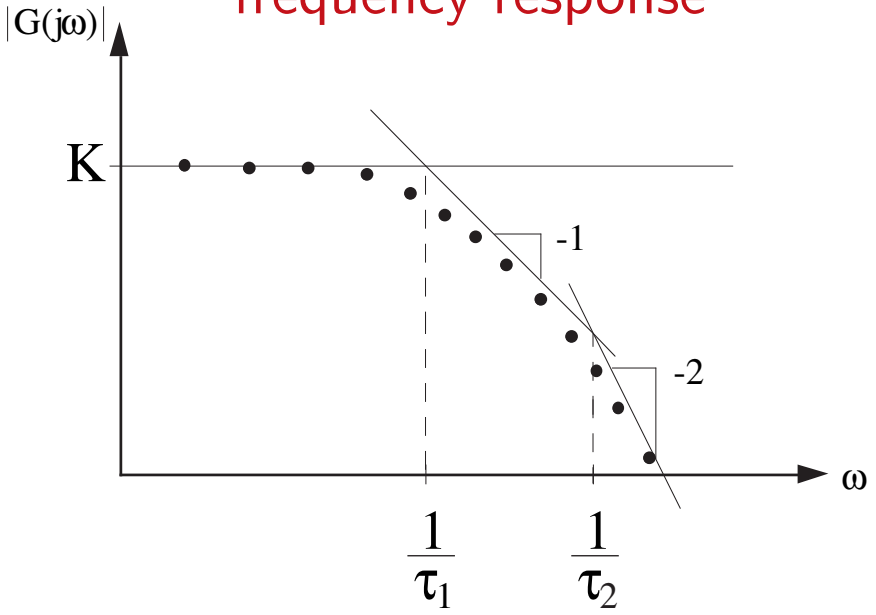
# Nonparametric models

**Parametric models :** Transfer functions and state-space models

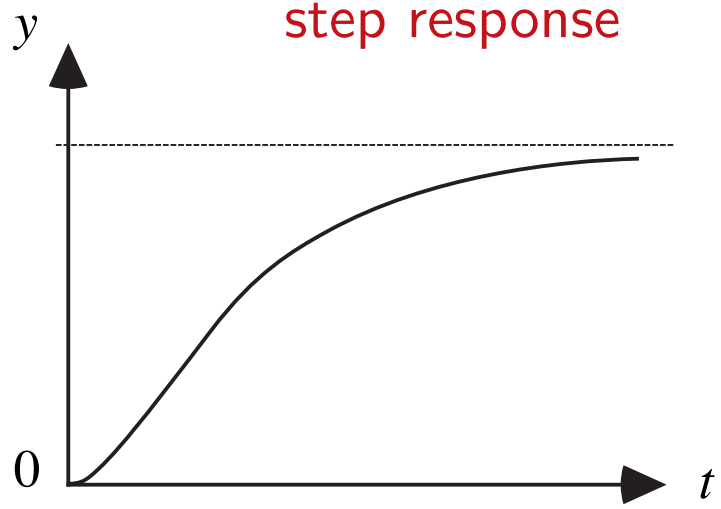
$$G(s) = \frac{K(s+1)}{(s+3)(s+5)} \quad ; \quad G(z) = \frac{z^2 + 2z + 3}{z^3 + 3z^2 + 3z + 8}$$

**Nonparametric models :** Graphs, data, etc.

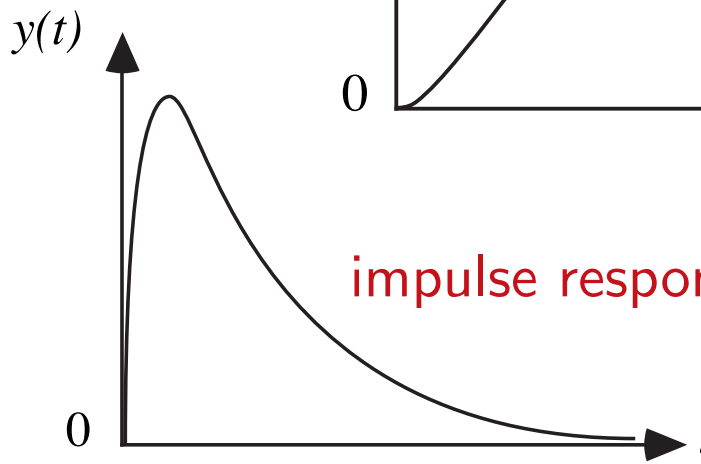
frequency response




step response




impulse response



# Parametric versus Nonparametric Models

**Parametric models :**

- ☺ Suitable for controller design, simulation, prediction etc.
- ☺ Different types of representation can be computed easily.
- ☹ Needs a priori assumptions about the model structure.
- ☹ Model parameters are the result of an optimization.
- ☺ Measurement noise can be filtered out in the optimization procedure.

**Nonparametric models :**

- ☺ Simple and fast, needs a few a priori information.
- ☹ Generally, no optimization is required. So, measurement noise comes directly to the model and cannot be filtered out by optimization.
- ☺ Suitable for controller design methods (PID, loop shaping, robust).
- ☹ Give only some general ideas about the system (bandwidth, order).

**Nonparametric models are necessary for parametric identification**

# Nonparametric Identification methods

## Time-domain methods :

**Step and impulse response are identified.**

- Transient response (step and impulse inputs)
- Deconvolution method (any input signal)
- Correlation approach (random signal)

## Frequency-domain methods :

**Frequency response is identified.**

- Frequency analysis (sinusoid signal )
- Fourier analysis (any input signal)
- Spectral analysis (random signal)

# Transient response

**Step response :** A step signal of amplitude $\alpha$ is applied to the plant.

- System should be in a stationary state.
- Noise level should be measured.
- How $\alpha$ should be selected ?
- Sampling period should be chosen as small as possible.

**Identify a simple parametric model from step response :**

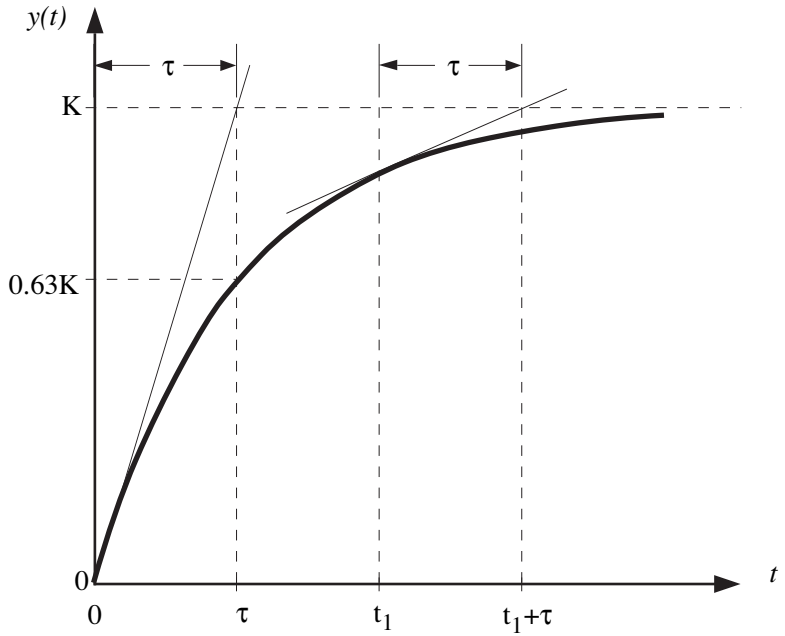
First order model :

