➋ Modele de identificare

➋.➋ Noțiuni de Statistică şi Prelucrare de Semnal

- Determinarea **gradului de corelare** existent între valorile unui set de date este extrem de importantă pentru aprecierea **predictibilității** acestuia.
- Dacă datele sunt **slab corelate între ele**, predicția valorilor viitoare nu se poate face decît în limitele unei **precizii scăzute**.

Cum se poate caracteriza corelarea dintre date?

Caracterizarea filozofică este fundamentată de următorul **principiu axiomatic**.

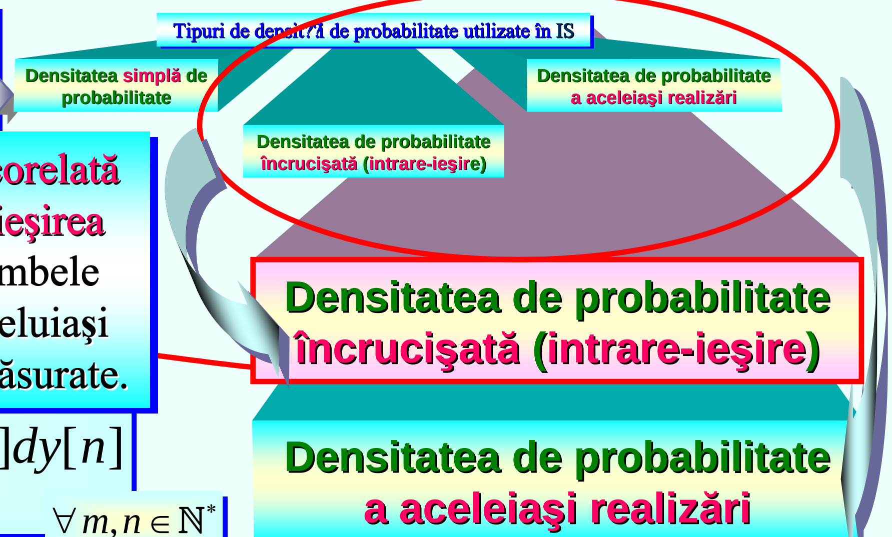
Principiul cauzalității din sistemul logic uman

☝ **Sistemul logic uman este incomplet!**

(Gödel, ~1933)

Prezentul şi trecutul apropiat influențează mai mult viitorul imediat decît o face trecutul îndepărtat.
Viitorul nu poate influența prezentul sau trecutul.

În **IS**, caracterizarea fenomenului de corelare dintre date se bazează pe:



Covarianță (încrucişată)

Arată **cît de corelată statistic este ieşirea cu intrarea**, ambele aparținînd aceluiaşi set de date măsurate.

$$r_{u,y}[m,n] \overset{def}{=} E\{u[m]y[n]\} = \int_{\mathcal{D}(u[m])}\int_{\mathcal{D}(y[n])} \boxed{p\big(u[m],y[n]\big)}\, u[m]y[n]\; du[m]dy[n]$$

pivoți

$\forall m,n \in \mathbb{N}^*$

☝ **Definiție formulată tot cu ajutorul operatorului de mediere statistică.**

❷ Modele de identificare

❷.❷ Noţiuni de Statistică şi Prelucrare de Semnal

În IS, caracterizarea fenomenului de corelare dintre date se bazează pe:

Tipuri de densităţi de probabilitate utilizate în IS

- Densitatea simplă de probabilitate
- Densitatea de probabilitate încrucişată (intrare-ieşire)
- Densitatea de probabilitate a aceleiaşi realizări

Densitatea de probabilitate încrucişată (intrare-ieşire)

Densitatea de probabilitate a aceleiaşi realizări

Auto-covarianţă

Arată cît de corelate statistic sunt ieşirile aparţinînd aceluiaşi set de date măsurate.

$$r_y[m,n] \overset{def}{=} E\{y[m]y[n]\} = \int_{\mathcal{D}(y[m])}\int_{\mathcal{D}(y[n])} \boxed{p\big(y[m],y[n]\big)}\, y[m]y[n]\, dy[m]dy[n]$$

pivoţi

$\forall m,n \in \mathbb{N}^*$

Similar

$$r_u[m,n] \overset{def}{=} E\{u[m]u[n]\}$$

(auto-covarianţa datelor de intrare)

$\forall m,n \in \mathbb{N}^*$

Nu este prea complicat?

