

Paolo Rocco

**Dispense ad uso degli studenti del
Politecnico di Milano
per i corsi da cinque crediti didattici**

**Controlli automatici
Ingegneria Aerospaziale**

E' vietato l'uso commerciale di questo materiale

Avvertenza

Questa dispensa è rivolta agli studenti del corso di Controlli Automatici per Ingegneria Aerospaziale, ovvero a studenti che hanno già frequentato un corso di base da cinque crediti didattici di Automatica.

La dispensa ha un taglio volutamente schematico e vuole proporsi come uno strumento agile offerto all'allievo per completare l'apprendimento della materia.

Milano, Gennaio 2004

Sommario

Lezione 1: Complementi di teoria dei sistemi

Lezione 2: Assegnamento degli autovalori

Lezione 3: Luogo delle radici

Lezione 4: Schemi avanzati di controllo

Lezione 5: Sistemi a tempo discreto

Lezione 6: Controllo digitale

Lezione 1

Complementi di teoria dei sistemi

Richiami sui sistemi dinamici

Come è noto, la scrittura dei modelli matematici in termini di “sistemi dinamici” consente di studiarne le proprietà con strumenti di analisi molto generali, che prescindono dalla natura specifica del sistema. Nel seguito riprenderemo rapidamente gli elementi essenziali alla base della nozione di sistema dinamico, già noti dal primo corso di Automatica.

Ricordiamo che un sistema dinamico si interfaccia con il “resto del mondo” per mezzo di una serie di variabili, che definiremo di ingresso, ed altre che definiremo di uscita. Definiamo di ingresso le variabili con cui dall'esterno si influenza il comportamento del sistema, di uscita quelle che caratterizzano il comportamento del sistema e sulle quali soffermiamo il nostro interesse (tipicamente perché costituiscono l'obiettivo del controllo).

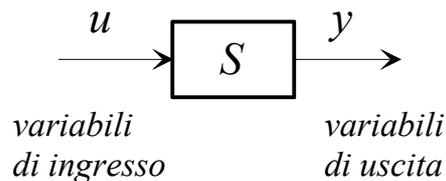


Fig. 1 : Ingressi e uscite di un sistema

E' anche noto che per conoscere il valore delle uscite ad un dato istante t è necessario che siano note un certo numero di condizioni iniziali assegnate ad un istante iniziale t_0 e l'andamento degli ingressi dall'istante iniziale a quello attuale. Il numero minimo di condizioni iniziali che occorre assegnare per determinare tutte le uscite del sistema, noti gli andamenti degli ingressi a partire dall'istante iniziale, prende il nome di *ordine* del sistema e si indica con la lettera n .

Lo **stato** del sistema ad un dato istante riassume tutta la storia passata del sistema fino a quell'istante ed è quindi quanto occorre conoscere per calcolare le uscite da quell'istante in poi, noti gli ingressi. Per quanto affermato sopra, lo stato si può esprimere per mezzo di n variabili, indicate con i simboli x_1, x_2, \dots, x_n , che prendono il nome di **variabili di stato**.

La formalizzazione matematica del sistema dinamico passa allora per la scrittura delle equazioni differenziali (*equazioni di stato*) di cui le variabili di stato sono le soluzioni, noti gli ingressi esterni, e del legame (espresso da trasformazioni algebriche) tra le variabili di uscita e quelle di stato e di ingresso (*trasformazioni di uscita*).

Sia m il numero delle variabili di ingresso e p il numero di variabili di uscita. Introdotti i tre vettori:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix},$$

e le due funzioni vettoriali:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \\ f_2(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \end{bmatrix},$$

$$\mathbf{g}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} g_1(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \\ g_2(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \\ \vdots \\ g_p(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m, t) \end{bmatrix}$$

la formulazione vettoriale del sistema dinamico è la seguente:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \\ \mathbf{y}(t) &= \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), t) \end{aligned}$$

Si osservi che la dipendenza esplicita dal tempo delle due funzioni vettoriali rende il sistema genericamente *tempo variante*. In realtà nel presente corso ci limiteremo allo studio di sistemi in cui questa dipendenza esplicita non esista (sistemi *tempo invarianti* o *stazionari*).

Definiremo poi come sistemi *SISO* (Single Input Single Output) i sistemi per cui $m=p=1$, e genericamente *MIMO* (Multiple Input Multiple Output) gli altri. Si dirà *strettamente proprio* un sistema in cui la funzione \mathbf{g} non dipende dall'ingresso \mathbf{u} , genericamente proprio un sistema in cui ciò non accade. Infine si definisce *lineare* un sistema in cui tutte le equazioni di stato e tutte le trasformazioni di uscita siano funzioni lineari delle variabili di stato e delle variabili di ingresso, *non lineari* tutti gli altri.

Ricordiamo che, assegnata una condizione iniziale all'istante t_0 :

$$\mathbf{x}(t_0) = \mathbf{x}_0,$$

e una funzione di ingresso a partire da t_0 :

$$\mathbf{u}(t) = \bar{\mathbf{u}}(t), \quad t \geq t_0,$$

diciamo **movimento dello stato** $\bar{\mathbf{x}}(t)$ la soluzione delle equazioni di stato corredate dalla condizione iniziale assegnata:

$$\begin{aligned} \dot{\bar{\mathbf{x}}}(t) &= \mathbf{f}(\bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), t) \\ \bar{\mathbf{x}}(t_0) &= \mathbf{x}_0 \end{aligned}$$

e movimento dell'uscita la conseguente uscita, ricavabile dalla trasformazione d'uscita:

$$\bar{\mathbf{y}}(t) = \mathbf{g}(\bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), t).$$

Ricordiamo infine che i particolari movimenti costanti nel tempo, associati a ingressi costanti, prendono il nome di **equilibri**. Per i sistemi tempo-invarianti, la ricerca di eventuali stati di equilibrio associati all'ingresso $\bar{\mathbf{u}}$ si conduce annullando le derivate nelle equazioni di stato e ricercando le eventuali soluzioni dell'equazione vettoriale implicita in $\bar{\mathbf{x}}$:

$$\mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}}) = 0.$$

Sistemi LTI: movimento

Si consideri un sistema LTI di equazioni:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)\end{aligned}$$

Siamo interessati a scrivere l'espressione del movimento dello stato e del movimento dell'uscita conseguenti ad una condizione iniziale \mathbf{x}_0 assegnata all'istante t_0 e ad un ingresso definito a partire dall'istante t_0 .

Per questo, occorre preliminarmente definire la nozione di **esponenziale di matrice**. Dati una matrice quadrata \mathbf{A} ed uno scalare τ definiamo esponenziale di matrice la matrice che si ottiene dal seguente sviluppo in serie:

$$e^{\mathbf{A}\tau} = \sum_{k=0}^{\infty} \frac{(\mathbf{A}\tau)^k}{k!} = \mathbf{I} + \mathbf{A}\tau + \frac{\mathbf{A}^2\tau^2}{2!} + \frac{\mathbf{A}^3\tau^3}{3!} + \dots$$

La nozione di esponenziale di matrice generalizza quindi la nozione di esponenziale scalare. Con questa condivide anche la regola di derivazione, dal momento che, definita come derivata rispetto a τ dell'esponenziale la matrice ottenuta derivando tutti gli elementi della matrice esponenziale rispetto a τ , risulta:

$$\frac{d}{d\tau} e^{\mathbf{A}\tau} = \mathbf{0} + \mathbf{A} + 2\frac{\mathbf{A}^2\tau}{2!} + 3\frac{\mathbf{A}^3\tau^2}{3!} + \dots = \mathbf{A} \left[\mathbf{I} + \mathbf{A}\tau + \frac{\mathbf{A}^2\tau^2}{2!} + \frac{\mathbf{A}^3\tau^3}{3!} + \dots \right] = \mathbf{A}e^{\mathbf{A}\tau}.$$

Il calcolo della matrice esponenziale risulta poi particolarmente agevole quando la matrice \mathbf{A} è diagonalizzabile, ossia quando è simile ad una matrice diagonale:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} = \text{diag}\{\lambda_i, i = 1, \dots, n\},$$

dove λ_i sono gli autovalori di \mathbf{A} . Risulta infatti:

$$\begin{aligned}e^{\mathbf{A}\tau} &= e^{\mathbf{T}^{-1}\hat{\mathbf{A}}\mathbf{T}\tau} = \mathbf{I} + \mathbf{T}^{-1}\hat{\mathbf{A}}\mathbf{T}\tau + \mathbf{T}^{-1}\hat{\mathbf{A}}\mathbf{T}\mathbf{T}^{-1}\hat{\mathbf{A}}\mathbf{T}\frac{\tau^2}{2!} + \dots = \mathbf{T}^{-1} \left[\mathbf{I} + \hat{\mathbf{A}}\tau + \hat{\mathbf{A}}^2\frac{\tau^2}{2!} + \dots \right] \mathbf{T} = \\ &= \mathbf{T}^{-1} e^{\hat{\mathbf{A}}\tau} \mathbf{T} = \mathbf{T}^{-1} \text{diag}\{e^{\lambda_i\tau}\} \mathbf{T}\end{aligned}$$

Fatta questa premessa, siamo in grado di scrivere le espressioni dei movimenti di stato e uscita, come segue:

$$\begin{aligned}\mathbf{x}(t) &= e^{\mathbf{A}(t-t_0)} \mathbf{x}_0 + \int_{t_0}^t e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau \\ \mathbf{y}(t) &= \mathbf{C}e^{\mathbf{A}(t-t_0)} \mathbf{x}_0 + \mathbf{C} \int_{t_0}^t e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau + \mathbf{D}\mathbf{u}(t)\end{aligned}$$

Per la verifica della correttezza dell'espressione del movimento dello stato (quello dell'uscita si ottiene per ovvia sostituzione nella trasformazione di uscita), è utile ricordare un risultato

noto dall'analisi matematica (derivazione sotto segno di integrale). Data una funzione $\boldsymbol{\varphi}(\tau, t)$ (scalare o matriciale) di due argomenti scalari, risulta:

$$\frac{d}{dt} \int_{t_0}^t \boldsymbol{\varphi}(\tau, t) d\tau = \int_{t_0}^t \frac{d}{dt} \boldsymbol{\varphi}(\tau, t) d\tau + \boldsymbol{\varphi}(t, t).$$

Assumendo come funzione $\boldsymbol{\varphi}$ l'espressione $e^{A(t-\tau)} \mathbf{B} \mathbf{u}(\tau)$, e ricordando la regola di derivazione dell'esponenziale di matrice, si ottiene:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \frac{d}{dt} \left[e^{A(t-t_0)} \mathbf{x}_0 \right] + \int_{t_0}^t \frac{d}{dt} \left[e^{A(t-\tau)} \mathbf{B} \mathbf{u}(\tau) \right] d\tau + \left[e^{A(t-\tau)} \mathbf{B} \mathbf{u}(\tau) \right]_{\tau=t} = \\ &= A e^{A(t-t_0)} \mathbf{x}_0 + \int_{t_0}^t A e^{A(t-\tau)} \mathbf{B} \mathbf{u}(\tau) d\tau + \mathbf{B} \mathbf{u}(t) = A \left[e^{A(t-t_0)} \mathbf{x}_0 + \int_{t_0}^t e^{A(t-\tau)} \mathbf{B} \mathbf{u}(\tau) d\tau \right] + \mathbf{B} \mathbf{u}(t) = \\ &= A \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) \end{aligned}$$

Inoltre:

$$\mathbf{x}(t_0) = \mathbf{x}_0.$$

Si osservi che, tanto il movimento dello stato, quanto quello dell'uscita, sono esprimibili come la somma di un termine che dipende (linearmente) solo dallo stato iniziale ed un termine che dipende (linearmente) solo dall'ingresso. Il primo termine prende il nome di moto libero, il secondo di moto forzato.