$$G(s) = \frac{\gamma}{\tau s + 1}$$

Theoretical step response

$$y(t) = K(1 - e^{-t/\tau})$$

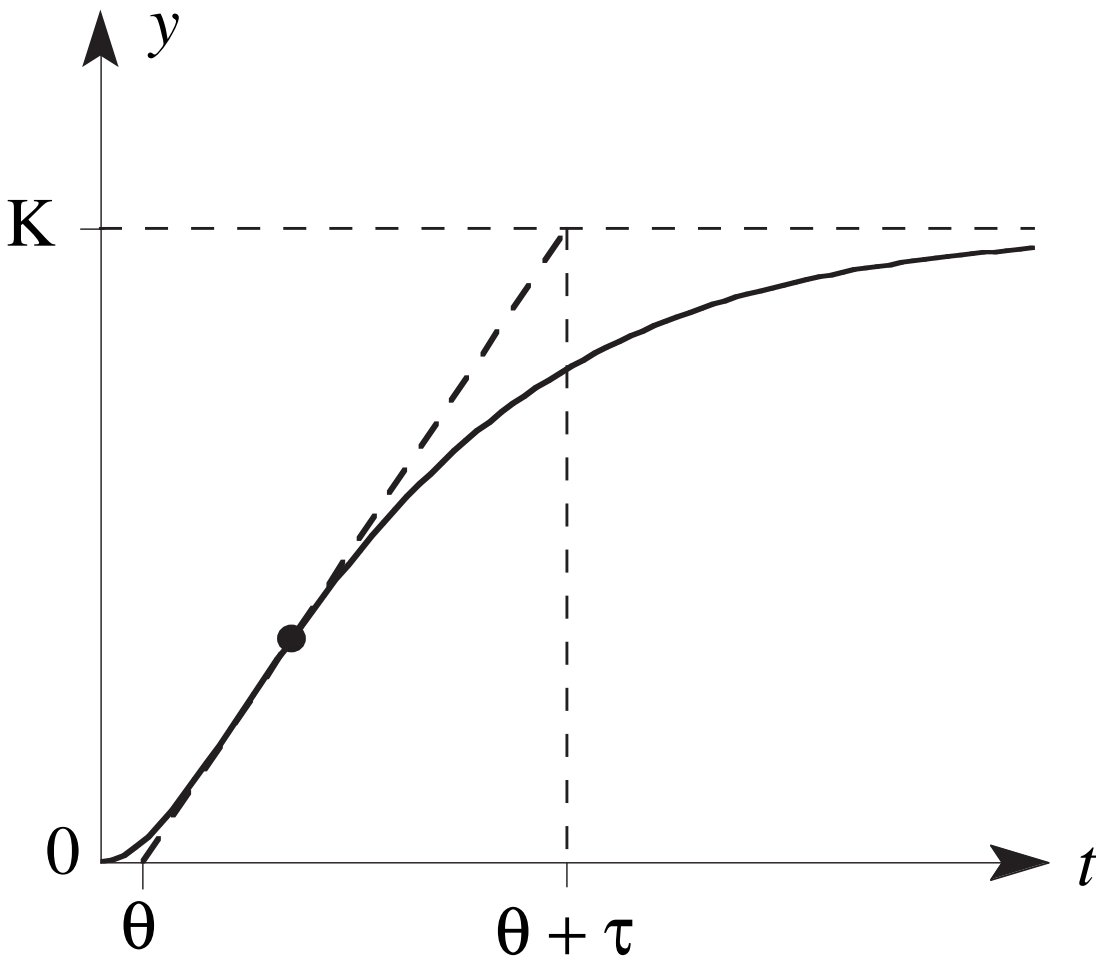
$$\Rightarrow \gamma = K/\alpha$$



# Step response

**Identify a simple parametric model from step response :**

First order model with delay (Ziegler-Nichols) : $G(s) \approx \dfrac{\gamma e^{-\theta s}}{\tau s + 1}$



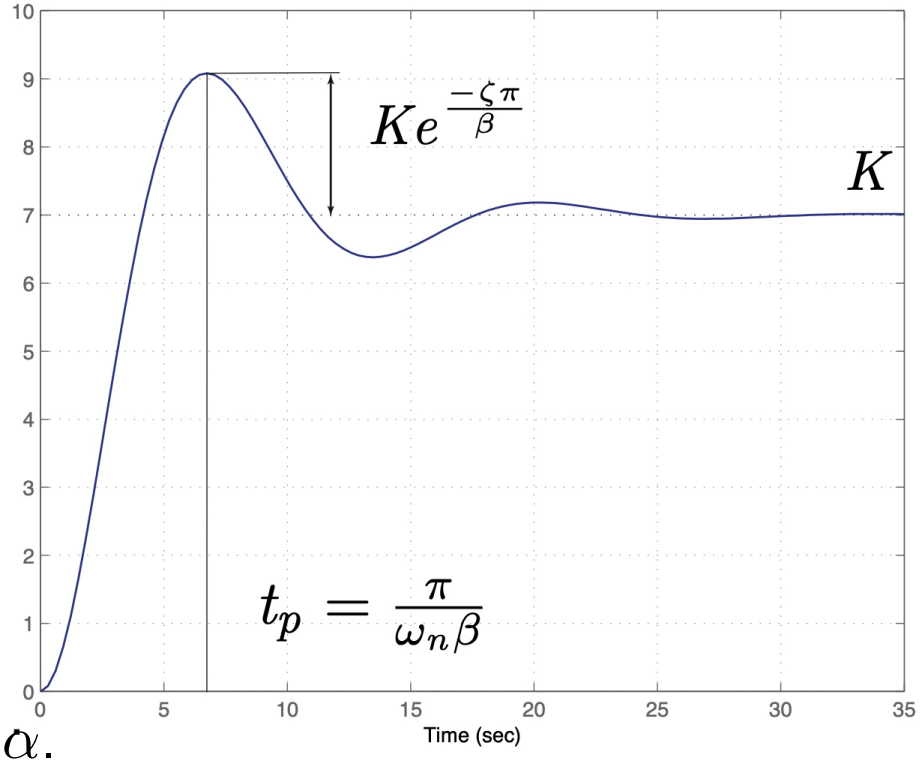
# Step response

**Identify a second-order model from step response :**

$$G(s) = \frac{\gamma\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

$$y(t) = K - \frac{K}{\beta} e^{-\zeta\omega_n t} \sin(\omega_n \beta t + \theta)$$

$$\beta = \sqrt{1-\zeta^2} \qquad \theta = \arccos\zeta$$

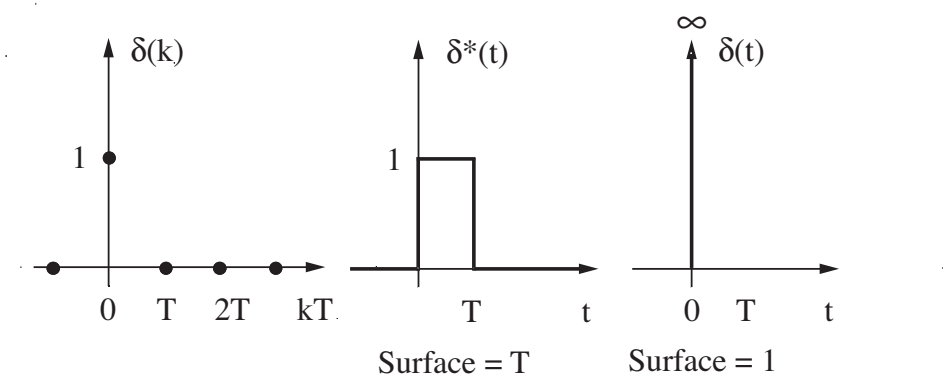


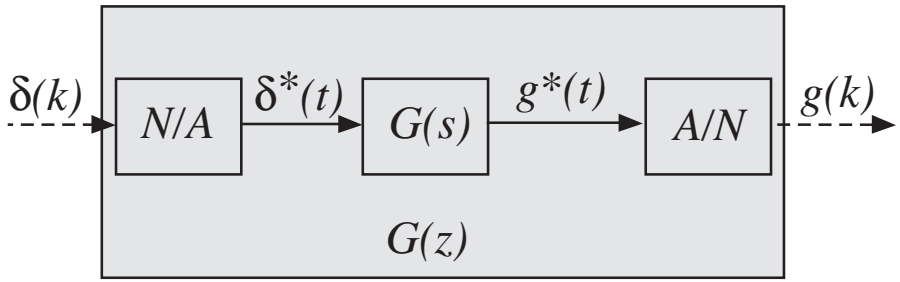

**Identification procedure :**

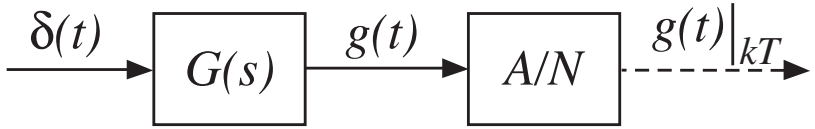
1. Measure $K$ and compute $\gamma = K/\alpha$.
2. Measure $t_p$ and $y(t_p)$ and compute the overshoot $M_p = \dfrac{y(t_p) - K}{K}$
3. Compute the damping factor $M_p = e^{-\zeta\pi/\beta} \quad \Rightarrow \quad \zeta^2 = \dfrac{\ln^2(M_p)}{\pi^2 + \ln^2(M_p)}$
4. Compute the natural frequency $\omega_n = \dfrac{\pi}{t_p\beta}$

# Impulse response

## Relation between $g(k)$ and $g(t)$




$$\lim_{T\to 0} \frac{1}{T} g^*(t) = g(t) \quad \Rightarrow \quad g(k) \approx Tg(t)\Big|_{kT}$$

# Impulse response

- It is impossible to apply the Dirac impulse, $\delta(t)$ to a system.
- The impulse response, $g(t)$ is the derivative of the step response.
- A scaled Kronecker delta, $\alpha\delta(k)$, can be applied to a discrete-time system.
- Then , the discrete-time impulse response is given by :

$$g(k) = \frac{1}{\alpha} g^*(t)|_{kT}$$

- If $y(k)$ is the step response of a step of amplitude $\alpha$, then :

$$g(k) = \frac{1}{\alpha}[y(k) - y(k-1)]$$

- The impulse response $g(t)$ is given by :

$$g(t) \approx \frac{1}{\alpha T} g^*(t)$$

**What can we do if the noise level is high ?**

# Numerical deconvolution

Use any input $u(k)$ and compute $g(k)$ from the convolution sum :

$$y(k) = g(k) * u(k) = \sum_{j=0}^{k} g(j)u(k-j) \qquad k = 0, 1, 2, \ldots, N-1$$

$$\begin{aligned} y(0) &= g(0)u(0) \\ y(1) &= g(0)u(1) + g(1)u(0) \\ y(2) &= g(0)u(2) + g(1)u(1) + g(2)u(0) \\ &\vdots \end{aligned}$$