Desigur, dacă se încearcă evaluarea plecînd de la definiţii. Simplificarea este posibilă tot datorită unei versiuni a IE.

Ipoteza Ergodică (de covarianţă)

Secvenţa de (auto)-covarianţă a unui proces nu depinde decît de diferenţa pivoţilor (adică procesul este staţionar în covarianţă) şi este egală cu (auto-)covarianţa temporală a oricărei realizări cu un număr infinit de date măsurate.

❷ Modele de identificare

❷.❷ Noțiuni de Statistică și Prelucrare de Semnal

IE

$$r_{u,y}[m,n] = r_{u,y}[m-n] = \lim_{N\to\infty} \frac{1}{\nu_{N,m,n}} \sum_{i=1+\max\{m,n\}}^{N+\min\{m,n\}} u[i-n]y[i-m] \quad \forall\, m,n \in \mathbb{N}^*$$

$$r_{y}[m,n] = r_{y}[m-n] = \lim_{N\to\infty} \frac{1}{\nu_{N,m,n}} \sum_{i=1+\max\{m,n\}}^{N+\min\{m,n\}} y[i-n]y[i-m] \quad \forall\, m,n \in \mathbb{N}^*$$

pivot

$$m - n \overset{def}{=} k \in \mathbb{Z}$$

☝ **Se poate observa cum se calculează diferența pivoților: prin scăderea argumentelor semnalelor implicate.**

$$r_{u,y}[k] = E\{u[n]y[n-k]\} \cong r_{u,y}^{N}[k] = \frac{1}{\nu_{N,k}} \sum_{n=1+\max\{0,k\}}^{N+\min\{0,k\}} u[n]y[n-k] \quad \forall\, k \in \mathbb{Z}$$

$$r_{y}[k] = E\{y[n]y[n-k]\} \cong r_{y}^{N}[k] = \frac{1}{\nu_{N,k}} \sum_{n=1+\max\{0,k\}}^{N+\min\{0,k\}} y[n]y[n-k] \quad \forall\, k \in \mathbb{Z}$$

Numitorul factorului ce precede oricare dintre sume este egal cu numărul de termeni ai sumei.

$$\nu_{N,m,n} \overset{def}{=} N + \min\{m,n\} - \max\{m,n\}$$

$$\nu_{N,k} \overset{def}{=} N + \min\{0,k\} - \max\{0,k\}$$

Aproximații similare

Numitorul factorului ce precede sumele este întotdeauna egal cu dimensiunea orizontului de măsură, N.

✍ **Convenție: semnalele se prelungesc cu zerouri în afara orizontului de măsură.**

(zero padding)

$$r_{u,y}[k] \cong \frac{1}{N} \sum_{n=1}^{N} u[n]y[n-k] \quad \forall\, k \in \mathbb{Z}$$

$$r_{y}[k] \cong \frac{1}{N} \sum_{n=1}^{N} y[n]y[n-k] \quad \forall\, k \in \mathbb{Z}$$

☝ **Cele două tipuri de aproximații sunt sensibil diferite pentru pivoți depărtați de origine.**

➋.➋ Noțiuni de Statistică şi Prelucrare de Semnal

Cît de precise sunt aceste aproximații?

Precizia aproximației creşte odată cu numărul de date achiziționate, dar scade odată cu îndepărtarea pivotului de origine (deoarece numărul de termeni ai sumei devine din ce în ce mai mic).

$$r_{u,y}^{N}[k] = \frac{1}{\nu_{N,k}} \sum_{n=1+\max\{0,k\}}^{N+\min\{0,k\}} u[n]\, y[n-k] \qquad \forall k \in \mathbb{Z}$$