Sistemi LTI: cambiamento di variabili di stato

I sistemi dinamici lineari tempo invarianti (LTI) possono essere sinteticamente rappresentati dalla quaterna di matrici $A(n \times n)$, $B(n \times m)$, $C(p \times n)$ e $D(p \times m)$ che definiscono le equazioni di stato e le trasformazioni di uscita:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)\end{aligned}$$

E' noto che, una volta definite le variabili di ingresso e uscita, con cui il sistema si interfaccia con l'esterno, la scelta delle variabili di stato è arbitraria. In un sistema LTI, il passaggio da una rappresentazione in variabili di stato ad una rappresentazione alternativa si ottiene sempre moltiplicando il vettore di stato originario per una matrice $(n \times n)$ non singolare:

$$\hat{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t), \quad \det(\mathbf{T}) \neq 0.$$

Potremo quindi esprimere anche il vecchio vettore di stato in termini del nuovo attraverso l'inversa della matrice \mathbf{T} :

$$\mathbf{x}(t) = \mathbf{T}^{-1}\hat{\mathbf{x}}(t).$$

Sostituendo questa espressione nelle equazioni che definiscono il sistema dinamico in \mathbf{x} , e premoltiplicando entrambi i membri dell'equazione vettoriale di stato per la matrice \mathbf{T} , si ottiene:

$$\begin{aligned}\dot{\hat{\mathbf{x}}}(t) &= \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t) \\ \mathbf{y}(t) &= \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)\end{aligned}$$

dove:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \quad \hat{\mathbf{B}} = \mathbf{T}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{T}^{-1}, \quad \hat{\mathbf{D}} = \mathbf{D}.$$

Si osservi che la relazione che sussiste tra le matrici \mathbf{A} e $\hat{\mathbf{A}}$ è di similitudine.

Esempio

Si consideri il circuito elettrico RLC di figura:

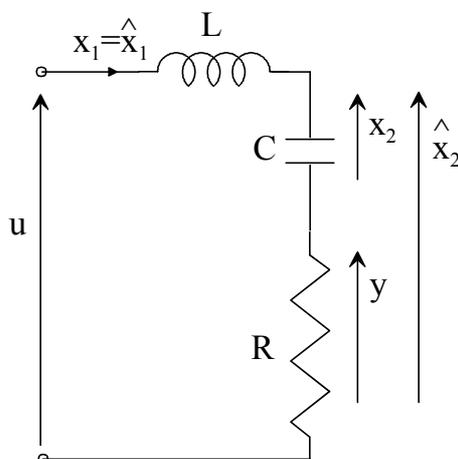


Fig. 2 : Differenti scelte di variabili di stato

Come prima scelta di variabili di stato si prenda la corrente nell'induttore (x_1) e la caduta di tensione ai capi del condensatore (x_2). Assunta come uscita del sistema la caduta di tensione ai capi della resistenza, le equazioni del sistema dinamico si ricavano facilmente:

$$\begin{cases} \dot{x}_1 = -\frac{R}{L}x_1 - \frac{1}{L}x_2 + \frac{1}{L}u \\ \dot{x}_2 = \frac{1}{C}x_1 \\ y = Rx_1 \end{cases} \Rightarrow \mathbf{A} = \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}, \mathbf{C} = [R \ 0], \mathbf{D} = 0$$

Prendendo ora come variabili di stato ancora la corrente nell'induttore (\hat{x}_1) e la caduta di tensione (\hat{x}_2) sulla serie tra resistenza e condensatore, si ottiene:

$$\begin{cases} \dot{\hat{x}}_1 = -\frac{1}{L}\hat{x}_2 + \frac{1}{L}u \\ \dot{\hat{x}}_2 = \frac{1}{C}\hat{x}_1 - \frac{R}{L}\hat{x}_2 + \frac{R}{L}u \\ y = Rx_1 \end{cases} \Rightarrow \hat{\mathbf{A}} = \begin{bmatrix} 0 & -\frac{1}{L} \\ \frac{1}{C} & -\frac{R}{L} \end{bmatrix}, \hat{\mathbf{B}} = \begin{bmatrix} \frac{1}{L} \\ \frac{R}{L} \end{bmatrix}, \hat{\mathbf{C}} = [R \ 0], \hat{\mathbf{D}} = 0.$$

Il legame tra i due insiemi di variabili di stato è evidentemente:

$$\begin{cases} \hat{x}_1 = x_1 \\ \hat{x}_2 = Rx_1 + x_2 \end{cases},$$

ovvero:

$$\hat{\mathbf{x}} = \mathbf{T}\mathbf{x}, \quad \mathbf{T} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix}.$$

Poiché $\det(\mathbf{T}) = 1$, il cambiamento di variabili di stato è ben posto, e si possono verificare le formule di trasformazione:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix} \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -R & 1 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{L} \\ \frac{1}{C} & -\frac{R}{L} \end{bmatrix},$$

$$\hat{\mathbf{B}} = \mathbf{T}\mathbf{B} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{L} \\ \frac{R}{L} \end{bmatrix},$$

$$\hat{\mathbf{C}} = \mathbf{C}\mathbf{T}^{-1} = \begin{bmatrix} R & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -R & 1 \end{bmatrix} = \begin{bmatrix} R & 0 \end{bmatrix},$$

$$\hat{\mathbf{D}} = \mathbf{D} = 0.$$

Concludiamo ricordando che una proprietà di un sistema LTI si dice **strutturale** se è valida qualunque sia la scelta effettuata sulle variabili di stato. In altre parole, descritto il sistema con un certo vettore \mathbf{x} di variabili di stato, ed effettuato il cambiamento di variabili di stato $\hat{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t)$, $\det(\mathbf{T}) \neq 0$, la proprietà verificata con il primo insieme di variabili di stato rimane valida anche per il nuovo insieme, qualsiasi sia la matrice di trasformazione \mathbf{T} .

Stabilità

Nel primo corso di Automatica si è fatto riferimento ad una nozione di stabilità valida per sistemi lineari tempo invarianti, che fa riferimento alla proprietà del sistema di reagire a perturbazioni che intervengono al suo ingresso, ritornando alla condizione preesistente alla perturbazione, o quantomeno non allontanandosene indefinitamente.

Discuteremo ora una nozione più generale di stabilità, che prevede che la perturbazione sulla base della quale si definisce la stabilità intervenga sullo **stato iniziale**. Questa teoria della stabilità (che va sotto il nome di stabilità alla **Lyapunov**, dal nome del matematico che l'ha elaborata) prende le mosse dalla definizione di stabilità di un movimento per un sistema dinamico genericamente non lineare e consente di pervenire all'analisi di stabilità dei sistemi LTI, i cui risultati sono peraltro identici a quelli ricavati con la definizione di stabilità semplificata basata sulla perturbazione impulsiva all'ingresso del sistema.

Prenderemo in considerazione solo sistemi dinamici **tempo invarianti**. Si farà inoltre uso della nozione di **norma** di un vettore che, per i nostri scopi, si potrà sempre definire come la norma Euclidea:

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}^T \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} .$$

Si consideri dunque un sistema dinamico **tempo invariante**, genericamente non lineare, che per $t \geq 0$ sia soggetto all'ingresso $\bar{\mathbf{u}}(t)$.

Definito con \mathbf{x}_{0n} lo stato iniziale (per $t=0$) nominale, sia $\mathbf{x}_n(t)$ il movimento che origina da \mathbf{x}_{0n} con ingresso $\bar{\mathbf{u}}(t)$: definiamo $\mathbf{x}_n(t)$ **movimento nominale**.

Sia ora \mathbf{x}_{0p} un diverso stato iniziale e $\mathbf{x}_p(t)$ il movimento che origina da \mathbf{x}_{0p} con *lo stesso ingresso* $\bar{\mathbf{u}}(t)$: definiamo $\mathbf{x}_p(t)$ **movimento perturbato**.

Ciò posto, il movimento $\mathbf{x}_n(t)$ si dice **stabile** se, $\forall \epsilon > 0$, esiste in corrispondenza un valore $\delta_\epsilon > 0$, tale che $\forall \mathbf{x}_{0p}$ che soddisfi la condizione:

$$\|\mathbf{x}_{0p} - \mathbf{x}_{0n}\| \leq \delta_\epsilon ,$$

risulti:

$$\|\mathbf{x}_p(t) - \mathbf{x}_n(t)\| \leq \epsilon, \quad \forall t \geq 0 .$$

La proprietà di stabilità del movimento formalizza quindi l'idea che perturbando lo stato iniziale del sistema (non l'ingresso) il movimento perturbato non si allontani indefinitamente dal movimento nominale o, più precisamente, che fissato arbitrariamente il massimo scostamento ammissibile tra movimento perturbato e nominale, sia possibile che lo scostamento risulti inferiore a questa soglia, pur di scegliere lo stato iniziale perturbato sufficientemente vicino allo stato iniziale nominale. Nel caso di sistema del primo ordine (nel quale quindi lo stato è rappresentato da una variabile scalare), si può dare un'espressiva rappresentazione grafica di un movimento stabile:

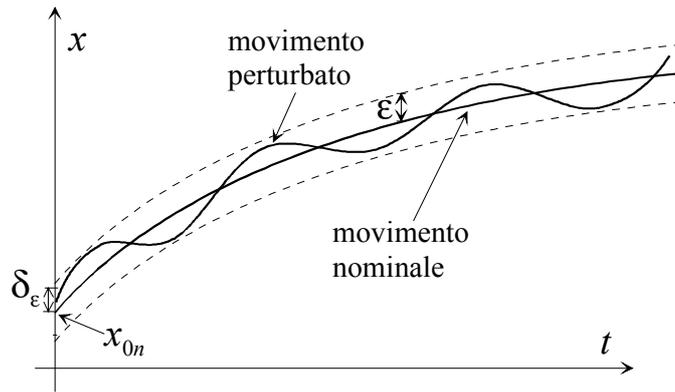


Fig. 3 : Movimento stabile

Un movimento si dirà **instabile** se non è stabile, ossia se non soddisfa la condizione di stabilità precedentemente enunciata. Fissato quindi il massimo scostamento ammissibile tramite il parametro ϵ , se il movimento è instabile non sarà possibile individuare un corrispondente intorno dello stato iniziale nominale, definito dal parametro δ_ϵ , comunque piccolo lo si scelga, in modo che per tutti gli stati iniziali perturbati il movimento perturbato non si scosti da quello nominale più del massimo scostamento ammissibile.