$g(k)$ can also be computed recursively, note that $u(0) \neq 0$ :

$$g(k) = \frac{1}{u(0)} \left[ y(k) - \sum_{j=0}^{k-1} g(j)u(k-j) \right] \qquad k = 0, 1, \ldots N-1$$

The approach is too sensitive to numerical errors, unless for some specific inputs. If $u(k)$ is a unit step : $g(k) = y(k) - \sum_{j=0}^{k-1} g(j) = y(k) - y(k-1)$

# Numerical deconvolution

In the matrix form, we have :

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} = \begin{bmatrix} u(0) & 0 & \cdots & 0 \\ u(1) & u(0) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ u(N-1) & u(N-2) & \cdots & u(0) \end{bmatrix} \begin{bmatrix} g(0) \\ g(1) \\ \vdots \\ g(N-1) \end{bmatrix}$$

$$Y = U\Theta \Rightarrow \Theta = U^{-1}Y$$

**Remark :** The asymmetric Toeplitz matrix $U$ is ill-conditioned so $\Theta$ cannot be computed.

There are two solutions to this problem :

1. Find the finite impulse response : If the length of impulse response is finite $K \ll N$, the number of unknowns, $g(k)$, can be reduced ($N$ equations, $K$ unknowns) and so the problem can be solved in the least square sense.
2. Regularization (Bias-Variance Trade-off) : When $U$ is close to singular the norm of $\Theta = U^{-1}Y$ will go to infinity. A solution is to minimize the norm of $\Theta$ in the least squares solution as well.

# Numerical deconvolution

**Find the finite impulse response :** For stable systems (without integrator), impulse response goes asymptotically to zero such that we can suppose $g(k) = 0 \quad \text{for} \quad k > K$. Then, we can write :

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} = \begin{bmatrix} u(0) & 0 & \cdots & 0 \\ u(1) & u(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ u(N-1) & u(N-2) & \cdots & u(N-K) \end{bmatrix} \begin{bmatrix} g(0) \\ g(1) \\ \vdots \\ g(K-1) \end{bmatrix}$$

or $Y = U_K \Theta_K$, where $U_K$ and $\Theta_K$ are the truncated version of $U$ and $\Theta$. Now, the number of equations, $N$, is more than the number of unknowns $K$ (i.e. $U_K$ is a $N \times K$ dimensional matrix). This problem can be solved in the least squares sense by computing the pseudo inverse of $U_K$ :

$$\Theta_K = U_K^{\dagger} Y = (U_K^T U_K)^{-1} U_K^T Y$$

# Numerical deconvolution

**Regularization (Bias-Variance Trade-off) :** Consider a modified least squares criterion as :

$$J(\Theta) = \mathcal{E}^T \mathcal{E} + \lambda \Theta^T \Theta$$

where $\mathcal{E} = Y - U\Theta$, and $\lambda$ is a weighting factor for bias-variance trade-off. Then we have :

$$\frac{\partial J}{\partial \Theta} = -2U^T Y + 2U^T U\Theta + 2\lambda\Theta = 0 \quad \Rightarrow \quad \Theta = (U^T U + \lambda I)^{-1} U^T Y$$

A non zero $\lambda$ makes $U^T U + \lambda I$ a regular matrix, which is invertible and leads to a finite norm $\Theta$.

A good value for $\lambda$ can be found by trial and error. Usually, we should have an idea about the length of impulse response to find a good value for $\lambda$.

# Correlation approach

## Cross-Correlation Function

For stochastic signals $x(k)$ and $y(k)$ is defined as

$$R_{xy}(h) = \mathbb{E}\{x(k)y(k-h)\} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} x(k)y(k-h)$$

## Independence

If $x$ and $y$ are independent :

$$R_{xy}(h) = \mathbb{E}\{x(k)\}\mathbb{E}\{y(k-h)\}$$

In addition, if one of them has zero mean :

$$R_{xy}(h) = \mathbb{E}\{x(k)y(k-h)\} = 0 \qquad \forall h$$

# Correlation approach

## Properties of the correlation functions

$$\boxed{R_{xx}(h) = R_{xx}(-h) \quad ; \quad R_{xy}(h) = R_{yx}(-h) \quad ; \quad |R_{xx}(h)| \leq R_{xx}(0)}$$

The first two properties can be proved by defining $k' = k - h$ :

$$\begin{aligned} R_{xx}(h) &= \mathbb{E}\{x(k)x(k-h)\} = \mathbb{E}\{x(k'+h)x(k')\} = \mathbb{E}\{x(k')x(k'-(-h))\} \\ &= R_{xx}(-h) \\ R_{xy}(h) &= \mathbb{E}\{x(k)y(k-h)\} = \mathbb{E}\{x(k'+h)y(k')\} = \mathbb{E}\{y(k')x(k'-(-h))\} \\ &= R_{yx}(-h) \end{aligned}$$

The last property can be proved using $\mathbb{E}\{[x(k) \pm x(k-h)]^2\} \geq 0$

$$\begin{aligned} &\Rightarrow \quad \mathbb{E}\{x^2(k)\} \pm 2\mathbb{E}\{x(k)x(k-h)\} + \mathbb{E}\{x^2(k-h)\} \geq 0 \\ &\Rightarrow \quad R_{xx}(0) \pm 2R_{xx}(h) + R_{xx}(0) \geq 0 \quad \Rightarrow \quad -R_{xx}(0) \leq R_{xx}(h) \leq R_{xx}(0) \\ &\Rightarrow \quad |R_{xx}(h)| \leq R_{xx}(0) \end{aligned}$$

# Correlation approach

**Principle :** The noisy output of an LTI system is given by :

$$y(k) = \sum_{j=0}^{\infty} g(j)u(k-j) + d(k) = g(k) * u(k) + d(k)$$

Take $u(k)$ as a random signal independent of the zero-mean $d(k)$.
Multiply $y(k)$ by $u(k-h)$ and take the expected value :

$$\mathbb{E}\{y(k)u(k-h)\} = \sum_{j=0}^{\infty} g(j)\mathbb{E}\{u(k-j)u(k-h)\} + \mathbb{E}\{d(k)u(k-h)\}$$

$$R_{yu}(h) = \sum_{j=0}^{\infty} g(j)R_{uu}(-j+h) + R_{du}(h)$$

Note that $R_{du}(h) = 0$, therefore :

$$R_{yu}(h) = \sum_{j=0}^{\infty} g(j)R_{uu}(h-j) \quad \Rightarrow \quad \boxed{R_{yu}(h) = g(h) * R_{uu}(h)}$$

# Correlation approach

## Removing the noise effect

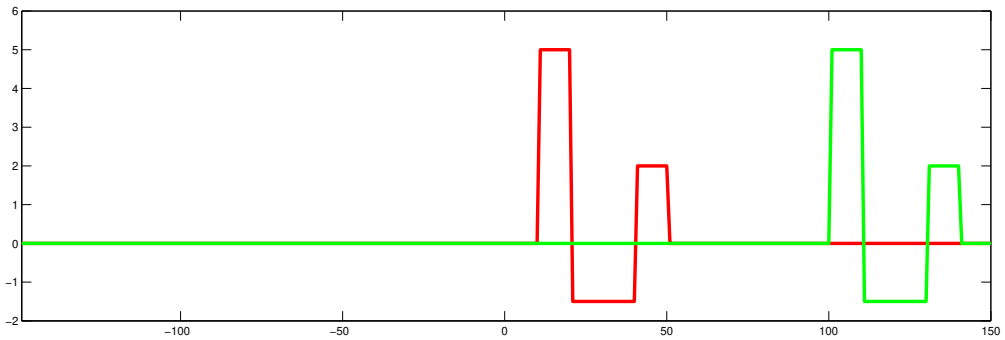
Input/output relation : $y(k) = g(k) * u(k) + d(k)$

Auto/cross-correlation relation : $R_{yu}(h) = g(h) * R_{uu}(h)$

$g(h)$ can be computed using the numerical deconvolution

**Main problem :** How to estimate the correlation functions ?

## Example



# Estimation of the correlation functions

## Autocorrelation function (biased estimate)

$$\hat{R}_{uu}(h) = \frac{1}{N}\sum_{k=h}^{N-1} u(k)u(k-h) \quad , \quad \hat{R}_{uu}(-h) = \hat{R}_{uu}(h) \quad , \quad h = 0, 1, 2, \ldots$$

This estimate is biased because :

$$\mathbb{E}\{\hat{R}_{uu}(h)\} = \frac{1}{N}\sum_{k=h}^{N-1} \mathbb{E}\{u(k)u(k-h)\} = \frac{N-h}{N} R_{uu}(h) \neq R_{uu}(h)$$