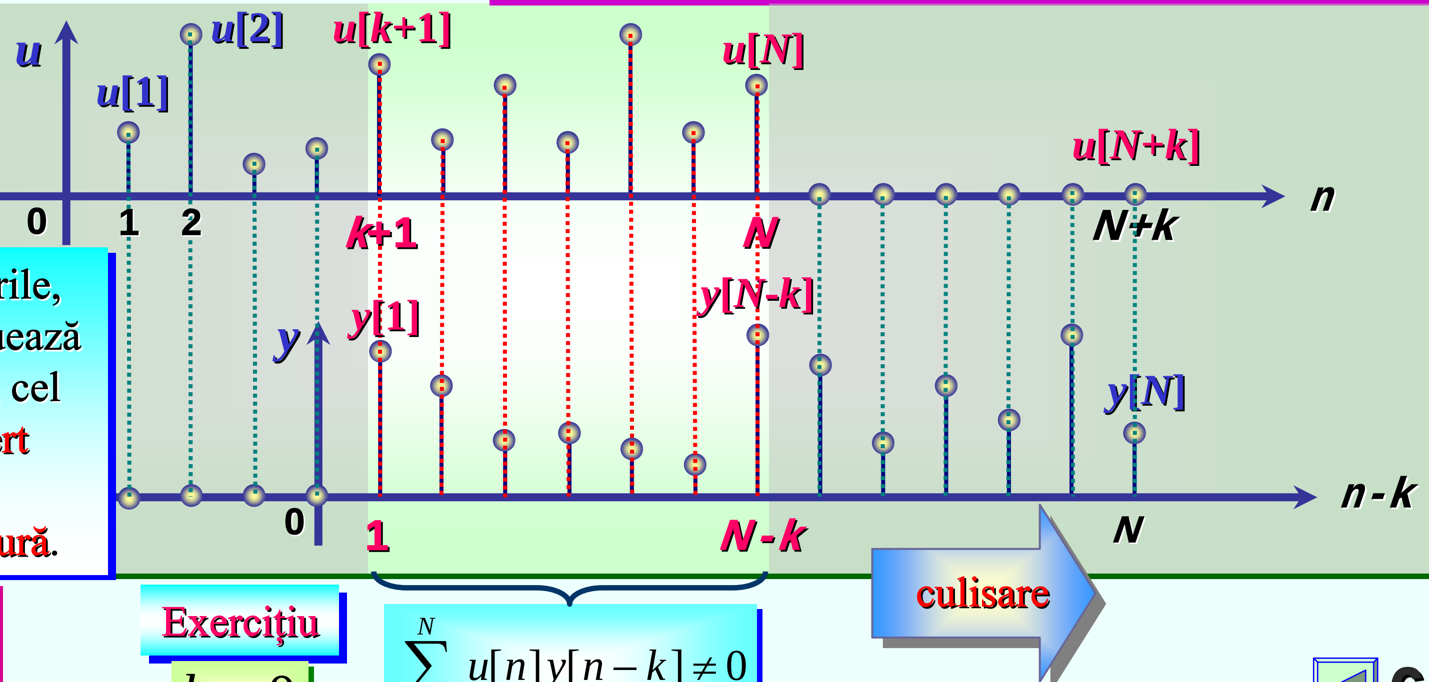
Două cazuri

eşantion întîrziat cu k paşi pentru $k \geq 0$

eşantion anticipat cu $-k$ paşi pentru $k < 0$

☝ Datele culisează de-a lungul orizontului de măsură.

Exemplu $k \geq 0$



$$\sum_{n=k+1}^{N} u[n]\,y[n-k] \neq 0$$

Pentru a limita erorile, în practică se evaluează un număr de valori cel mult egal cu un sfert din dimensiunea orizontului de măsură.

$$k \in \overline{0, \left\lfloor \frac{N}{4} \right\rfloor}$$

Exercițiu $k < 0$

❷ Modele de identificare

❷.❷ Noțiuni de Statistică şi Prelucrare de Semnal

IE ➔ **Proprietăţi elementare ale secvenţei de (auto-)covarianţă** | **Exerciţii**

➔ **Simetrie**

Covarianţă încrucişată

$$r_{u,y}[-k] = r_{y,u}[k] \quad \forall k \in \mathbb{Z}$$

se schimbă cu

Auto-covarianţă

$$r_{y}[-k] = r_{y}[k] \quad \forall k \in \mathbb{Z}$$

☝ Este suficientă evaluarea pentru pivoţi nenegativi.

$$r_y^N[k] = \frac{1}{N-k}\sum_{n=k+1}^{N} y[n]y[n-k] \quad \forall k \in \mathbb{N}$$

Aceste proprietăți se demonstrează cu ajutorul **IE**, completînd seriile de date cu zerouri şi extinzînd sumele la infinit.