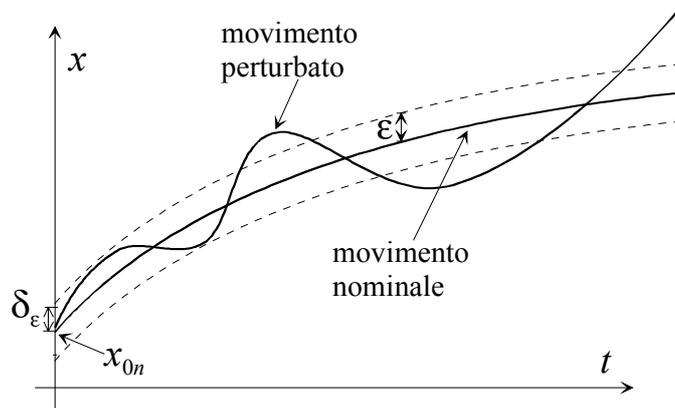


Fig. 4 : Movimento instabile

Infine un movimento $\mathbf{x}_n(t)$ si dice **asintoticamente stabile** se, oltre ad essere stabile, soddisfa la condizione:

$$\lim_{t \rightarrow \infty} \|\mathbf{x}_p(t) - \mathbf{x}_n(t)\| = 0, \quad \forall \mathbf{x}_{0p} : \|\mathbf{x}_{0p} - \mathbf{x}_{0n}\| < \delta_\epsilon.$$

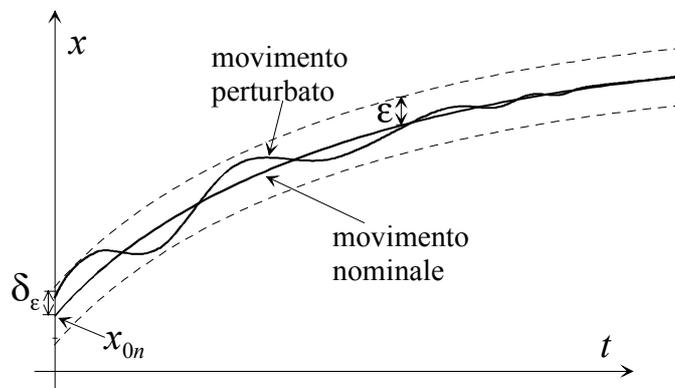


Fig. 5 : Movimento asintoticamente stabile

La proprietà di asintotica stabilità richiede quindi che i movimenti perturbati, oltre a non allontanarsi indefinitamente dal movimento nominale, convergano a questo per $t \rightarrow \infty$.

La definizione di stabilità del movimento si può naturalmente applicare anche a quei particolari movimenti costanti nel tempo che sono gli **equilibri**. Si parlerà quindi di stato di equilibrio stabile, instabile, asintoticamente stabile. Uno stesso sistema dinamico può ammettere in corrispondenza dello stesso ingresso costante diversi stati di equilibrio, con differenti proprietà di stabilità. Si pensi a titolo di esempio ad un pendolo nel piano verticale, soggetto a coppia esterna nulla. Come è noto, il sistema ammette due stati equilibrio fisicamente distinguibili, nel primo dei quali il pendolo è fermo rivolto verso il basso, nel secondo dei quali il pendolo è fermo e rivolto verso l'alto. Dalle definizioni di stabilità precedentemente date, si deduce facilmente che il primo stato di equilibrio è stabile (asintoticamente in presenza di dissipazione), il secondo (pendolo rivolto verso l'alto) è instabile.

Per gli stati di equilibrio asintoticamente stabili si definisce anche la nozione di **regione di attrazione**, intesa come regione nello spazio di stato \mathcal{R}^n in cui può essere scelto lo stato iniziale perturbato perché il movimento perturbato converga allo stato di equilibrio. Se la regione di attrazione coincide con \mathcal{R}^n , ossia se comunque si scelga lo stato iniziale il movimento perturbato converge allo stato di equilibrio, lo stato di equilibrio si dice **globalmente asintoticamente stabile**.

Esempio

Si consideri un sistema dinamico del primo ordine soggetto ad un ingresso costante \bar{u} , e si supponga che il legame tra x e \dot{x} imposto dalla relazione:

$$\dot{x} = f(x, \bar{u})$$

sia rappresentato dalla curva in figura:

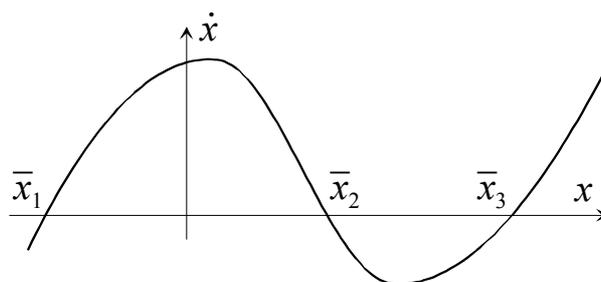


Fig. 6 : Derivata dello stato

Poiché la derivata si annulla per tre valori di x , il sistema ammette tre stati di equilibrio, $\bar{x}_1, \bar{x}_2, \bar{x}_3$.

Non è difficile convincersi che gli stati \bar{x}_1 e \bar{x}_3 sono instabili, in quanto perturbando lo stato iniziale ad un valore di poco superiore al valore di equilibrio, lo stato, avendo derivata positiva, tenderà ad allontanarsi dallo stato di equilibrio stesso, e viceversa per perturbazioni negative. Con ragionamento analogo si conclude che lo stato \bar{x}_2 è invece asintoticamente stabile e la sua regione di attrazione è data dall'insieme (\bar{x}_1, \bar{x}_3) .

Stabilità nei sistemi LTI

La stabilità alla Lyapunov discussa nel precedente paragrafo è una proprietà dei singoli movimenti di un sistema e non una proprietà del sistema stesso. In altre parole, per un generico sistema non lineare, l'affermazione "il sistema è stabile (instabile, asintoticamente stabile)" è priva di significato. Le cose stanno diversamente per i sistemi lineari, come una semplice analisi consente di evidenziare. Consideriamo quindi un sistema LTI di equazioni di stato:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t).$$

Assegnato l'ingresso $\bar{\mathbf{u}}(t)$ per $t \geq 0$, si considerino un movimento nominale:

$$\mathbf{x}_n(t), \quad \mathbf{x}_n(0) = \mathbf{x}_{n0},$$

ed un movimento perturbato:

$$\mathbf{x}_p(t), \quad \mathbf{x}_p(0) = \mathbf{x}_{p0}.$$

Entrambi i movimenti soddisfano le equazioni di stato del sistema dinamico, per cui risulterà:

$$\dot{\mathbf{x}}_n(t) = \mathbf{A}\mathbf{x}_n(t) + \mathbf{B}\bar{\mathbf{u}}(t), \quad \mathbf{x}_n(0) = \mathbf{x}_{n0},$$

$$\dot{\mathbf{x}}_p(t) = \mathbf{A}\mathbf{x}_p(t) + \mathbf{B}\bar{\mathbf{u}}(t), \quad \mathbf{x}_p(0) = \mathbf{x}_{p0}.$$

Sottraendo membro a membro (vettorialmente) le due equazioni, e le relative condizioni iniziali, si ottiene:

$$\dot{\mathbf{x}}_p(t) - \dot{\mathbf{x}}_n(t) = \mathbf{A}(\mathbf{x}_p(t) - \mathbf{x}_n(t)), \quad \mathbf{x}_p(0) - \mathbf{x}_n(0) = \mathbf{x}_{p0} - \mathbf{x}_{n0}.$$

Definiamo ora lo **scostamento** tra il movimento perturbato e quello nominale, e il relativo scostamento tra le condizioni iniziali:

$$\delta\mathbf{x}(t) := \mathbf{x}_p(t) - \mathbf{x}_n(t), \quad \delta\mathbf{x}_0 := \mathbf{x}_{p0} - \mathbf{x}_{n0}.$$

La precedente equazione differenziale potrà quindi essere riscritta in termini della variabile scostamento, come segue:

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t), \quad \delta\mathbf{x}(0) = \delta\mathbf{x}_0.$$

Riprendendo ora la *definizione* di stabilità, e parafrasandola in termini della variabile scostamento, potremo dire che il movimento \mathbf{x}_n è stabile se $\forall \varepsilon > 0$, esiste in corrispondenza un valore $\delta_\varepsilon > 0$, tale che $\forall \delta\mathbf{x}_0$ che soddisfi la condizione:

$$\|\delta\mathbf{x}_0\| \leq \delta_\varepsilon,$$

risulti:

$$\|\delta\mathbf{x}(t)\| \leq \varepsilon, \quad \forall t \geq 0.$$

Il movimento è instabile se non soddisfa questa condizione di stabilità, ed è asintoticamente stabile se, oltre ad essere stabile, soddisfa la condizione:

$$\lim_{t \rightarrow \infty} \|\delta\mathbf{x}(t)\| = 0, \quad \forall \delta\mathbf{x}_0 : \|\delta\mathbf{x}_0\| < \delta_\varepsilon.$$

Appare del tutto evidente che qualunque sia il movimento nominale x_n preso in considerazione, l'analisi della sua stabilità conduce sempre all'analisi del comportamento delle soluzioni dell'equazione differenziale $\delta\dot{x}(t) = A\delta x(t)$, al variare della condizione iniziale. Poiché il risultato di questa analisi è ovviamente lo stesso qualunque sia il movimento x_n di cui si studia la stabilità, potremo concludere che in un sistema dinamico LTI i movimenti sono tutti stabili, o tutti instabili o tutti asintoticamente stabili. In questo senso si può parlare di **stabilità del sistema**: un sistema dinamico LTI sarà stabile se tutti i suoi movimenti sono stabili (analogamente per l'instabilità e l'asintotica stabilità).

