	plant	$[\quad]$
$\hat{G} = \frac{y}{u}$	estimated plant	$[\quad]$
$Y(e^{j\omega})$	output spectrum	$[\quad]$
$U(e^{j\omega})$	input spectrum	$[\quad]$
ZOH	zero order hold	
DAC	digital analog converter	
ADC	analog digital converter	

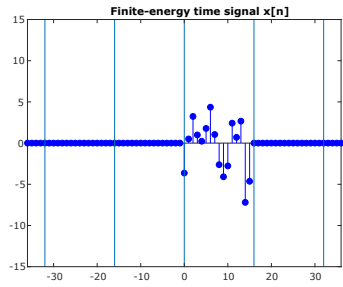
16 OVERVIEWS

16.1 TRANSFER FUNCTIONS

Nonperiodic Time Domain Signal

$$x[n] = \int_0^1 \hat{x}(\theta) e^{2\pi i n \theta} d\theta$$

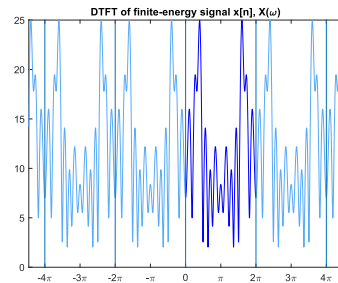
$$x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega k} d\omega$$



Discrete Time Fourier Transform (DTFT)

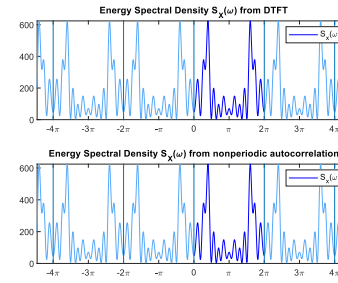
$$\hat{x}(\theta) = \sum_{n=-\infty}^{\infty} x[n] e^{-2\pi i n \theta}$$

$$X(\omega) = \sum_{k=-\infty}^{\infty} x(k) e^{-j\omega k}$$



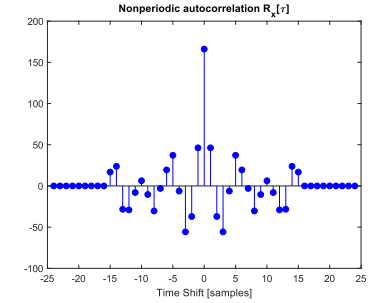
Energy Spectral Density

$$S_x(\omega) = |X(\omega)|^2 = \sum_{\tau=-\infty}^{\infty} R_x(\tau) e^{-j\omega \tau}$$



Nonperiodic Autocorrelation

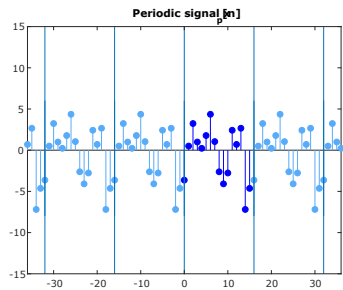
$$R_x(\tau) = \sum_{k=-\infty}^{\infty} x(k) x(k - \tau)$$



Periodic Time Domain Signal

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{x}[k] e^{2\pi i n \frac{k}{N}}$$

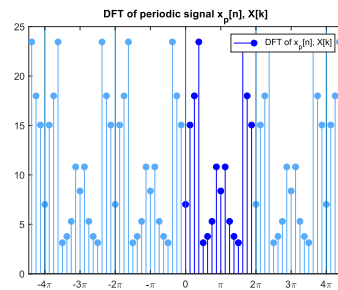
$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} X(\omega_n) e^{j\omega_n k}, \quad \omega_n = 2\pi \frac{n}{N}$$



Discrete Fourier Transform

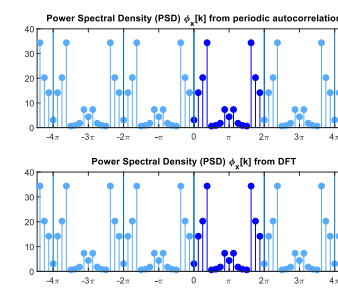
$$\hat{x}[k] = \sum_{n=0}^{N-1} x[n] e^{-2\pi i k \frac{n}{N}}$$

$$X(\omega_n) = \sum_{k=0}^{N-1} x(k) e^{-j\omega_n k}, \quad \omega_n = 2\pi \frac{n}{N}$$



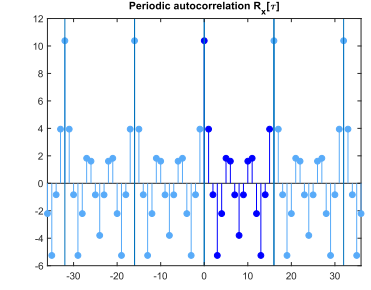
Power Spectral Density (PSD)

$$\phi(\omega_n) = \frac{1}{N} |X(\omega_n)|^2 = \sum_{\tau=0}^{N-1} R_x(\tau) e^{-j\omega_n \tau}$$



Periodic Autocorrelation

$$R_x(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) x(k - \tau)$$



fill in coolness chart