➔ **Mărginirea auto-covarianţei**

$$\left|r_y[k]\right| \le r_y[0] = \sigma_y^2 \quad \forall k \in \mathbb{Z}$$

(valoarea maximă a auto-covarianţei este egală cu dispersia datelor)

Auto-corelație (auto-covarianţa normalizată)

$$\rho_y[k] \stackrel{def}{=} \frac{r_y[k]}{r_y[0]} = \frac{r_y[k]}{\sigma_y^2} \quad \forall k \in \mathbb{Z}$$

$$\left|\rho_y[k]\right| \le 1 \quad \forall k \in \mathbb{Z}$$

Similar

Corelație (încrucişată) (covarianţa (încrucişată) normalizată)

$$\rho_{u,y}[k] \stackrel{def}{=} \frac{r_{u,y}[k]}{\sqrt{r_u[0]}\sqrt{r_y[0]}} = \frac{r_{u,y}[k]}{\sigma_u\sigma_y} \quad \forall k \in \mathbb{Z}$$

❷ Modele de identificare

❷.❷ Noţiuni de Statistică şi Prelucrare de Semnal

IE → Proprietăţi elementare ale secvenţei de (auto-)covarianţă (continuare) | Exerciţii

➔ Periodicitatea auto-covarianţei

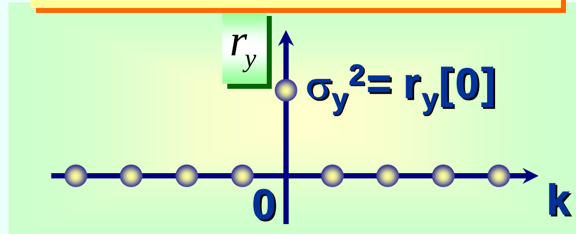
$y[n] = y[n \pm P]$ $\forall n \in \mathbb{Z}$ ⇒ $r_y[k] = r_y[k \pm P]$ $\forall k \in \mathbb{Z}$ (cu aceaşi perioadă ca a datelor)

perioada datelor

impulsul unitar discret (simbolul lui Kronecker)

Proces ne(auto-)corelat ⇒ $r_y[k] = E\{y[n]y[n-k]\} = \sigma_y^2 \boxed{\delta_0[k]}$ $\forall k \in \mathbb{Z}$

O singură valoare nenulă, în origine, egală cu dispersia



☝ Acest tip de proces este complet impredictibil (viitorul nu depinde nici de prezent, nici de trecut).

Proces (statistic) independent ⇒ $p(y[m], y[n]) = p(y[m]) \cdot p(y[n])$ $\forall m,n \in \mathbb{N}^*$

Probabilitatea de apariţie simultană a două valori ale ieşirii pe aceeaşi realizare este egală cu produsul probabilităţilor de apariţie independentă a valorilor pe acea realizare.

Propoziţia 1 | Demonstraţie → Exerciţiu

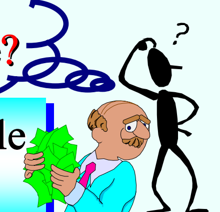
Un proces independent de medie nulă este şi necorelat.

Propoziţia 2 | Demonstraţie → Complicat !

Un proces necorelat avînd distribuţie Gaussiană este şi independent.

Ce legătură există între ele?

Două rezultate remarcabile relevă această legătură.

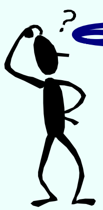


❷ Modele de identificare

❷.❷ Noţiuni de Statistică şi Prelucrare de Semnal

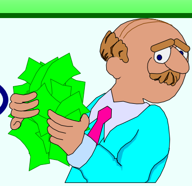
- Datele achiziționate sunt în realitate **semnale** care transportă o anumită **informație** referitoare la comportamentul procesului care le-a produs.

Semnal ⇒ Entitate ce transportă informaţie cu privire la starea sau comportarea unui sistem, atît în **timp** cît şi în **frecvenţă**.



Timp şi frecvenţă?
Nu sunt concepte independente?

- Aparent, **DA**.
- În esență, **NU**.