Si osservi inoltre che l'equazione differenziale $\delta\dot{x}(t) = A\delta x(t)$, le cui soluzioni al variare dello stato iniziale definiscono la stabilità di tutti i movimenti del sistema dinamico oggetto di studio, è formalmente identica all'equazione di stato del sistema originario, pur di annullare l'ingresso. Le soluzioni di questa equazione sono quindi i moti liberi del sistema, che si generano al variare della condizione iniziale: in altre parole, l'analisi di stabilità di un sistema LTI coinvolge solo il **moto libero** del sistema.

Ricordando che l'espressione del moto libero del sistema (a partire dall'istante iniziale $t=0$ con condizione iniziale δx_0) è la seguente:

$$\delta x_l(t) = e^{At} \delta x_0,$$

si deduce che la stabilità del sistema si analizzerà studiando le proprietà della matrice A .

D'altra parte il fatto che il moto libero si scali proporzionalmente alla condizione iniziale x_0 porta a concludere che la condizione di **stabilità** è verificata **se tutti i moti liberi** (cioè tutti i movimenti che si generano al variare della condizione iniziale) **sono limitati** nel tempo: in questo caso infatti sarà possibile soddisfare la condizione di stabilità precedentemente enunciata, qualunque sia ϵ , pur di scegliere in corrispondenza un valore δ_ϵ sufficientemente piccolo. Se questo non accade, ossia se c'è almeno un moto libero del sistema che non è limitato nel tempo, il sistema è instabile. Infine se tutti i moti liberi del sistema, oltre ad essere limitati, tendono a zero per $t \rightarrow \infty$, il sistema è asintoticamente stabile (in questo caso, infatti, lo scostamento tra movimento perturbato e movimento nominale decade a zero, qualunque sia lo stato iniziale perturbato).

Stabilità ed autovalori

Lo studio della stabilità alla Lyapunov di un sistema LTI ha portato a concludere che occorre studiare le soluzioni dell'equazione¹:

$$\mathbf{x}_l(t) = e^{At} \mathbf{x}_0$$

(moti liberi del sistema) al variare della condizione iniziale \mathbf{x}_0 . Ci si chiede in particolare se tali soluzioni siano limitate nel tempo e, in caso affermativo, se convergano a zero.

Ci muoveremo nell'ipotesi semplificativa che la matrice A sia diagonalizzabile, ossia che esista una matrice non singolare T tale che:

$$\hat{A} = TAT^{-1} = \text{diag}\{\lambda_i, i = 1, \dots, n\},$$

dove λ_i sono gli autovalori di A . Ricordiamo che condizione sufficiente (non necessaria) perché la matrice A sia diagonalizzabile è che gli n autovalori siano distinti, mentre condizione necessaria e sufficiente è che agli n autovalori siano associati n autovettori linearmente indipendenti.

Come già dimostrato, l'espressione del moto libero in termini della matrice diagonale \hat{A} è la seguente:

$$\mathbf{x}_l(t) = e^{At} \mathbf{x}_0 = e^{T^{-1}\hat{A}Tt} \mathbf{x}_0 = T^{-1}e^{\hat{A}t}T\mathbf{x}_0 = T^{-1}\text{diag}\{e^{\lambda_i t}\}T\mathbf{x}_0.$$

Pertanto i moti liberi del sistema sono combinazioni lineari, attraverso gli elementi delle matrici T e T^{-1} e le componenti del vettore \mathbf{x}_0 , degli esponenziali $e^{\lambda_i t}$.

Si ricorda che gli autovalori prendono valore nel campo complesso, e formano un insieme simmetrico rispetto all'asse reale (se esiste un autovalore a parte immaginaria non nulla, esiste anche il complesso coniugato). Dette α_i e β_i rispettivamente la parte reale e immaginaria dell'autovalore λ_i , risulta:

$$e^{\lambda_i t} = e^{\alpha_i t} e^{j\beta_i t} = e^{\alpha_i t} (\cos(\beta_i t) + j \sin(\beta_i t)).$$

Si osservi però che la combinazione lineare attraverso gli elementi delle matrici T e T^{-1} fa naturalmente elidere tutte le parti immaginarie dei numeri complessi, dal momento che l'esponenziale della matrice A non può che avere elementi reali.

Possiamo a questo punto concludere che se tutti gli autovalori di A hanno parte reale negativa ($\alpha_i < 0, \forall i$), tutti gli esponenziali $e^{\lambda_i t}$ sono limitati nel tempo e tendono a zero, monotonicamente se $\beta_i = 0$, oscillando se $\beta_i \neq 0$. Ne consegue che tutti i moti liberi del sistema (qualunque siano le componenti dello stato iniziale) sono limitati nel tempo e si esauriscono per $t \rightarrow \infty$, e quindi che il sistema è asintoticamente stabile.

Se invece tutti gli autovalori di A hanno parte reale negativa o nulla ($\alpha_i \leq 0, \forall i$), e vi è almeno un autovalore a parte reale nulla, tutti gli esponenziali $e^{\lambda_i t}$ sono limitati nel tempo ma quelli associati agli autovalori a parte reale nulla non convergono a zero, rimanendo costanti per $\beta_i = 0$, e oscillando permanentemente se $\beta_i \neq 0$. Essendo sempre possibile individuare

¹ A questo punto, l'indicazione dello scostamento δ non ha ovviamente più alcun rilievo.

almeno una condizione iniziale per cui il moto libero del sistema non converge, in norma, a zero, fermo restando che tutti i moti liberi sono limitati nel tempo, possiamo concludere che il sistema è stabile, non asintoticamente.

Infine se esiste almeno un autovalore di A a parte reale positiva ($\exists j : \alpha_j > 0$), c'è almeno un termine esponenziale $e^{\lambda_i t}$ che diverge (non è limitato nel tempo). Essendo sempre possibile individuare almeno una condizione iniziale per cui il moto libero del sistema non è limitato nel tempo possiamo concludere che il sistema è instabile.

Quanto sopra esposto costituisce dimostrazione del seguente **teorema**.

Un sistema dinamico LTI con matrice A diagonalizzabile è:

asintoticamente stabile: se e solo se tutti gli autovalori di A hanno parte reale negativa;

stabile: se e solo se tutti gli autovalori di A hanno parte reale negativa o nulla e ne esistono a parte reale nulla;

instabile: se e solo se esistono autovalori di A a parte reale positiva.

L'analisi di stabilità di un sistema LTI si conduce quindi studiando la posizione nel piano complesso degli autovalori della matrice A .

Osservazioni

- 1) Ricordiamo che a seguito di un cambiamento di variabili di stato descritto da una matrice di trasformazione T , la matrice A del sistema si trasforma secondo una relazione di similitudine ($\hat{A} = TAT^{-1}$). Poiché matrici simili hanno gli stessi autovalori, l'analisi della stabilità è del tutto indipendente dalla scelta delle variabili di stato. In altre parole, la proprietà di stabilità è una **proprietà strutturale** del sistema dinamico.
- 2) Se la matrice A non è diagonalizzabile, può essere messa in relazione di similitudine con una forma canonica (forma di Jordan). Seguendo questa strada si giunge alla conclusione che il teorema precedentemente enunciato va corredato dalla precisazione che se vi sono autovalori multipli a parte reale nulla (e non vi sono autovalori a parte reale positiva), il sistema è instabile se per almeno uno degli autovalori a parte reale nulla la cosiddetta *molteplicità geometrica* (numero degli autovettori linearmente indipendenti associati all'autovalore) è inferiore alla *molteplicità algebrica* (molteplicità con cui l'autovalore è radice del polinomio caratteristico).

Si consideri a titolo d'esempio il sistema con matrice A :

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Il polinomio caratteristico della matrice è ovviamente $\chi_A(\lambda) = \lambda^2$, per cui l'autovalore $\lambda=0$ ha molteplicità algebrica 2 (è radice doppia del polinomio caratteristico). Per calcolare gli autovettori associati all'autovalore, impostiamo l'equazione:

$$\lambda \mathbf{x} = A\mathbf{x}$$

da cui:

$$\begin{cases} x_2 = 0 \\ 0 = 0 \end{cases}.$$

Pertanto tutti gli autovalori hanno seconda componente nulla, ovvero tutti gli autovettori sono linearmente dipendenti dall'autovettore:

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

La molteplicità geometrica dell'autovalore è quindi pari a 1, inferiore alla molteplicità algebrica: pertanto il sistema è instabile.

A scopo di verifica, calcoliamo l'esponenziale della matrice A :

$$e^{At} = \mathbf{I} + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \dots = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

L'espressione del moto libero è quindi:

$$\mathbf{x}_l(t) = e^{At} \mathbf{x}_0 = \begin{cases} x_{01} + x_{02}t \\ x_{02} \end{cases},$$

il che conferma che è sufficiente scegliere lo stato iniziale con componente x_{02} diversa da zero per generare un moto libero non limitato: il sistema è quindi instabile.

- 3) L'analisi di stabilità condotta sulla base della risposta all'impulso del sistema conduce alle stesse conclusioni della presente analisi. Nello studio della risposta all'impulso si era fatto riferimento ai poli della funzione di trasferimento, i quali d'altra parte coincidono con gli autovalori della matrice A in assenza di cancellazioni di radici tra numeratore e denominatore della funzione di trasferimento. Si osservi che autovalori multipli a parte reale nulla con molteplicità geometrica inferiore alla molteplicità algebrica danno luogo a poli multipli nella funzione di trasferimento che, anche in assenza di poli a parte reale positiva, comportano comunque l'instabilità del sistema.

Stabilità: analisi della matrice A

Nello studio della stabilità di un sistema LTI si può ricorrere, in taluni casi particolari, a criteri che, evitando il calcolo diretto degli autovalori, consentono di concludere circa la stabilità del sistema dinamico. Alcuni criteri basati sull'**ispezione della matrice A** sono i seguenti:

- 1) Se A è triangolare, gli autovalori sono sulla diagonale della matrice, per cui l'analisi di stabilità è immediata (si osservi che in questo caso gli autovalori non possono che essere reali).
- 2) Si ricorda che la traccia di una matrice è la somma degli elementi sulla diagonale principale. Poiché risulta:

$$\text{tr}(A) = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n \text{Re}[\lambda_i],$$

se la traccia di A è positiva o nulla il sistema non è asintoticamente stabile ed è senz'altro instabile se la traccia è positiva (c'è almeno un autovalore a parte reale positiva).