## Autocorrelation function (unbiased estimate)

$$\hat{R}_{uu}(h) = \frac{1}{N-h}\sum_{k=h}^{N-1} u(k)u(k-h) \quad , \quad \hat{R}_{uu}(-h) = \hat{R}_{uu}(h) \quad , \quad h = 0, 1, 2, \ldots$$

# Correlation approach

## Example

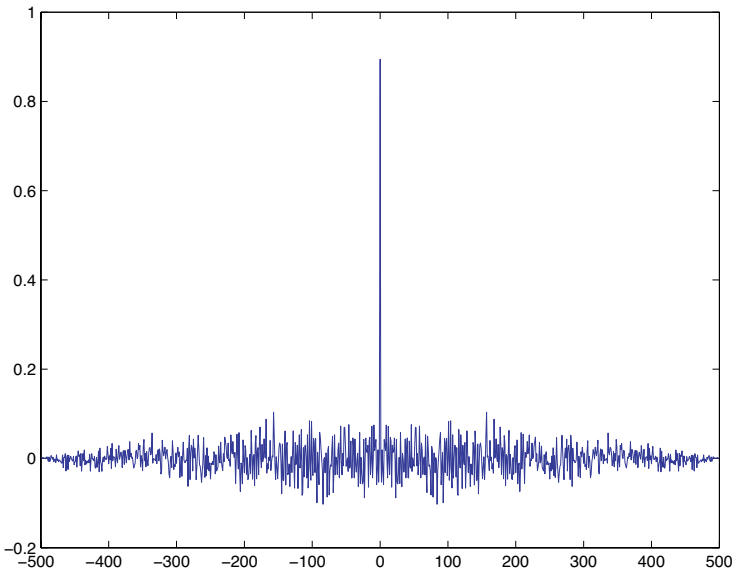
Autocorrelation function estimate of a zero mean discrete white noise $e(k)$ with $N = 500$.

Biased estimate

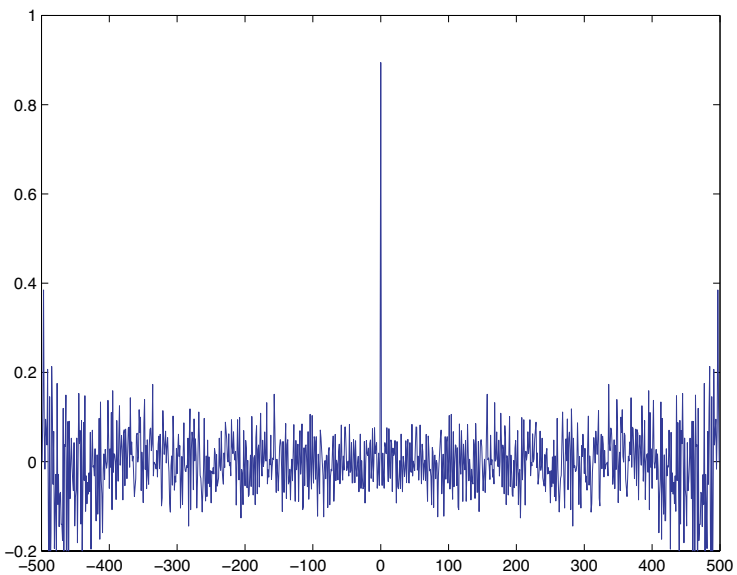
$$\hat{R}_{ee}(h) = \frac{1}{N} \sum_{k=h}^{N-1} e(k)e(k-h)$$

Unbiased estimate

$$\hat{R}_{ee}(h) = \frac{1}{N-h} \sum_{k=h}^{N-1} e(k)e(k-h)$$




(a)




(b)

# Correlation approach

## Cross-correlation function (biased estimate)

$$\hat{R}_{yu}(h) = \frac{1}{N} \sum_{k=\max(h,0)}^{N-1+\min(h,0)} y(k)u(k-h) \quad , \quad h = 0, \pm 1, \pm 2, \ldots$$

## Cross-correlation function (unbiased estimate)

$$\hat{R}_{yu}(h) = \frac{1}{N-|h|} \sum_{k=\max(h,0)}^{N-1+\min(h,0)} y(k)u(k-h) \quad , \quad h = 0, \pm 1, \pm 2, \ldots$$

## Impulse response estimation

If $u(k)$ is white noise :

$$R_{yu}(h) = g(h) * \delta(h) R_{uu}(0) = g(h) R_{uu}(0) \quad \Rightarrow \quad g(h) = \frac{\hat{R}_{yu}(h)}{\hat{R}_{uu}(0)}$$

# Correlation approach

## Impulse response estimation

If $u(k)$ is not white, numerical deconvolution can be used. It is assumed that for $j \geq K$ we have $g(j) = 0$.

$$R_{yu}(h) = \sum_{j=0}^{K-1} g(j) R_{uu}(h-j) \qquad\qquad h = 0, 1, 2, \ldots$$

$$\begin{aligned}
\hat{R}_{yu}(0) &= \hat{R}_{uu}(0)g(0) + \hat{R}_{uu}(-1)g(1) + \cdots + \hat{R}_{uu}(-K+1)g(K-1) \\
\hat{R}_{yu}(1) &= \hat{R}_{uu}(1)g(0) + \hat{R}_{uu}(0)g(1) + \cdots + \hat{R}_{uu}(-K+2)g(K-1) \\
&\vdots
\end{aligned}$$

$$\begin{bmatrix} \hat{R}_{yu}(0) \\ \hat{R}_{yu}(1) \\ \vdots \\ \hat{R}_{yu}(K-1) \end{bmatrix} = \begin{bmatrix} \hat{R}_{uu}(0) & \hat{R}_{uu}(1) & \cdots & \hat{R}_{uu}(K-1) \\ \hat{R}_{uu}(1) & \hat{R}_{uu}(0) & \cdots & \hat{R}_{uu}(K-2) \\ \vdots & \vdots & & \vdots \\ \hat{R}_{uu}(K-1) & \hat{R}_{uu}(K-2) & \cdots & \hat{R}_{uu}(0) \end{bmatrix} \begin{bmatrix} g(0) \\ g(1) \\ \vdots \\ g(K-1) \end{bmatrix}$$

# Frequency-domain methods

## Frequency analysis

The input signal is a sinusoidal signal which leads, in steady state, to a sinusoidal output.

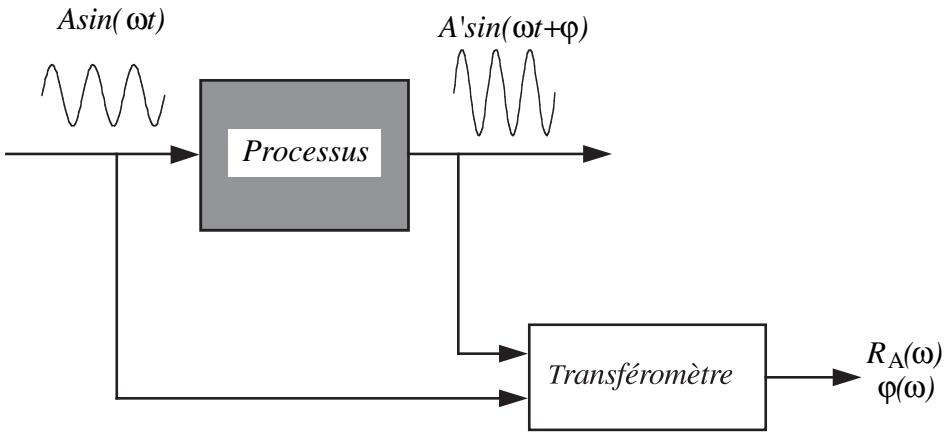
## Fourier analysis

A deterministic signal, preferably periodic, that excites a large number of frequencies (sum of sinusoids or PRBS) is applied as input.