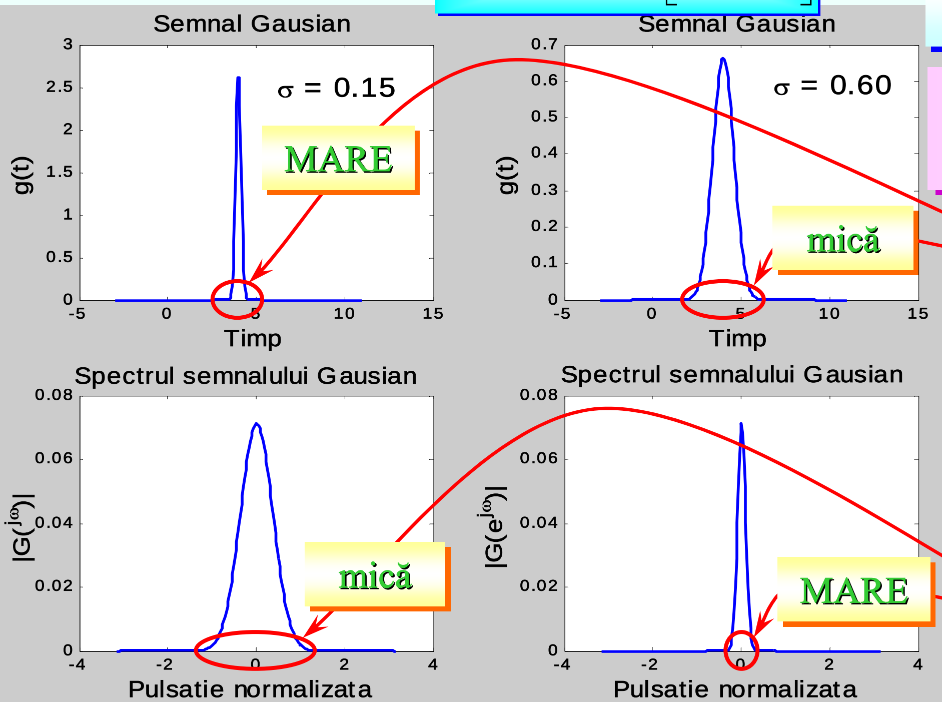
Domeniul "**Timp**" şi domeniul "**Frecvenţă**" sunt **duale**.

- Există mai multe manifestări ale dualității dintre ele.

Exemplu

O "Gaussiană"

$$g(t) \overset{def}{=} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(t-t_0)^2}{2\sigma^2} \right]$$



Principiul de incertitudine GABOR-HEISENBERG ⇒ **Semnalul şi spectrul său nu pot avea simultan suporturi compacte / finite.**

☞ **Ambele pot avea, însă, suporturi infinite.**

Rezoluție în "**Timp**"

Informație temporală **episodică** (energie concentrată pe o durată scurtă).

Produsul rezoluțiilor ≤ o constantă.

Informație frecvențială **persistentă** (energie disipată pe o bandă largă).

Rezoluție în "**Frecvență**"

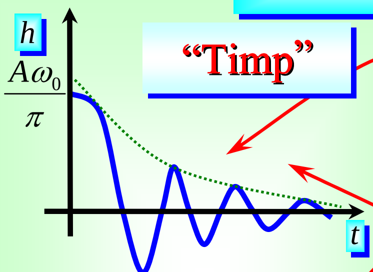
❷ Modele de identificare

❷.❷ Noțiuni de Statistică şi Prelucrare de Semnal

- Informația transportată de semnal poate fi mai uşor de interpretat într-un domeniu decît în altul.
- Compresia semnalului poate fi mai uşor de realizat într-un domeniu decît în altul.

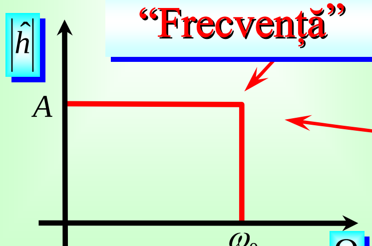
Exemplu

Un FTJ ideal $h(t) \stackrel{def}{=} \dfrac{A\sin\omega_0 t}{\pi t}$

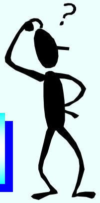


"Timp" — Complicată — Interpretare — Simplă

"Frecvență" — MARE — Rată de compresie — mică



Cum poate fi transferată informația între domeniile Timp şi Frecvență?