- 3) Poiché:

$$\det(A) = \prod_{i=1}^n \lambda_i,$$

se il determinante della matrice A è nullo il sistema non è asintoticamente stabile (c'è almeno un autovalore nullo).

Criterio di Routh

Il criterio di Routh fornisce una **condizione necessaria e sufficiente** di asintotica stabilità di un sistema dinamico LTI sulla base di un'analisi dei coefficienti del polinomio caratteristico della sua matrice A . In presenza di strumenti di calcolo numerico di diffusione ormai vastissima, in grado di calcolare le radici di polinomi di grado arbitrario, la rilevanza del criterio può apparire modesta, anche in considerazione di una certa laboriosità nella sua applicazione. E' invece indiscutibile il rilievo del criterio nella soluzione di problemi parametrici, ovvero in quei casi in cui i coefficienti del polinomio caratteristico dipendano da uno o più parametri e si vogliano ricavare condizioni di asintotica stabilità del sistema dinamico nello spazio dei parametri.

Sia dunque dato il polinomio caratteristico:

$$\chi_A(\lambda) = \varphi(\lambda) = \varphi_0 \lambda^n + \varphi_1 \lambda^{n-1} + \varphi_2 \lambda^{n-2} + \dots + \varphi_n$$

E' noto che sussiste una **condizione necessaria** perché tutte le radici del polinomio abbiano parte reale negativa, e quindi perché il sistema sia asintoticamente stabile: se il sistema è asintoticamente stabile, tutti i coefficienti $\varphi_0, \varphi_1, \dots, \varphi_n$ del polinomio caratteristico sono concordi (hanno lo stesso segno).

Essendo la condizione solo necessaria, non consente di concludere nulla circa l'asintotica stabilità del sistema quando è verificata.

Per l'applicazione del criterio di Routh, occorre preliminarmente formare una tabella (**tabella di Routh**). La tabella è composta di $n+1$ righe ed è inizializzata con due righe, la prima delle quali è formata dai coefficienti φ_i a partire da φ_0 , presi uno sì e uno no, la seconda dai restanti coefficienti φ_i :

$$\begin{array}{cccc} \varphi_0 & \varphi_2 & \varphi_4 & \dots \\ \varphi_1 & \varphi_3 & \varphi_5 & \dots \\ \vdots & \vdots & \vdots & \\ h_1 & h_2 & h_3 & \dots \\ k_1 & k_2 & k_3 & \dots \\ l_1 & l_2 & l_3 & \dots \\ \vdots & \vdots & \vdots & \end{array}$$

Per formare una generica riga della tabella (la l -sima) ci si serve delle due righe immediatamente precedenti (la k -sima e la h -sima), calcolando ogni elemento della riga secondo la formula:

$$l_i = -\frac{1}{k_1} \det \begin{bmatrix} h_1 & h_{i+1} \\ k_1 & k_{i+1} \end{bmatrix} = h_{i+1} - \frac{h_1 k_{i+1}}{k_1}.$$

Occorre quindi calcolare il determinante della matrice formata dagli elementi in prima colonna delle due righe immediatamente precedenti e dagli elementi di queste due righe nella colonna immediatamente successiva a quella cui appartiene l'elemento che si sta calcolando,

cambiare segno al determinante e dividere il risultato per il primo elemento della riga immediatamente precedente. Si osservi che le righe della tabella possono essere completate con zeri laddove questo sia richiesto per il calcolo delle righe successive.

Se il primo elemento di una colonna risulta nullo, la tabella si dice non definita e il procedimento si arresta, altrimenti prosegue fino alla scrittura della $n+1$ -sima riga.

Il **criterio di Routh** afferma che il sistema è asintoticamente stabile se e solo se la tabella di Routh è ben definita e tutti gli elementi della prima colonna hanno lo stesso segno.

Esempio

Si consideri il polinomio:

$$\varphi(\lambda) = \lambda^3 + 4\lambda^2 + 5\lambda + 2.$$

Essendo tutti i coefficienti del polinomio positivi, la condizione necessaria di stabilità non consente di concludere nulla. Costruiamo la tabella di Routh:

$$1 \quad 5 \quad 0$$

$$4 \quad 2 \quad 0$$

$$k_1 \quad k_2$$

$$h_1$$

I coefficienti della tabella indicati simbolicamente si calcolano come illustrato precedentemente:

$$k_1 = -\frac{1}{4} \det \begin{bmatrix} 1 & 5 \\ 4 & 2 \end{bmatrix} = 4.5, \quad k_2 = -\frac{1}{4} \det \begin{bmatrix} 1 & 0 \\ 4 & 0 \end{bmatrix} = 0$$

$$h_1 = -\frac{1}{k_1} \det \begin{bmatrix} 4 & 2 \\ k_1 & k_2 \end{bmatrix} = 2$$

La tabella è quindi la seguente:

$$1 \quad 5 \quad 0$$

$$4 \quad 2 \quad 0$$

$$4.5 \quad 0$$

$$2$$

Poiché tutti i coefficienti della prima colonna sono positivi, il sistema è asintoticamente stabile.

Esempio

Si consideri il polinomio:

$$\varphi(\lambda) = \lambda^4 + 2\lambda^3 + \lambda^2 + \lambda + a,$$

dove a è un parametro reale. La condizione necessaria richiede che il parametro a sia positivo per l'asintotica stabilità del sistema. Costruiamo la tabella di Routh:

$$\begin{array}{ccc} 1 & 1 & a \\ 2 & 1 & 0 \\ k_1 & k_2 & \\ h_1 & h_2 & \\ l_1 & & \end{array}$$

I coefficienti della tabella indicati simbolicamente si calcolano come illustrato precedentemente:

$$k_1 = -\frac{1}{2} \det \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = 0.5, \quad k_2 = -\frac{1}{2} \det \begin{bmatrix} 1 & a \\ 2 & 0 \end{bmatrix} = a$$

$$h_1 = -\frac{1}{k_1} \det \begin{bmatrix} 2 & 1 \\ k_1 & k_2 \end{bmatrix} = \frac{0.5 - 2a}{0.5} = 1 - 4a, \quad h_2 = -\frac{1}{k_1} \det \begin{bmatrix} 2 & 0 \\ k_1 & 0 \end{bmatrix} = 0.$$

$$l_1 = -\frac{1}{h_1} \det \begin{bmatrix} k_1 & k_2 \\ h_1 & h_2 \end{bmatrix} = a$$

La tabella è quindi la seguente:

$$\begin{array}{ccc} 1 & 1 & a \\ 2 & 1 & 0 \\ 0.5 & a & \\ 1-4a & 0 & \\ a & & \end{array}$$

Affinché tutti i coefficienti della prima colonna siano concordi (positivi) devono essere soddisfatte le due condizioni:

$$\begin{cases} 1 - 4a > 0 \\ a > 0 \end{cases},$$

ossia:

$$0 < a < \frac{1}{4}.$$

Stabilità degli stati di equilibrio

Discutendo della stabilità alla Lyapunov, si è dato significato al concetto di stabilità di un movimento ed in particolare di uno stato di equilibrio. Mentre la verifica di stabilità di un movimento per un sistema non lineare è decisamente complessa, per l'analisi di stabilità di uno stato di equilibrio esistono delle semplici condizioni deducibili dall'analisi del **sistema linearizzato** nell'intorno dello stato di equilibrio.

Si consideri dunque un sistema non lineare di equazioni di stato:

$$\dot{x}(t) = f(x(t), u(t)),$$

ed, in corrispondenza dell'ingresso costante \bar{u} esista lo stato di equilibrio \bar{x} (risulta quindi $f(\bar{x}, \bar{u}) = 0$). Si consideri la matrice A del sistema linearizzato nell'intorno dello stato di equilibrio:

$$A = \left. \frac{\partial f}{\partial x} \right|_{\bar{x}, \bar{u}}$$

Si hanno i seguenti risultati:

- 1) Se la matrice A ha tutti autovalori a parte reale negativa (ossia se il sistema linearizzato è asintoticamente stabile) lo stato di equilibrio \bar{x} è **asintoticamente stabile**.
- 2) Se la matrice A ha almeno un autovalore a parte reale positiva lo stato di equilibrio \bar{x} è **instabile**.

Resta come caso indecidibile sulla base dell'analisi del sistema linearizzato quello in cui la matrice A non ha autovalori a parte reale positiva ma ne ha a parte reale nulla. In questo caso occorrono approssimazioni del sistema originario non lineare estese a termini di ordine superiore al primo per decidere circa la stabilità dello stato di equilibrio.

Esempio

Si riprenda l'esempio del primo ordine in cui il legame tra x e \dot{x} imposto dalla relazione:

$$\dot{x} = f(x, \bar{u})$$

è rappresentato dalla curva in figura 6.

Si è già osservato che il sistema ammette tre stati di equilibrio, $\bar{x}_1, \bar{x}_2, \bar{x}_3$.

Il sistema linearizzato nell'intorno di questi punti di equilibrio assume espressione:

$$\delta\dot{x}(t) = a\delta x(t) + b\delta u(t),$$

con:

$$a = \left. \frac{\partial f(x, u)}{\partial x} \right|_{\bar{x}_i, \bar{u}} = \left. \frac{df(x, \bar{u})}{dx} \right|_{\bar{x}_i},$$

unico autovalore del sistema. Per gli stati \bar{x}_1 e \bar{x}_3 risulta $a > 0$, per cui, in base al secondo dei due risultati precedentemente enunciati, i due stati di equilibrio sono instabili. Per lo stato \bar{x}_2

risulta invece $a < 0$, per cui, sulla base del primo risultato, lo stato di equilibrio è asintoticamente stabile.

Invarianza della funzione di trasferimento

Dato un sistema dinamico LTI descritto dalle equazioni:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$

è noto che una rappresentazione alternativa del sistema si ottiene introducendo i vettori $\mathbf{U}(s)$ e $\mathbf{Y}(s)$, rispettivamente vettori delle trasformate di Laplace degli ingressi e delle uscite del sistema dinamico. Assunto lo stato iniziale del sistema nullo, il legame tra i due vettori è espresso dalla **funzione di trasferimento**:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D},$$

matrice a p righe (quante sono le uscite) e m colonne (quanti sono gli ingressi), funzione della variabile complessa s .