## Spectral analysis

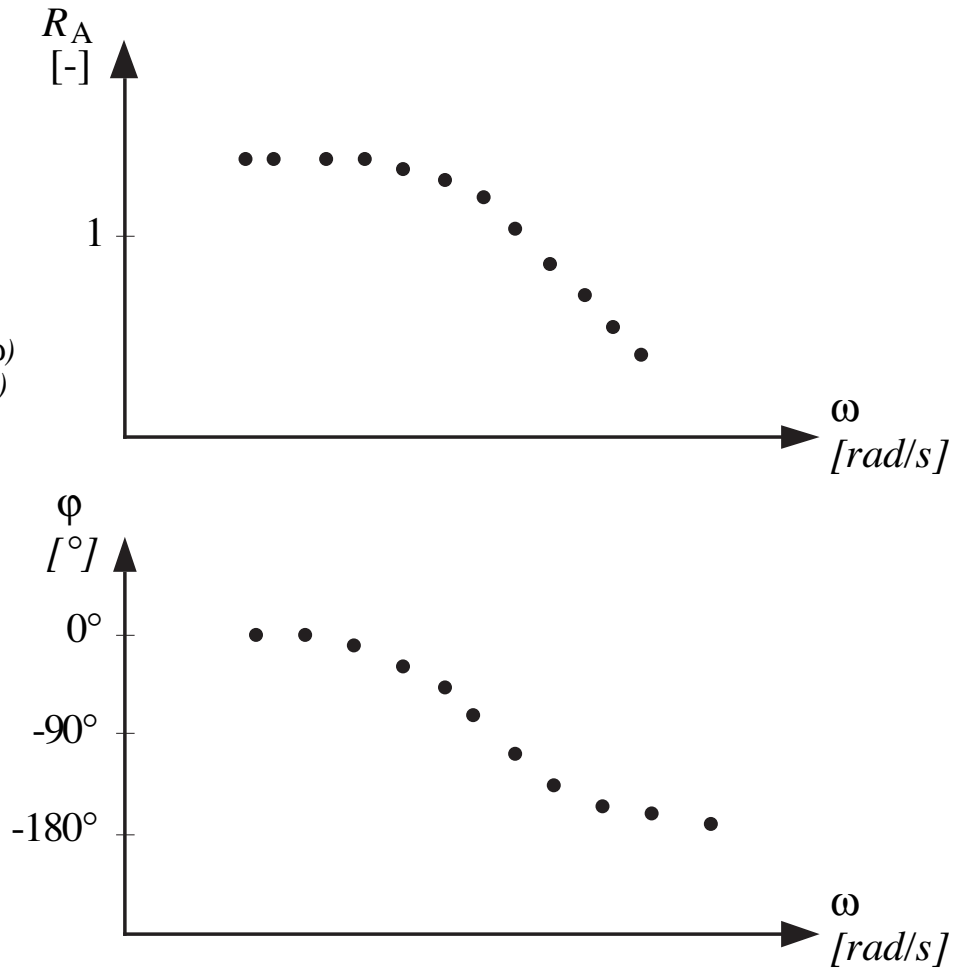
The input and output of the system are considered as stochastic signals (input is a random process and the output is affected by measurement noise).

# Frequency analysis




$$R_A(\omega) = \frac{A'}{A} = |G(j\omega)|$$

$$\varphi(\omega) = \arg[G(j\omega)]$$



# Frequency analysis

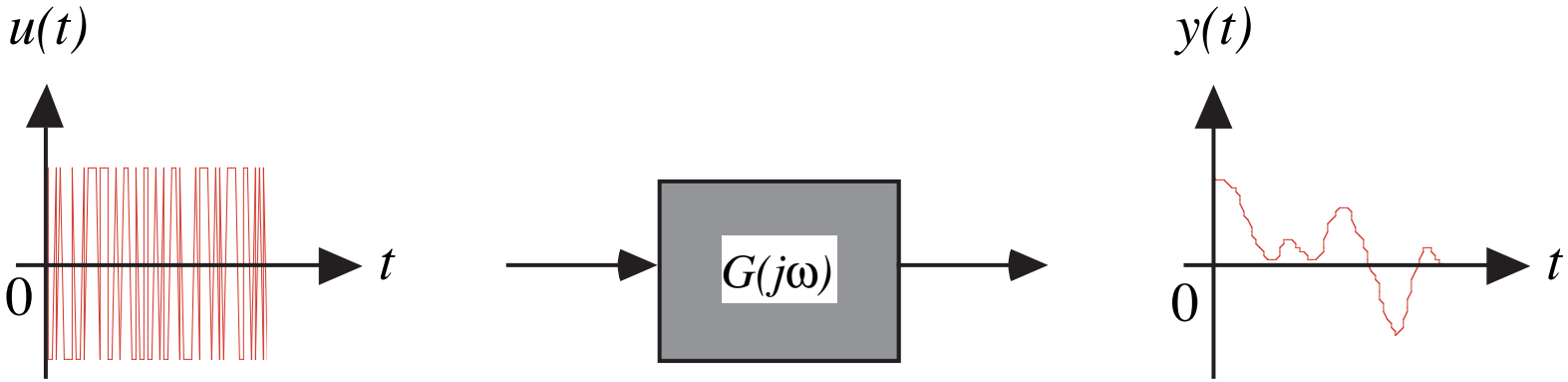
## Cautions

- Frequency range should cover at least 2 or 3 decades with more than 10 frequency points per decade.
- The magnitude of input signal should be chosen regarding the noise level and system nonlinearity.
- The measurements should be done in steady state.

## Advantages and disadvantages

- Simple and intuitive.
- Needs only a few a priori information about the plant.
- The experimentation time is too long especially for slow systems.

# Fourier analysis




**Principle :** For LTI continuous-time systems, we have :

$$y(t) = \int_0^\infty g(\tau)u(t-\tau)d\tau = g(t) * u(t)$$

Taking the Fourier transform, we obtain :

$$Y(j\omega) = G(j\omega)\,U(j\omega) \quad \Rightarrow \quad G(j\omega) = \frac{Y(j\omega)}{U(j\omega)}$$

where

$$U(j\omega) = \int_{-\infty}^{\infty} u(t)e^{-j\omega t}\,dt \quad ; \quad Y(j\omega) = \int_{-\infty}^{\infty} y(t)e^{-j\omega t}\,dt$$

# Fourier analysis

**Principle :** For LTI discrete-time systems, we have :

$$y(k) = \sum_{j=0}^{\infty} g(j)u(k-j) = g(k) * u(k)$$

Taking the Fourier transform, we obtain :

$$Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) \quad \Rightarrow \quad G(e^{j\omega}) = \frac{Y(e^{j\omega})}{U(e^{j\omega})}$$

where

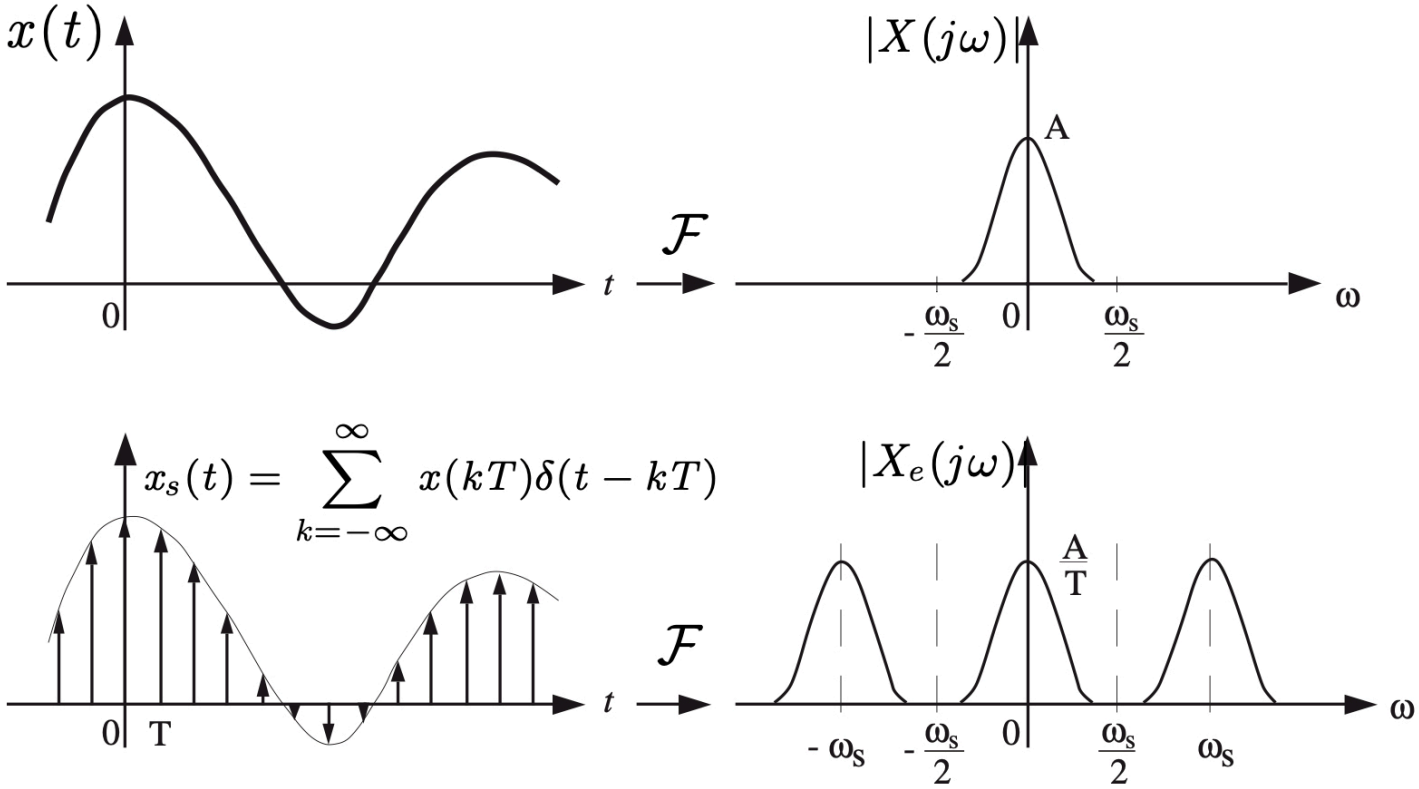
$$U(e^{j\omega}) = \sum_{k=-\infty}^{\infty} u(k)e^{-jk\omega T} \quad ; \quad Y(e^{j\omega}) = \sum_{k=-\infty}^{\infty} y(k)e^{-jk\omega T}$$

**Three types of error in Fourier analysis :**

1. Sampling error
2. Measurement error
3. Truncation error

# Sampling error

**Source of error :** We identify $G(e^{j\omega})$ instead of $G(j\omega)$.




If the condition of Shannon Theorem is satisfied then $G(e^{j\omega}) = G(j\omega)$.