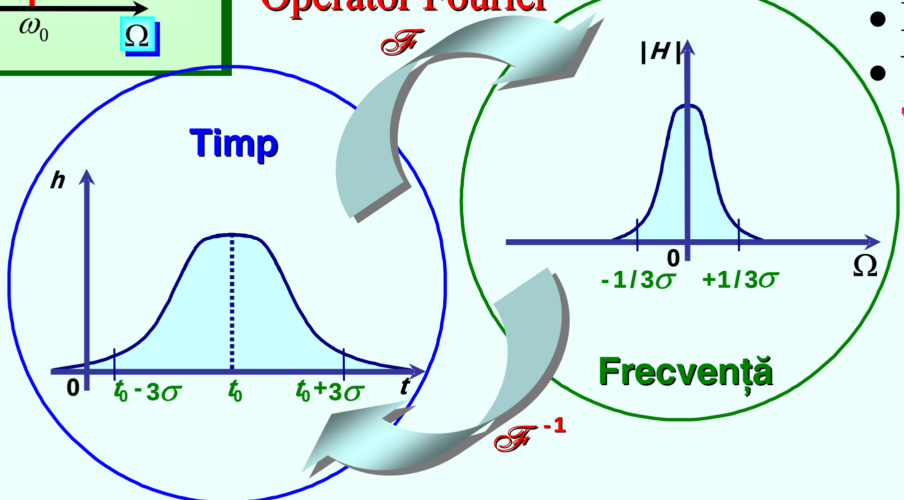


Cu ajutorul transformatelor.

Transformată

Operator inversabil care transformă un semnal temporal în unul frecvențial.

Operator Fourier



- Există mai multe clase de transformate.
- Una dintre cele mai diverse este clasa Transformatelor Fourier (TF).

PS

❷.❷ Noţiuni de Statistică şi Prelucrare de Semnal

Definiţii ale Operatorului Fourier (OF) direct (de analiză) şi invers (de sinteză)

Semnale continue

Analiză

minus

$$\mathscr{F}(f)(j\Omega) \overset{def}{=} F(j\Omega) \overset{def}{=} \int_{-\infty}^{+\infty} f(t)\,\mathrm{e}^{-j\Omega t}\,dt \qquad \forall\,\Omega\in\mathbb{R}$$

Sinteză

plus

$$f(t) = \mathscr{F}^{-1}(F)(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(j\Omega)\,\mathrm{e}^{+j\Omega t}\,d\Omega \qquad \forall\,t\in\mathbb{R}$$

Semnale discrete

Analiză

minus

$$\mathscr{F}(x)(\mathrm{e}^{j\omega}) \overset{def}{=} X(\mathrm{e}^{j\omega}) \overset{def}{=} \sum_{n\in\mathbb{Z}} x[n]\,\mathrm{e}^{-j\omega n} \qquad \forall\,\omega\in\mathbb{R}$$

Sinteză

plus

$$x[n] = \mathscr{F}^{-1}(X)[n] = \frac{1}{2\pi}\int_{-\pi}^{+\pi} X(\mathrm{e}^{j\omega})\,\mathrm{e}^{+j\omega n}\,d\omega \qquad \forall\,n\in\mathbb{Z}$$

De regulă, semnalul frecvenţial returnat de OF este o funcţie cu valori complexe.

Semnale continue

$$F(j\Omega) = \mathrm{Re}\big[F(j\Omega)\big] + j\,\mathrm{Im}\big[F(j\Omega)\big] = \boxed{\big|F(j\Omega)\big|}\,\mathrm{e}^{j\varphi_F(\omega)} \qquad \forall\,\Omega\in\mathbb{R}$$

Spectru **Fază**

Semnale discrete

$$X(\mathrm{e}^{j\omega}) = \mathrm{Re}\big[X(\mathrm{e}^{j\omega})\big] + j\,\mathrm{Im}\big[X(\mathrm{e}^{j\omega})\big] = \boxed{\big|X(\mathrm{e}^{j\omega})\big|}\,\mathrm{e}^{j\varphi_X(\omega)} \qquad \forall\,\omega\in\mathbb{R}$$

Caracterizarea în frecvenţă a sistemelor liniare continue

$$\mathscr{L}(f)(s) \overset{def}{=} F(s) \overset{def}{=} \int_0^{+\infty} f(t)e^{-st}\,dt$$

Transformata Laplace (TL)