Si supponga ora di effettuare un cambiamento di variabili di stato:

$$\hat{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t), \quad \det(\mathbf{T}) \neq 0.$$

Il sistema nelle nuove variabili di stato è retto dalle equazioni:

$$\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{A}}\hat{\mathbf{x}}(t) + \hat{\mathbf{B}}\mathbf{u}(t)$$

$$\mathbf{y}(t) = \hat{\mathbf{C}}\hat{\mathbf{x}}(t) + \hat{\mathbf{D}}\mathbf{u}(t)$$

dove:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \quad \hat{\mathbf{B}} = \mathbf{T}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{T}^{-1}, \quad \hat{\mathbf{D}} = \mathbf{D}.$$

Calcoliamo la funzione di trasferimento per questo sistema:

$$\begin{aligned} \hat{\mathbf{G}}(s) &= \hat{\mathbf{C}}(s\mathbf{I}_n - \hat{\mathbf{A}})^{-1}\hat{\mathbf{B}} + \hat{\mathbf{D}} = \mathbf{C}\mathbf{T}^{-1}(s\mathbf{T}\mathbf{T}^{-1} - \mathbf{T}\mathbf{A}\mathbf{T}^{-1})^{-1}\mathbf{T}\mathbf{B} + \mathbf{D} = \\ &= \mathbf{C}\mathbf{T}^{-1}[\mathbf{T}(s\mathbf{I}_n - \mathbf{A})\mathbf{T}^{-1}]^{-1}\mathbf{T}\mathbf{B} + \mathbf{D} = \mathbf{C}\mathbf{T}^{-1}\mathbf{T}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{T}\mathbf{B} + \mathbf{D} = \\ &= \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} = \mathbf{G}(s) \end{aligned}$$

Pertanto l'espressione della funzione di trasferimento non muta al variare della rappresentazione di stato, ed è quindi **proprietà strutturale** del sistema dinamico.

Dinamiche nascoste

Facendo riferimento ad un sistema SISO, per il quale la funzione di trasferimento è scalare, è noto che la funzione di trasferimento stessa è una funzione razionale in s , ossia esprimibile come il rapporto di due polinomi:

$$G(s) = \frac{N(s)}{D(s)}.$$

Nel formare l'espressione della funzione di trasferimento a partire dalle matrici del sistema dinamico espresso nel dominio del tempo, possono intervenire **cancellazioni** tra termini a numeratore e denominatore. Il risultato netto è che il denominatore della funzione di trasferimento è un polinomio di grado inferiore all'ordine n del sistema.

Vedremo ora alcuni esempi in cui intervengono cancellazioni e l'interpretazione sistemistica che si può dare a questa circostanza.

Esempio 1

Si consideri il sistema del secondo ordine:

$$\begin{cases} \dot{x}_1 = -2x_1 \\ \dot{x}_2 = x_1 - x_2 + u \\ y = 2x_1 + 3x_2 \end{cases} \Rightarrow \mathbf{A} = \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [2 \quad 3], \quad \mathbf{D} = 0,$$

e calcoliamone la funzione di trasferimento:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} = \dots = \frac{3(s+2)}{(s+2)(s+1)} = \frac{3}{s+1}.$$

Si manifesta quindi una cancellazione della radice -2 . Si osservi che la variabile di stato x_1 non è in alcun modo influenzata dall'ingresso del sistema: essa influenza l'uscita (sia direttamente che indirettamente attraverso la seconda equazione di stato), ma limitatamente al moto libero. Se lo stato iniziale è nullo, la variabile x_1 è identicamente nulla.

Nel sistema è quindi presente una "dinamica nascosta", associata alla prima variabile di stato, che non partecipa al legame ingresso-uscita e rende il sistema dinamico esternamente di ordine inferiore a quanto è in realtà.

Esempio 2

Si consideri il sistema del secondo ordine:

$$\begin{cases} \dot{x}_1 = -x_1 + x_2 + u \\ \dot{x}_2 = -x_2 + u \\ y = x_2 \end{cases} \Rightarrow \mathbf{A} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [0 \quad 1], \quad \mathbf{D} = 0,$$

e calcoliamone la funzione di trasferimento:

$$G(s) = C(sI_n - A)^{-1}B + D = \dots = \frac{s+1}{(s+1)^2} = \frac{1}{s+1}.$$

Si manifesta quindi una cancellazione della radice -1 . Si osservi che la variabile di stato x_1 non influenza l'uscita del sistema, né direttamente, né indirettamente attraverso la seconda variabile di stato. La variabile di stato x_1 è influenzata dall'ingresso, ma la sua evoluzione non ha alcun riscontro nell'uscita y .

Nel sistema è quindi presente una "dinamica nascosta", associata alla prima variabile di stato, che non partecipa al legame ingresso-uscita e rende il sistema dinamico esternamente di ordine 1.

Esempio 3

Si consideri la rete elettrica riportata in figura:

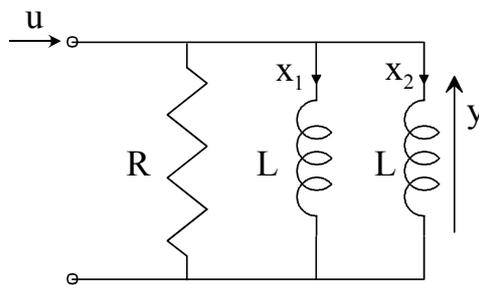


Fig. 7 : Rete elettrica

Dette x_1 e x_2 le correnti nei due induttori, il sistema dinamico è costituito dalle equazioni:

$$\begin{cases} L\dot{x}_1 = R(u - x_1 - x_2) \\ L\dot{x}_2 = R(u - x_1 - x_2) \\ y = R(u - x_1 - x_2) \end{cases} \Rightarrow A = \begin{bmatrix} -\frac{R}{L} & -\frac{R}{L} \\ \frac{R}{L} & \frac{R}{L} \\ -\frac{R}{L} & -\frac{R}{L} \end{bmatrix}, B = \begin{bmatrix} \frac{R}{L} \\ \frac{R}{L} \\ \frac{R}{L} \end{bmatrix}, C = [-R \quad -R], D = R,$$

Per l'analisi del sistema risulta più comodo effettuare il seguente cambiamento di variabili di stato:

$$\begin{cases} \hat{x}_1 = x_1 + x_2 \\ \hat{x}_2 = x_1 - x_2 \end{cases}$$

Sommando e sottraendo le equazioni si ottiene:

$$\begin{cases} L\dot{\hat{x}}_1 = 2R(u - \hat{x}_1) \\ L\dot{\hat{x}}_2 = 0 \\ y = R(u - \hat{x}_1) \end{cases} \Rightarrow \hat{A} = \begin{bmatrix} -2\frac{R}{L} & 0 \\ 0 & 0 \end{bmatrix}, \hat{B} = \begin{bmatrix} 2\frac{R}{L} \\ 0 \end{bmatrix}, \hat{C} = [-R \quad 0], \hat{D} = R$$

e quindi la funzione di trasferimento:

$$G(s) = \hat{G}(s) = \hat{C}(sI_n - \hat{A})^{-1}\hat{B} + \hat{D} = \dots = \frac{RLs}{Ls + 2R}.$$

La variabile di stato \hat{x}_2 non è influenzata dall'ingresso e non influenza l'uscita. Pertanto la dinamica ad essa associata (legata alla differenza tra le correnti negli induttori) non partecipa al legame ingresso-uscita e si configura come una "dinamica nascosta". Agli effetti esterni, la rete elettrica originaria è del tutto equivalente alla rete elettrica semplificata con un solo induttore di induttanza $L/2$ riportata in figura:

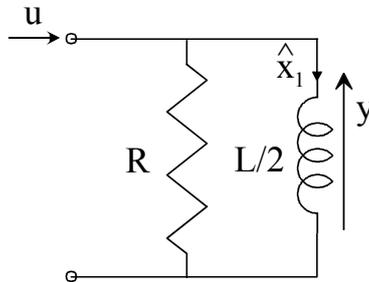


Fig. 8 : Rete elettrica equivalente agli effetti esterni

Raggiungibilità ed osservabilità

La presenza di “dinamiche nascoste” in un sistema, che non si manifestano nel legame ingresso-uscita, può essere formalizzata ricorrendo alla teoria della raggiungibilità e dell’osservabilità, di cui daremo alcuni cenni, anche perché propedeutica alla tecnica di controllo ad assegnamento degli autovalori che verrà trattata nel successivo capitolo.

Diremo che un sistema dinamico lineare tempo invariante è (completamente) **raggiungibile** se è possibile trasferire lo stato del sistema dall’origine a qualsiasi punto in \mathcal{R}^n , con un’opportuna scelta dell’ingresso, in un tempo finito arbitrario.

Un sistema dinamico LTI si dice invece (completamente) **osservabile** se qualunque sia il vettore in \mathcal{R}^n che costituisce lo stato iniziale, l’uscita libera che origina da tale punto differisce, su qualunque intervallo di tempo finito arbitrario, dall’uscita identicamente nulla.

La raggiungibilità è quindi una proprietà che ha a che fare con la possibilità di pilotare il sistema dinamico dall’ingresso in modo da conseguire un’evoluzione arbitraria delle variabili di stato. L’osservabilità concerne invece la possibilità, osservando l’uscita libera del sistema di distinguere lo stato iniziale dall’origine dello spazio di stato: si vedrà in seguito che di fatto la proprietà di osservabilità consente di ricostruire l’andamento delle variabili di stato del sistema a partire da osservazioni sulle sue variabili di uscita.

Se un sistema di ordine n non è completamente raggiungibile, può tuttavia essere scomposto, previo un opportuno cambiamento di variabili di stato, in due parti, una delle quali è completamente raggiungibile, l’altra completamente non raggiungibile (per la quale cioè non è possibile spostare lo stato dall’origine in alcun punto dello spazio di stato agendo sull’ingresso).

Analogamente, un sistema non completamente osservabile può essere scomposto in due parti, una completamente osservabile, l’altra completamente non osservabile (per la quale cioè non è possibile distinguere alcun stato iniziale dall’origine dello spazio di stato sulla base dell’uscita libera).

Se ne deduce che un qualsiasi sistema LTI può essere messo, con un opportuno cambiamento di variabili di stato in una forma canonica, in cui sono poste in evidenza quattro parti, raggiungibili e no, osservabili e no, eventualmente vuote. Il risultato di questo procedimento, che va sotto il nome di **scomposizione canonica di Kalman** e per il quale sono richieste conoscenze di algebra lineare non fornite in questo corso, è rappresentato dallo schema a blocchi in figura, nel quale sono rappresentate le quattro parti², con le mutue influenze e le influenze con ingresso e uscita:

² Le parti non raggiungibili sono indicate con \bar{R} , quelle non osservabili con \bar{O} .