**The condition of Shannon Theorem is never satisfied for real systems ! There is always a sampling error.**

# Measurement error

**Source of error :** Noise and disturbance.

$$y(k) = \sum_{j=0}^{\infty} g(j)u(k-j) + d(k) \quad \overset{\mathcal{F}}{\Rightarrow} \quad Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) + D(e^{j\omega})$$

$$\hat{G}(e^{j\omega}) = \frac{Y(e^{j\omega})}{U(e^{j\omega})} = G(e^{j\omega}) + \frac{D(e^{j\omega})}{U(e^{j\omega})}$$

- The frequency response estimate is unbiased if $d(k)$ is zero-mean.

$$\mathbb{E}\{\hat{G}(e^{j\omega}) - G(e^{j\omega})\} = \frac{\mathbb{E}\{D(e^{j\omega})\}}{U(e^{j\omega})} = 0$$

- The variance of the frequency response estimates are given by :

$$\text{var}[\hat{G}(e^{j\omega})] = \frac{\mathbb{E}\{D(e^{j\omega})D(e^{-j\omega})\}}{U(e^{j\omega})U(e^{-j\omega})} = \frac{\mathbb{E}\{|D(e^{j\omega})|^2\}}{|U(e^{j\omega})|^2} = \frac{\Phi_{dd}(\omega)}{\Phi_{uu}(\omega)}$$

- **Measurement error is large in frequencies that the noise spectrum is large and the spectrum of input signal is small.**
- **Increasing number of data will not decrease the variance.**

# Truncation error

**Fourier transform for aperiodic discrete-time signals :**

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

**Fourier transform for periodic discrete-time signals :**

$$X(e^{j\omega_n}) = \frac{1}{M}\sum_{k=0}^{M-1} x(k)e^{-jk\omega_n T} \qquad \omega_n = \frac{2\pi n}{MT} \qquad n = 0, 1, \ldots, M-1$$

**Source of error :** For aperiodic signals the Fourier transform is defined from $-\infty$ to $\infty$ but it is computed in a truncated interval $[0\,,\ N-1]$.

## Solutions

- Excitation with energy signals
- Excitation with periodic signals
- Windowing

# Truncation error

## Excitation with energy signals

Excite the system with a time limited signal such that $u(k) = 0$ for $k < 0$ and $k \geq M$ then the Fourier transform can be computed exactly (no truncation error).

$$U(e^{j\omega}) = \sum_{k=0}^{M-1} u(k) e^{-jk\omega T}$$

- If the system is stable, then $y(k) \approx 0$ for large $k \geq N$.
- So wait until the output returns to the steady state and record all data.
- This way, the truncation error will be negligible for the output (only the noise will be truncated !).

# Truncation error

## Excitation with periodic signals

- If the input signal is periodic with period $M$, its Fourier series can be computed exactly (no truncation error) :

$$U(e^{j\omega_n}) = \frac{1}{M}\sum_{k=0}^{M-1} u(k)e^{-jk\omega_n T} \qquad \omega_n = \frac{2\pi n}{MT} \qquad n = 0, 1, \ldots, M-1$$

- The output at steady-state will be almost periodic (ignoring the noise) :

$$Y(e^{j\omega_n}) = \frac{1}{M}\sum_{k=L}^{L+M-1} y(k)e^{-jk\omega_n T} \qquad \omega_n = \frac{2\pi n}{MT} \qquad n = 0, 1, \ldots, M-1$$

where $L$ is larger than the settling time of the system.

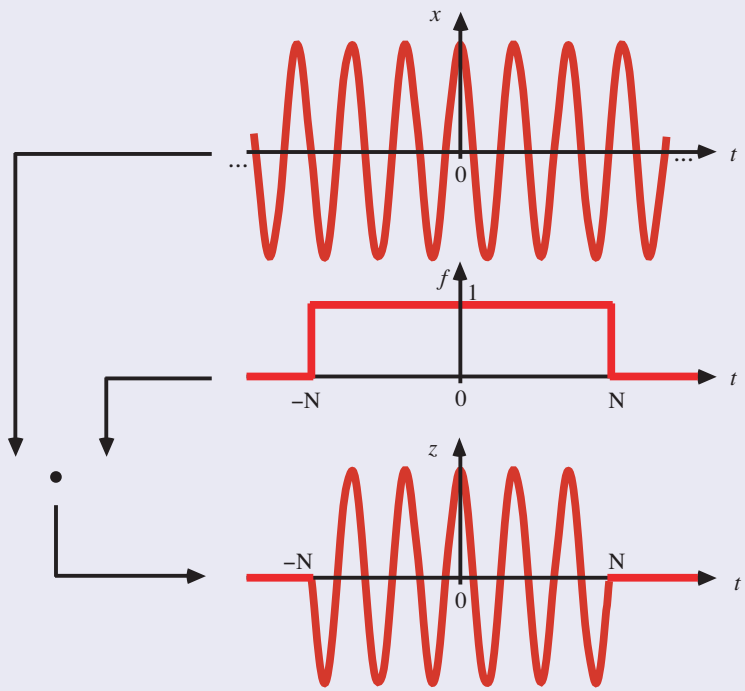
- The noise effect can be reduced by averaging over $p$ periods :

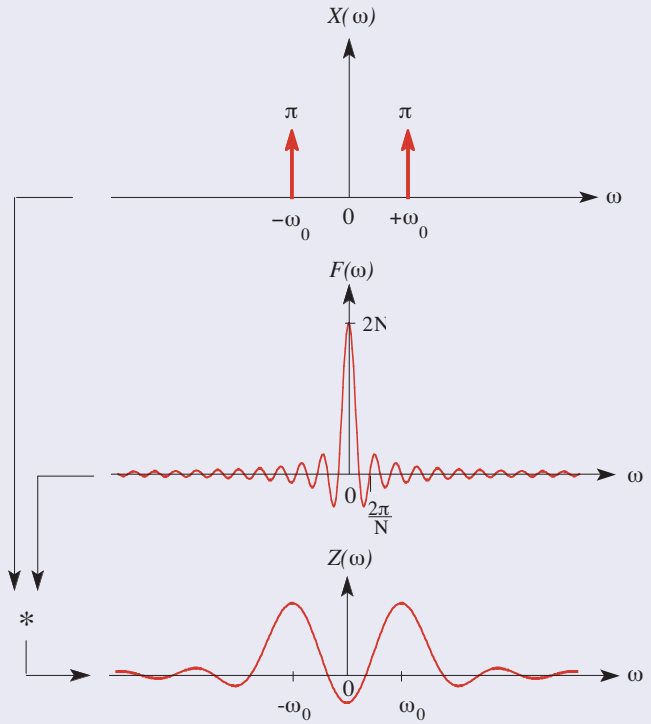
$$G(e^{j\omega_n}) = \frac{\frac{1}{p}\sum_{i=1}^{p} Y_i(e^{j\omega_n})}{U(e^{j\omega_n})} \qquad \omega_n = \frac{2\pi n}{MT} \qquad n = 0, 1, \ldots, M-1$$

# Truncation error

## Windowing

Truncation can be viewed as multiplying the signal in the time-domain by a rectangular window. It results in a convolution in the frequency domain.




$$z(t) = x(t)f(t) \quad \Rightarrow Z(\omega) = X(\omega) * F(\omega) = \int_{-\infty}^{\infty} X(\nu)F(\omega - \nu)d\nu$$

# Truncation error

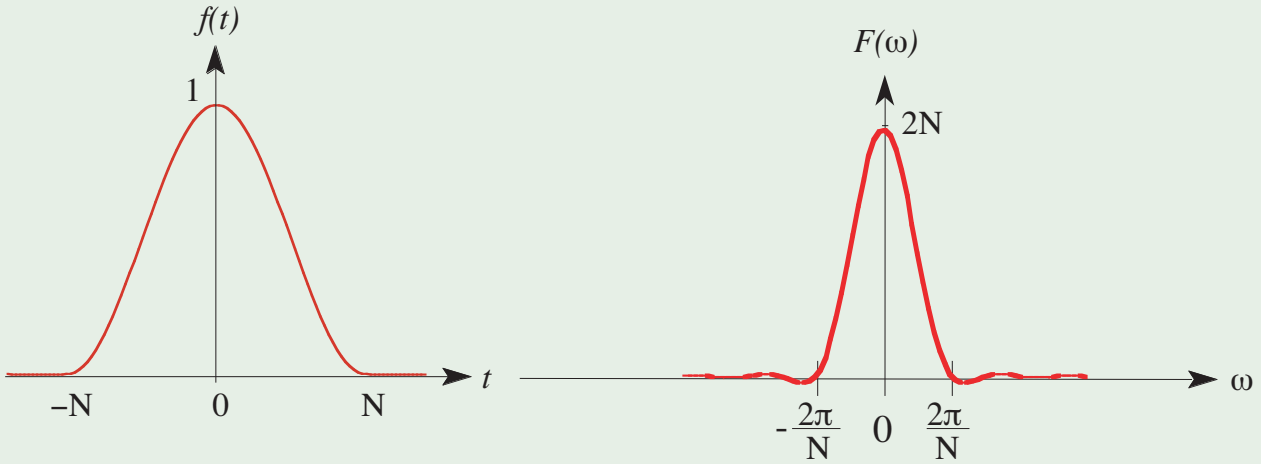
## Windowing introduces two errors :

Loss of resolution : is related to the main lobe width (MLW) and has a smoothing effect.