← instrumentul de caracterizare a transferului intrare-ieşire pentru sisteme liniare continue

funcţie original (continuală cauzală) oarecare

$f(t)=0, \quad \forall t<0$

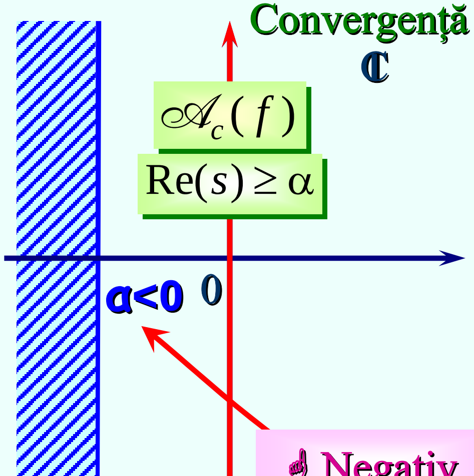
$|f(t)| \le e^{\alpha t}, \quad \forall t > t_0$

Convergenţă

$\mathbb{C}$

$\mathscr{A}_c(f)$

$\mathrm{Re}(s) \ge \alpha$



☝ Negativ.

indicele descreşterii relative normalizate

restricţia TL la axa imaginară (dacă aparţine zonei de convergenţă)

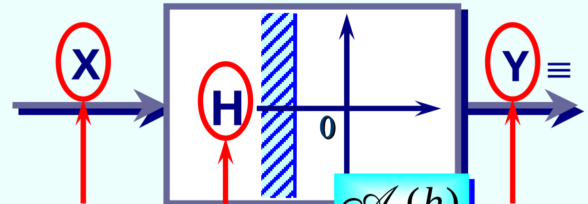
pulsaţie $\in \mathbb{R}$

$$F(s)\big|_{s=j\Omega} = \int_0^{+\infty} f(t)e^{-j\Omega t}\,dt$$

$$\mathscr{F}(f)(j\Omega) \overset{def}{=} F(j\Omega) \overset{def}{=} \int_0^{+\infty} f(t)e^{-j\Omega t}\,dt$$

$\forall \Omega \in \mathbb{R}$

Caracterizare în domeniul complex



TL a intrării

funcţie de transfer

TL a ieşirii

$$H(s) = \frac{Y(s)}{X(s)} = \mathscr{L}(h)(s)$$

☝ Sistem cauzal şi stabil.

$$H(j\Omega) = \frac{Y(j\Omega)}{X(j\Omega)} = \int_0^{+\infty} h(t)e^{-j\Omega t}\,dt$$

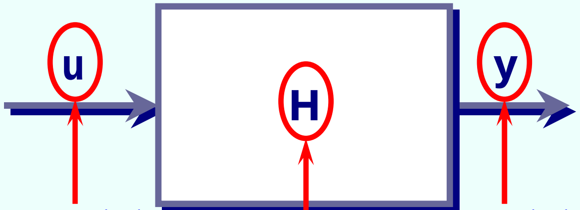
$\forall \Omega \in \mathbb{R}$

Răspuns în frecvenţă

TF a secvenţei pondere

Caracterizarea în frecvenţă a sistemelor liniare discrete

Sistem liniar



semnal de intrare — funcţie de sistem — semnal de ieşire

$$y[n] \stackrel{def}{=} (u*h)[n] = (h*u)[n] = \sum_{k\in\mathbb{Z}} h[k]u[n-k] = \sum_{k\in\mathbb{Z}} h[k]e^{j\omega(n-k)} \quad \forall n\in\mathbb{Z}$$

(convoluţie)

$$y[n] = e^{j\omega n}\boxed{\sum_{k\in\mathbb{Z}} h[k]e^{-j\omega k}} \quad \forall n\in\mathbb{Z}$$

$$e^{j\omega n} \;*\; h \;\rightarrow\; y[n] = e^{j\omega n}\, H(e^{j\omega})$$

$$H(e^{j\omega}) \stackrel{def}{=} \sum_{k\in\mathbb{Z}} h[k]e^{-j\omega k} \stackrel{def}{=} \mathscr{F}(h)(e^{j\omega}) \quad \forall \omega\in\mathbb{R}$$

secvenţă pondere

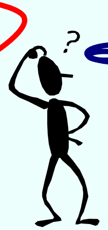
armonică de pulsaţie normalizată $\omega\in\mathbb{R}$

convoluţie

Este răspunsul în frecvenţă corect definit?

Răspuns în frecvenţă

TF a secvenţei pondere



- DA, în cazul sistemelor stabile.
- În acest caz, operaţia de convoluţie este de asemenea corect definită.

Exerciţii

- Deduceţi expresia răspunsului în frecvenţă al unui sistem liniar continuu, respectiv discret, apelînd la un raţionament similar celui utilizat pentru sistemele liniare discrete, respectiv continue.