Nel secondo esempio risulta:

$$\mathbf{K}_r = [\mathbf{B} \quad \mathbf{AB}] = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \Rightarrow \det(\mathbf{K}_r) \neq 0$$

$$\mathbf{K}_o = [\mathbf{C}^T \quad \mathbf{A}^T \mathbf{C}^T] = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \Rightarrow \det(\mathbf{K}_o) = 0$$

Il sistema è quindi completamente raggiungibile ma non osservabile. Sarà quindi scomponibile in una parte raggiungibile ed osservabile ed una parte raggiungibile e non osservabile.

Per quanto riguarda il terzo esempio, utilizzando le variabili di stato \hat{x}_1, \hat{x}_2 , risulta:

$$\hat{\mathbf{K}}_r = [\hat{\mathbf{B}} \quad \hat{\mathbf{A}}\hat{\mathbf{B}}] = \begin{bmatrix} 2\frac{R}{L} & -4\frac{R^2}{L^2} \\ 0 & 0 \end{bmatrix} \Rightarrow \det(\hat{\mathbf{K}}_r) = 0$$

$$\hat{\mathbf{K}}_o = [\hat{\mathbf{C}}^T \quad \hat{\mathbf{A}}^T \hat{\mathbf{C}}^T] = \begin{bmatrix} -R & 2\frac{R^2}{L} \\ 0 & 0 \end{bmatrix} \Rightarrow \det(\hat{\mathbf{K}}_o) = 0$$

Il sistema non è quindi né completamente raggiungibile, né completamente osservabile. Sarà quindi scomponibile in una parte raggiungibile ed osservabile ed una parte non raggiungibile e non osservabile.

Traendo spunto anche dal terzo esempio, ci si può chiedere se le proprietà di raggiungibilità e osservabilità dipendano dalla scelta delle variabili di stato.

Si supponga quindi di effettuare un cambiamento di variabili di stato:

$$\hat{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t), \quad \det(\mathbf{T}) \neq 0.$$

Il sistema nelle nuove variabili di stato è caratterizzato dalle matrici:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \quad \hat{\mathbf{B}} = \mathbf{T}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{T}^{-1}, \quad \hat{\mathbf{D}} = \mathbf{D}.$$

La matrice di raggiungibilità per il sistema nelle nuove variabili di stato sarà quindi:

$$\begin{aligned} \hat{\mathbf{K}}_r &= [\hat{\mathbf{B}} \quad \hat{\mathbf{A}}\hat{\mathbf{B}} \quad \hat{\mathbf{A}}^2\hat{\mathbf{B}} \quad \dots \quad \hat{\mathbf{A}}^{n-1}\hat{\mathbf{B}}] = \\ &= [\mathbf{T}\mathbf{B} \quad \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{B} \quad \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{B} \quad \dots \quad (\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\dots\mathbf{T}\mathbf{A}\mathbf{T}^{-1})\mathbf{T}\mathbf{B}] = \\ &= \mathbf{T}[\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}] = \mathbf{T}\mathbf{K}_r \end{aligned}$$

Poiché \mathbf{T} è non singolare, il rango di $\hat{\mathbf{K}}_r$ coincide con il rango di \mathbf{K}_r . Pertanto il sistema nelle nuove variabili di stato è completamente raggiungibile se e solo se lo è quello nelle variabili di stato originarie, ovvero la **raggiungibilità è proprietà strutturale**.

Analogamente, la matrice di osservabilità per il sistema nelle nuove variabili di stato sarà:

$$\begin{aligned}
\hat{\mathbf{K}}_o &= \begin{bmatrix} \hat{\mathbf{C}}^T & \hat{\mathbf{A}}^T \hat{\mathbf{C}}^T & \hat{\mathbf{A}}^{T^2} \hat{\mathbf{C}}^T & \dots & \hat{\mathbf{A}}^{T^{n-1}} \hat{\mathbf{C}}^T \end{bmatrix} = \\
&= \begin{bmatrix} \mathbf{T}^{-T} \mathbf{C}^T & \mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T \mathbf{T}^{-T} \mathbf{C}^T & \mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T \mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T \mathbf{T}^{-T} \mathbf{C}^T & \dots \\ (\mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T \mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T \dots \mathbf{T}^{-T} \mathbf{A}^T \mathbf{T}^T) \mathbf{T}^{-T} \mathbf{C}^T \end{bmatrix} \\
&= \mathbf{T}^{-T} \begin{bmatrix} \mathbf{C}^T & \mathbf{A}^T \mathbf{C}^T & \mathbf{A}^{T^2} \mathbf{C}^T & \dots & \mathbf{A}^{T^{n-1}} \mathbf{C}^T \end{bmatrix} = \mathbf{T}^{-T} \mathbf{K}_o
\end{aligned}$$

Poiché \mathbf{T} (e quindi \mathbf{T}^{-T}) è non singolare, il rango di $\hat{\mathbf{K}}_o$ coincide con il rango di \mathbf{K}_o . Pertanto il sistema nelle nuove variabili di stato è completamente osservabile se e solo se lo è quello nelle variabili di stato originarie, ovvero la **osservabilità è proprietà strutturale**.

Realizzazione

Data una funzione di trasferimento, si definisce **realizzazione** la scrittura di un sistema dinamico nel dominio del tempo (equazioni di stato e trasformazioni di uscita) che ammetta come funzione di trasferimento quella data. Ci limiteremo, nella trattazione del problema, a sistemi SISO (un ingresso ed un'uscita).

E' evidente che il problema di realizzazione ammette infinite soluzioni. Tra tutte le soluzioni, quelle nelle quali la matrice A ha dimensioni uguali al denominatore della funzione di trasferimento prendono il nome di **realizzazioni minime**. Per quanto discusso precedentemente circa la scomposizione canonica di Kalman, si deduce che le realizzazioni minime sono costituite da sistemi dinamici LTI completamente raggiungibili e osservabili.

Assumono poi particolare rilevanza, tra le realizzazioni minime, alcune **forme canoniche**, immediatamente ricavabili dall'espressione della funzione di trasferimento. Queste forme canoniche semplificano, come si vedrà, la soluzione di alcuni problemi rilevanti nello studio dei sistemi nello spazio di stato, quali l'assegnamento degli autovalori e la stima dello stato.

Consideriamo dunque una generica funzione di trasferimento razionale:

$$G(s) = \frac{N(s)}{D(s)},$$

con $D(s)$ di grado n . Se anche il numeratore è di grado n (sistema non strettamente proprio), sarà sempre possibile riscrivere $G(s)$ come:

$$G(s) = \frac{\tilde{N}(s)}{D(s)} + D,$$

dove D è lo scalare che lega $u(t)$ a $y(t)$ nella trasformazione d'uscita, mentre $\tilde{N}(s)$ è di grado inferiore a n . Possiamo quindi concentrarci sulla seguente espressione di G :

$$G(s) = \frac{b_n s^{n-1} + b_{n-1} s^{n-2} + \dots + b_2 s + b_1}{s^n + a_n s^{n-1} + a_{n-1} s^{n-2} + \dots + a_2 s + a_1}.$$

La **forma canonica di raggiungibilità** (anche nota come **forma canonica di controllo**) è la seguente realizzazione minima:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 & -a_2 & -a_3 & \dots & -a_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$\mathbf{C} = [b_1 \quad b_2 \quad b_3 \quad \dots \quad b_n]$$

E' facile verificare, trasformando secondo Laplace le singole equazioni di stato e procedendo per sostituzione, che la forma canonica di controllo è una realizzazione della funzione di trasferimento.

Esempio

Sia:

$$G(s) = \frac{s^2 + 1}{s^2 + 4s + 3}.$$

Scomponendo la frazione, otteniamo:

$$G(s) = \frac{-4s - 2}{s^2 + 4s + 3} + 1.$$

Pertanto la forma canonica di controllo è la seguente:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}, & \mathbf{B} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ \mathbf{C} &= [-2 \quad -4], & \mathbf{D} &= 1. \end{aligned}$$

La forma canonica di controllo è sempre **completamente raggiungibile** per costruzione. Infatti:

$$\mathbf{K}_r = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 & * \\ \vdots & 0 & \dots & \vdots & \vdots \\ 0 & 1 & * & \dots & * \\ 1 & * & * & \dots & * \end{bmatrix},$$

dove con * si indicano termini il cui valore è inessenziale. Il determinante di \mathbf{K}_r vale 1, per cui il sistema è completamente raggiungibile. Per quanto osservato prima, il sistema è anche completamente osservabile, purché la funzione di trasferimento sia ben assegnata (non vi siano radici in comune tra numeratore e denominatore).

Un problema di non immediata soluzione è l'individuazione di un **cambiamento di variabili di stato** che porta un sistema LTI completamente raggiungibile in forma canonica di controllo.

Sia (\mathbf{A}, \mathbf{B}) la coppia di matrici del sistema di partenza, la cui matrice di raggiungibilità è:

$$\mathbf{K}_r = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix},$$

e $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ la coppia di matrici del sistema di arrivo, la cui matrice di raggiungibilità è:

$$\hat{\mathbf{K}}_r = \begin{bmatrix} \hat{\mathbf{B}} & \hat{\mathbf{A}}\hat{\mathbf{B}} & \hat{\mathbf{A}}^2\hat{\mathbf{B}} & \dots & \hat{\mathbf{A}}^{n-1}\hat{\mathbf{B}} \end{bmatrix} = \mathbf{TK}_r.$$

Pertanto la matrice \mathbf{T} di trasformazione è data da:

$$\mathbf{T} = \hat{\mathbf{K}}_r \mathbf{K}_r^{-1}.$$

Un **algoritmo** per ricavare la matrice T può dunque essere il seguente:

Data (A, B) coppia raggiungibile:

- Calcolare $\det(sI - A) = s^n + a_n s^{n-1} + a_{n-1} s^{n-2} + \dots + a_2 s + a_1$
- Ricavare la coppia (\hat{A}, \hat{B}) dai coefficienti a_i
- Calcolare K_r e \hat{K}_r
- Calcolare $T = \hat{K}_r K_r^{-1}$

Una forma canonica utilizzata in alternativa alla forma canonica di controllo è la **forma canonica di osservabilità** (o **forma canonica di ricostruzione**).