Spectral leakage : is related to the second lobe relative amplitude (SLA), and generates some errors in the neighboring frequencies.

**Trade-off :** The spectral leakage can be reduced by increasing the MLW, which reduces the resolution. Since, the loss of resolution has an smoothing effect and reduces the noise, increasing MLW is usually preferred.

## Example (Hann window)



# Practical Aspects of Fourier Analysis

**Discrete Fourier Transform (DFT) :** The implementation formula for computing the Fourier transform of a sampled signal of length $N$ is :

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-jk\frac{2\pi n}{N}} \qquad n = 0, 1, \ldots, N-1$$

The complex vector $X$ can be associated to a frequency vector as :

$$\left[0, \frac{\omega_s}{N}, \frac{2\omega_s}{N}, \cdots, \frac{(N-1)\omega_s}{N}\right]$$

where $\omega_s$ is the sampling frequency.

- DFT gives an approximation of the Fourier transform for continuous-time signals.
- DFT samples uniformly $N$ points on the Fourier transform of an aperiodic discrete-time signal.
- DFT gives $N$ times the exact value of Fourier transform for discrete-time periodic signals.

# Practical Aspects of Fourier Analysis

- Input signal should be rich enough to excite the system in a large frequency band. A good choice is a discrete white noise or a PRBS which has uniform spectrum.
- The best choice is to apply several periods of a periodic PRBS which is rich in all frequencies and has no truncation error. Moreover, if we use $m$ periods, the measurement error will be divided by $m$.
- In order to obtain a periodic signal for the output (for the deterministic part) we should let the transient response be finished.
- If we cannot apply a periodic signal, we should use Hamming or Hann or other windows to reduce the truncation error.
- It is always better to have less frequency data with more accuracy. For example, if we have enough data, we can divide data to 10 parts and compute $G_1$ to $G_{10}$ and then compute the average.

# Spectral analysis

**For a linear system we have :**

$$y(k) = g(k) * u(k) + d(k) \quad \Rightarrow \quad Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) + D(e^{j\omega})$$

**Using the correlation approach we have :**

$$R_{yu}(h) = g(h) * R_{uu}(h) \quad \Rightarrow \quad \Phi_{yu}(\omega) = G(e^{j\omega})\Phi_{uu}(\omega)$$

$$G(e^{j\omega}) = \frac{\Phi_{yu}(\omega)}{\Phi_{uu}(\omega)} = \frac{\sum_{h=-\infty}^{\infty} R_{yu}(h)e^{-j\omega hT}}{\sum_{h=-\infty}^{\infty} R_{uu}(h)e^{-j\omega hT}}$$

**In practice DFT is used to compute the estimates :**

$$\hat{G}(e^{j\omega_n}) = \frac{\hat{\Phi}_{yu}(\omega_n)}{\hat{\Phi}_{uu}(\omega_n)} = \frac{\sum_{h=0}^{N-1} \hat{R}_{yu}(h)e^{-jh\frac{2\pi n}{N}}}{\sum_{h=0}^{N-1} \hat{R}_{uu}(h)e^{-jh\frac{2\pi n}{N}}}; \quad \omega_n = \frac{n\omega_s}{N}; \quad n = 0, 1, \ldots, N-1$$

# Spectral analysis

## Estimating the noise spectrum

For an LTI discrete-time system, we have :

$$y(k) = g(k) * u(k) + d(k) \quad \Rightarrow \quad Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) + D(e^{j\omega})$$

Multiplying it with its conjugate and taking the expectation, leads to

$$\mathbb{E}\{|Y(e^{j\omega})|^2\} = \mathbb{E}\{|G(e^{j\omega})|^2|U(e^{j\omega})|^2\} + \mathbb{E}\{|D(e^{j\omega})|^2\}$$

Therefore, by using $\Phi_{yy}(\omega) = \mathbb{E}\{|Y(e^{j\omega})|^2\}$ :

$$\Phi_{dd}(\omega) = \Phi_{yy}(\omega) - \frac{|\Phi_{yu}(\omega)|^2}{\Phi_{uu}(\omega)}$$

**Remark :** For periodic signals, Fourier and Spectral analysis are equivalent !

$$G(e^{j\omega_n}) = \frac{\Phi_{yu}(\omega_n)}{\Phi_{uu}(\omega_n)} = \frac{Y(e^{j\omega_n})U(e^{-j\omega_n})}{U(e^{j\omega_n})U(e^{-j\omega_n})} = \frac{Y(e^{j\omega_n})}{U(e^{j\omega_n})}$$

# Improving the spectral analysis

## Averaging over several estimates

$$G(e^{j\omega_n}) = \frac{\frac{1}{m}\sum_{i=1}^{m}\Phi_{yu,i}(\omega_n)}{\frac{1}{m}\sum_{i=1}^{m}\Phi_{uu,i}(\omega_n)}$$

Averaging is done on spectral density functions and not in $G$, because using unbiased estimates for correlation functions the spectral density estimates are also unbiased so "averaging" will reduce the variance.

## Windows on correlation functions

Since for large $h$, the correlation function estimates are not accurate, a Hamming or Hann window on the correlation functions can be used. These windows give small weighting to the non accurate parts of correlation function estimates.

# Improving the spectral analysis

**Windowing :** The unbiased estimate of the correlation function has large variance for large $|h|$. They can be weighted with a window as :

$$\hat{\Phi}_{yu,f}(\omega_n) = \sum_{h=0}^{N-1} \hat{R}_{yu}(h) f(h) e^{-j2\pi nh/N}$$

where, e.g. $$f(h) = \begin{cases} 0.5 + 0.5\cos(\pi h/M) & h = 0, 1, \ldots, M \\ 0 & h > M \end{cases}$$

For a Hann window the variance of the estimates is :

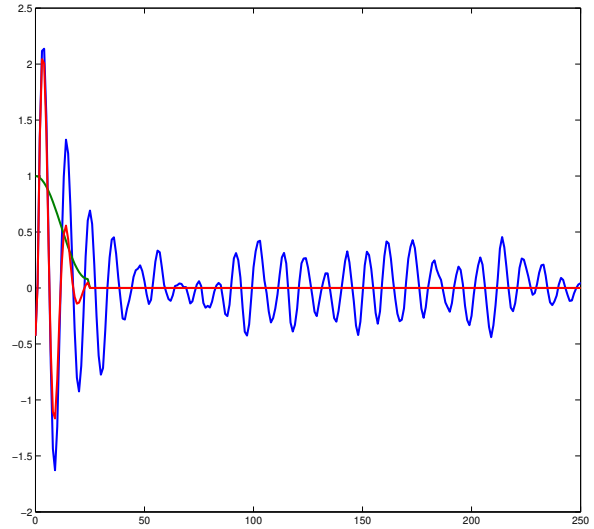
$$\text{var}[\hat{G}(j\omega)] \approx \frac{0.7M}{N} \frac{\Phi_{dd}(\omega)}{\Phi_{uu}(\omega)}$$

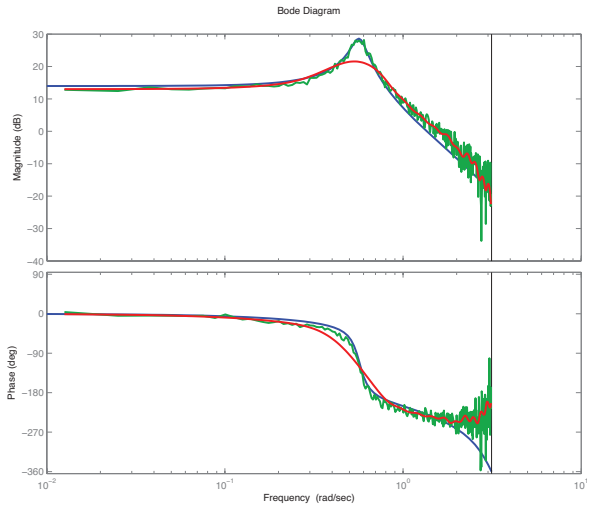
## Choice of $M$

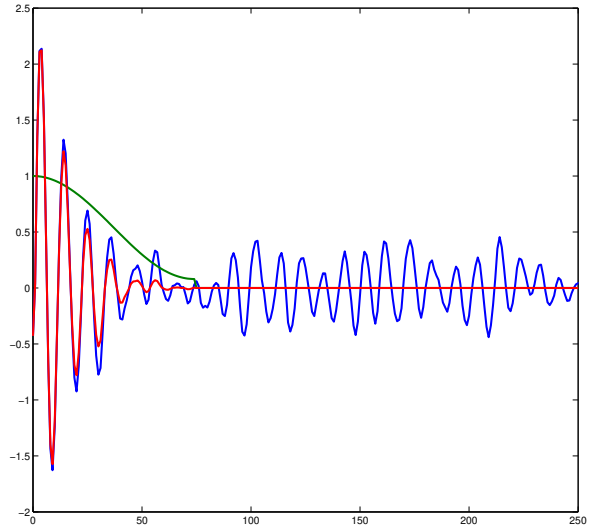
- Large $M$ increases the resolution as well as spectral leakage.
- Small $M$ reduces the noise effect and smooth the estimates.
- If the input signal is white (or PRBS), $M$ should be greater than the settling time of the impulse response of the system (the settling time $K$ can be estimated by observing $R_{yu}(h) \approx 0$ for $h > K$).