Transformata Z (TZ):
$$\mathcal{Z}(x)(z) \stackrel{def}{=} X(z) \stackrel{def}{=} \sum_{n\in\mathbb{Z}} x[n]z^{-n}$$

(convergentă pe o coroană circulară)

Convoluţie continuă:
$$(u*h)(t) \stackrel{def}{=} \int_{-\infty}^{+\infty} u(\tau)h(t-\tau)\,d\tau \quad t\in\mathbb{R}$$

Caracterizarea în frecvență a sistemelor liniare discrete (continuare)

$q^{-1} \Longleftrightarrow e^{-j\omega}$

☞ Semnal frecvențial continual.

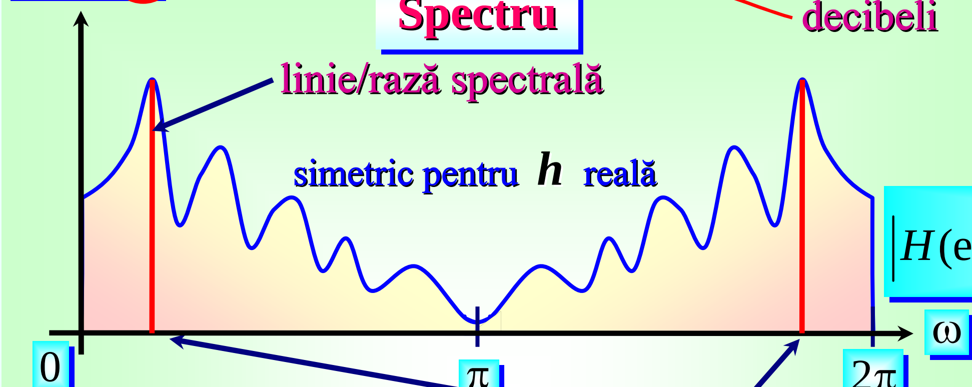
$$H(q^{-1}) \overset{def}{=} \sum_{k\in\mathbb{Z}} h[k]q^{-k}$$

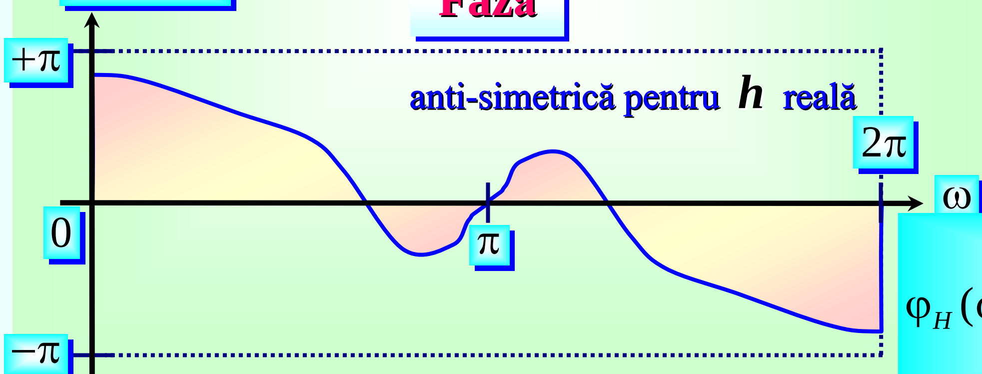
Funcție de sistem

$$H(e^{j\omega}) \overset{def}{=} \sum_{k\in\mathbb{Z}} h[k]e^{-j\omega k} \quad \forall\,\omega\in\mathbb{R}$$

Răspuns în frecvență

$$[a]_{dB} \overset{def}{=} 20\lg a$$





Proprietăți elementare

- Funcție analitică (indefinit derivabilă).
- Periodicitate: $H(e^{j\omega}) = H(e^{j(\omega+2\pi)}) \quad \forall\,\omega\in\mathbb{R}$

$$|H(e^{j\omega})| = \sqrt{\mathrm{Re}^2\left[H(e^{j\omega})\right] + \mathrm{Im}^2\left[H(e^{j\omega})\right]}$$

$$\varphi_H(\omega) = \mathrm{atan2}\left(\frac{\mathrm{Im}\left[H(e^{j\omega})\right]}{\mathrm{Re}\left[H(e^{j\omega})\right]}\right)$$

De ce "decibeli"?

☝ Pentru echilibrarea rezoluției grafice.