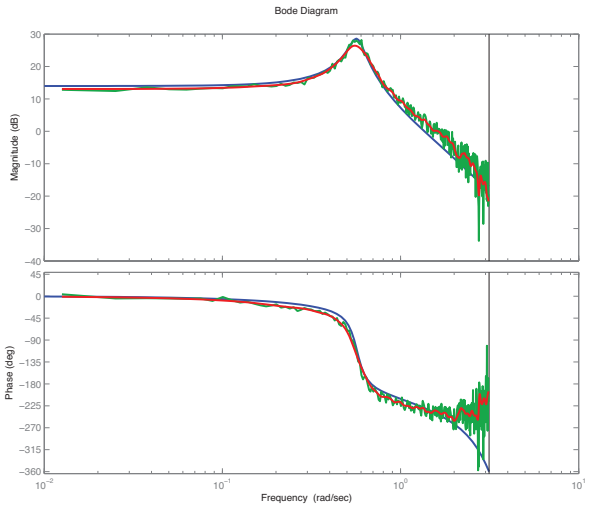
# Spectral analysis

## Example



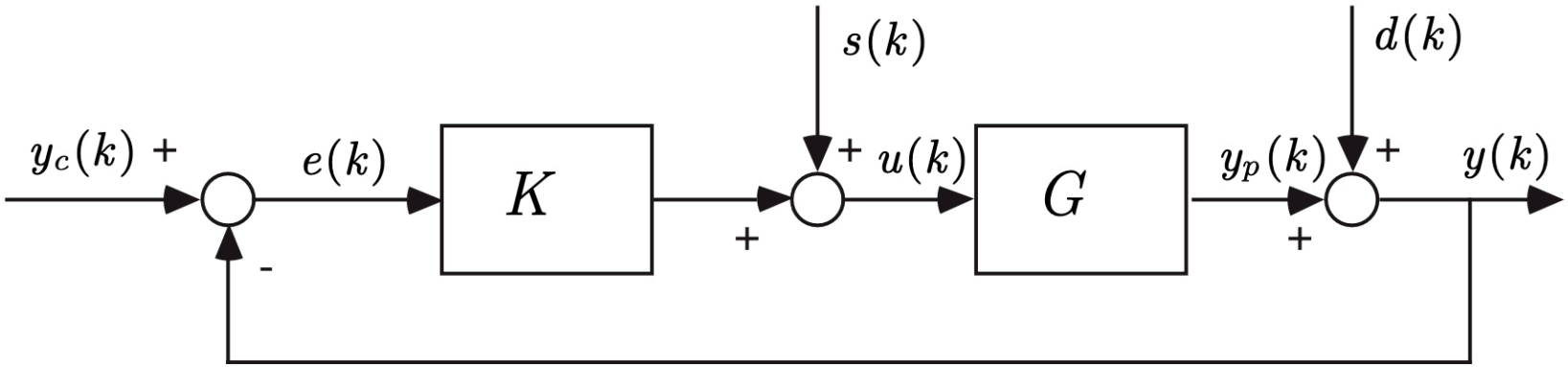



**Why are we interested in closed-loop identification ?**

- There are systems that are unstable in open-loop operation.
- The output drift occurs in some systems in open-loop operation.
- It is always better to identify an accurate model in the frequency zone interesting for control.




**Problem :** $u$ is correlated with $d$ via feedback $\left(R_{du}(h) \neq 0\right)$ so the spectral analysis cannot be applied directly on input/output data.

$$\hat{G}(e^{j\omega}) = \frac{\Phi_{yu}(\omega)}{\Phi_{uu}(\omega)} = \frac{\Phi_{y_p u}(\omega) + \Phi_{du}(\omega)}{\Phi_{uu}(\omega)} = G(e^{j\omega}) + \frac{\Phi_{du}(\omega)}{\Phi_{uu}(\omega)}$$

# Closed-loop identification

## Consider two extreme cases :

1. No disturbance : $d(k) = 0 \Rightarrow \Phi_{du}(\omega) = 0$, no error.
2. No excitation : $e(k) = -y(k)$ ; $\Phi_{uu}(\omega) = K(e^{j\omega})\Phi_{eu}(\omega)$.

$$\hat{G}(e^{j\omega}) = \frac{\Phi_{yu}(\omega)}{\Phi_{uu}(\omega)} = \frac{-\Phi_{eu}(\omega)}{\Phi_{uu}(\omega)} = \frac{-\Phi_{eu}(\omega)}{K(e^{j\omega})\Phi_{eu}(\omega)} = \frac{-1}{K(e^{j\omega})}$$

We identify the inverse of the controller !

**Solution :** Excite the system at $s$.

1. Identify the TF between $s$ and $y$ : $\dfrac{\hat{G}(e^{j\omega})}{1 + K(e^{j\omega})\hat{G}(e^{j\omega})} = \dfrac{\Phi_{ys}(\omega)}{\Phi_{ss}(\omega)}$
2. Identify the TF between $s$ and $u$ : $\dfrac{1}{1 + K(e^{j\omega})\hat{G}(e^{j\omega})} = \dfrac{\Phi_{us}(\omega)}{\Phi_{ss}(\omega)}$
3. Divide two models : $\hat{G}(e^{j\omega}) = \dfrac{\Phi_{ys}(\omega)}{\Phi_{us}(\omega)}$

# Parametric model from frequency data

**A nonparametric model in frequency domain is available.**

$$G(j\omega_k) \qquad k = 1, \ldots, N$$

**Find a parametric model for the system :**

$$\hat{G}(s) = \frac{B(s)}{A(s)} = \frac{b_m s^m + \ldots + b_1 s + b_0}{a_n s^n + \ldots + a_1 s + 1} \qquad m \leq n$$

**such that** $J = \sum_{k=1}^{N} |\varepsilon(\omega_k)|^2$ **is minimized, where**

$$\varepsilon(\omega_k) = G(j\omega_k) - \frac{\hat{B}(j\omega_k)}{\hat{A}(j\omega_k)}$$

**This is a nonlinear optimization problem in the complex plane.**

**Solution :** Define a linear error with respect to the parameters of A and B : $\varepsilon_\ell(\omega_k) = \hat{A}(j\omega_k)G(j\omega_k) - \hat{B}(j\omega_k) \qquad k = 1, \cdots, N$

Consider the following system of linear equations :

$$\begin{aligned} \hat{A}(j\omega_1)G(j\omega_1) - \hat{B}(j\omega_1) &= 0 \\ \hat{A}(j\omega_2)G(j\omega_2) - \hat{B}(j\omega_2) &= 0 \\ &\vdots \\ \hat{A}(j\omega_N)G(j\omega_N) - \hat{B}(j\omega_N) &= 0 \end{aligned}$$

This system of equations can be solved in the least squares sense but the solution, the parameters of $\hat{A}$ and $\hat{B}$, may be complex.

Therefore, the real and imaginary part of error are considered separately :

$$\begin{aligned} R_e\{\hat{A}(j\omega_k)G(j\omega_k)\} - R_e\{\hat{B}(j\omega_k)\} &= 0 \quad \text{for} \quad k = 1, \ldots, N \\ I_m\{\hat{A}(j\omega_k)G(j\omega_k)\} - I_m\{\hat{B}(j\omega_k)\} &= 0 \quad \text{for} \quad k = 1, \ldots, N \end{aligned}$$

# Parametric model from frequency data

## Problem

In the least squares solution, however, the real error is weighted by a frequency dependent weighting function :

$$|\varepsilon_\ell(j\omega_k)| = |[a_n(j\omega_k)^n + \ldots + a_1(j\omega_k) + 1]\varepsilon(j\omega_k)|$$

Note that $|a_n(j\omega_k)^n + \ldots + a_1(j\omega_k) + 1|$ is typically large in high frequencies (why ?).

This problem lead to very good fit in HF and very bad fit in LF.

## Solution

Use an iterative approach. At $i$-th iteration minimize :

$$J(\theta_i) = \sum_{k=1}^{N} \left| \frac{\varepsilon_\ell(\omega_k, \theta_i)}{\hat{A}(\theta_{i-1})} \right|^2$$

where $\hat{A}(\theta_{i-1})$ is computed in the previous iteration.