Dato un sistema SISO strettamente proprio di matrici (A, B, C) , la sua funzione di trasferimento, scalare, coinciderà ovviamente con la sua trasposta:

$$G(s) = C(sI_n - A)^{-1} B = G^T(s) = [C(sI_n - A)^{-1} B]^T = B^T (sI_n - A)^{-T} C^T = B^T (sI_n - A^T)^{-1} C^T.$$

Posto allora:

$$\tilde{A} = A^T, \quad \tilde{B} = C^T, \quad \tilde{C} = B^T,$$

risulta:

$$G(s) = \tilde{C}(sI_n - \tilde{A})^{-1} \tilde{B},$$

per cui il sistema di matrici $(\tilde{A}, \tilde{B}, \tilde{C})$, che prende il nome di sistema **duale** del sistema dato, costituisce una realizzazione di $G(s)$. La forma canonica di osservabilità è la duale della forma canonica di raggiungibilità:

$$\tilde{A} = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_1 \\ 1 & 0 & \dots & 0 & -a_2 \\ 0 & 1 & \dots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -a_n \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix},$$

$$\tilde{C} = [0 \quad 0 \quad \dots \quad 0 \quad 1].$$

E' facile dimostrare che la realizzazione è completamente osservabile per costruzione e completamente raggiungibile se la funzione di trasferimento è assegnata correttamente.

Esempio

Per l'esempio precedente, la forma canonica di osservabilità è la seguente:

$$\mathbf{A} = \begin{bmatrix} 0 & -3 \\ 1 & -4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -2 \\ -4 \end{bmatrix},$$
$$\mathbf{C} = [0 \quad 1], \quad \mathbf{D} = 1.$$

Lezione 2

Assegnamento degli autovalori

Introduzione

Con la tecnica dell'assegnamento degli autovalori ci proponiamo, dato un sistema lineare tempo invariante, di progettare una legge di controllo in retroazione tale da posizionare gli autovalori del sistema in anello chiuso in punti desiderati del piano complesso. E' noto che gli autovalori di un sistema ne determinano le caratteristiche dei transitori, in particolare dei transitori associati al moto libero. Disporre della possibilità di posizionare gli autovalori del sistema in anello chiuso in punti diversi da quelli occupati dagli autovalori del sistema sotto controllo consente di risolvere due tipologie di problemi, entrambe di primaria importanza:

- **Stabilizzare** un sistema instabile, posizionando gli autovalori in anello chiuso nel semipiano sinistro aperto del piano complesso;
- **Modificare le caratteristiche dei transitori**, anche quelli associati ad autovalori a parte reale negativa, rendendoli più smorzati o rapidi.

Naturalmente i punti del piano complesso prescelti per gli autovalori del sistema in anello chiuso dovranno essere coerenti con il fatto che gli autovalori sono radici di un polinomio a coefficienti reali (il polinomio caratteristico), e quindi dovranno formare un insieme simmetrico rispetto all'asse reale (dovranno quindi essere reali o a coppie complessi e coniugati).

Per la soluzione del problema distinguiamo due scenari:

A) Informazione completa (stato accessibile)

Si assume che il controllore abbia a disposizione le misure di tutte le variabili di stato.

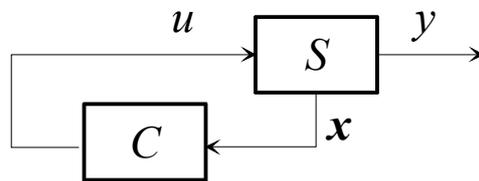


Fig. 1 : Assegnamento degli autovalori con informazione completa

B) Informazione parziale (stato non accessibile)

Si assume che il controllore abbia a disposizione le misure solo delle variabili di uscita, non delle variabili di stato.

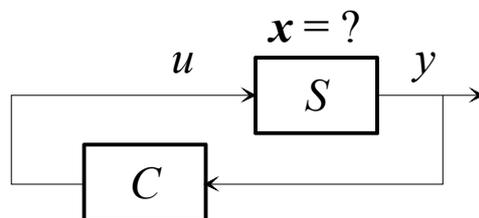


Fig. 2 : Assegnamento degli autovalori con informazione parziale

Supporremo inoltre, salvo esplicita indicazione contraria, che il sistema sotto controllo sia SISO (un'ingresso ed un'uscita).

Assegnamento con stato accessibile

Ipotizziamo che il controllore abbia accesso a tutte le variabili di stato. Per quanto noto sul significato dello stato del sistema, la cui conoscenza costituisce un'informazione completa sull'evoluzione della dinamica del sistema, è ragionevole supporre che il problema possa essere risolto da un **regolatore non dinamico**, ovvero da una legge di controllo puramente proporzionale tra stato e variabile di controllo:

$$u(t) = \mathbf{K}\mathbf{x}(t).$$

Siano dunque:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t)$$

le equazioni di stato del sistema sotto controllo.

Osserviamo che sarà possibile spostare gli autovalori solo della parte raggiungibile del sistema, dal momento che la parte non raggiungibile non è influenzata dall'ingresso.

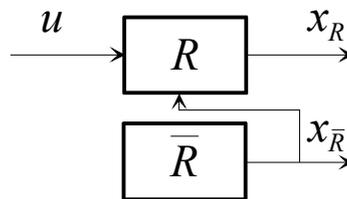


Fig. 3 : Parti raggiungibile e non raggiungibile

Possiamo quindi direttamente assumere la coppia (\mathbf{A}, \mathbf{B}) **raggiungibile**.

A questo punto dobbiamo operare una distinzione a seconda che la coppia (\mathbf{A}, \mathbf{B}) sia in forma canonica di controllo o no.

1) (\mathbf{A}, \mathbf{B}) in forma canonica di controllo

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 & -a_2 & \dots & \dots & -a_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Il polinomio caratteristico di \mathbf{A} è quindi:

$$\chi_{\mathbf{A}}(s) = \det(s\mathbf{I}_n - \mathbf{A}) = s^n + a_n s^{n-1} + \dots + a_2 s + a_1$$

Sia $u(t) = \mathbf{K}\mathbf{x}(t)$ con $\mathbf{K} = [k_1 \quad k_2 \quad \dots \quad k_n]$ la legge di controllo. Le equazioni del sistema in anello chiuso si ottengono sostituendo la legge di controllo nelle equazioni del sistema sotto controllo. Si ottiene immediatamente:

$$\dot{\mathbf{x}}(t) = (\mathbf{A} + \mathbf{B}\mathbf{K})\mathbf{x}(t).$$

Vediamo che struttura assume la matrice $A+BK$:

$$\begin{aligned}
 A+BK &= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1 & -a_2 & \cdots & \cdots & -a_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} [k_1 \quad k_2 \quad \cdots \quad k_{n-1} \quad k_n] = \\
 &= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1+k_1 & -a_2+k_2 & \cdots & \cdots & -a_n+k_n \end{bmatrix}.
 \end{aligned}$$

Poiché la matrice $A+BK$ è in forma canonica di controllo, il suo polinomio caratteristico è immediatamente determinabile:

$$\chi_{A+BK}(s) = \det(sI_n - (A+BK)) = s^n + (a_n - k_n)s^{n-1} + \cdots + (a_2 - k_2)s + (a_1 - k_1).$$

Siano ora λ_i^o , $i=1, \dots, n$ gli autovalori desiderati per il sistema in anello chiuso, radici del **polinomio caratteristico desiderato**:

$$\chi^o(s) = \prod_{i=1}^n (s - \lambda_i^o) = s^n + b_n s^{n-1} + \cdots + b_2 s + b_1.$$

Uguagliando questo polinomio a quello già determinato per la matrice dinamica del sistema in anello chiuso $A+BK$, si ottengono le relazioni:

$$b_i = a_i - k_i, \quad i = 1 \cdots n$$

Pertanto esiste una e una sola legge di controllo $u(t) = Kx(t)$ che risolve il problema dell'assegnamento degli autovalori, i cui coefficienti sono:

$$\boxed{k_i = a_i - b_i, \quad i = 1 \cdots n}$$

2) (A, B) coppia raggiungibile generica

In questo caso occorre effettuare un **cambiamento di variabili di stato** per riportare il sistema in forma canonica di controllo:

$$\hat{x}(t) = T x(t).$$

Le matrici del sistema nelle nuove variabili di stato sono, come è noto, le seguenti:

$$\hat{A} = TAT^{-1}, \quad \hat{B} = TB.$$

La matrice T si ottiene a partire dalle matrici di raggiungibilità prima e dopo il cambiamento di variabili di stato. Infatti la matrice di raggiungibilità originaria è:

$$K_r = [B \quad AB \quad A^2B \quad \cdots],$$

mentre quella del sistema in forma canonica di controllo è:

$$\hat{\mathbf{K}}_r = \begin{bmatrix} \hat{\mathbf{B}} & \hat{\mathbf{A}}\hat{\mathbf{B}} & \hat{\mathbf{A}}^2\hat{\mathbf{B}} & \dots \end{bmatrix} = \begin{bmatrix} \mathbf{T}\mathbf{B} & \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{B} & \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{B} & \dots \end{bmatrix} = \\ = \mathbf{T} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots \end{bmatrix} = \mathbf{T}\mathbf{K}_r$$

Pertanto la matrice del cambiamento di variabili è:

$$\mathbf{T} = \hat{\mathbf{K}}_r \mathbf{K}_r^{-1}$$

A questo punto si risolve il problema dell'assegnamento degli autovalori per il sistema in forma canonica di controllo con la metodologia indicata al punto precedente, trovando:

$$u(t) = \hat{\mathbf{K}}\hat{\mathbf{x}}(t)$$

Tornando alle variabili di stato originarie (che per ipotesi sono anche quelle direttamente misurabili), si ottiene la legge di controllo:

$$u(t) = \hat{\mathbf{K}}\mathbf{T}\mathbf{x}(t) = \mathbf{K}\mathbf{x}(t),$$

con:

$$\boxed{\mathbf{K} = \hat{\mathbf{K}}\mathbf{T}}$$

Conclusion: Se lo stato del sistema è completamente accessibile, esiste **una e una sola** legge di controllo che consente di assegnare arbitrariamente gli n autovalori del sistema in anello chiuso se e solo se il sistema è **completamente raggiungibile**.

Esempio

Dato il sistema:

$$\begin{cases} \dot{x}_1 = -x_1 + 3x_2 + u \\ \dot{x}_2 = 2x_2 + 2u \end{cases} \\ y = x_1 + x_2$$

in cui si suppone lo stato completamente misurabile, si vuole progettare una legge di controllo che sposti l'autovalore positivo nella stessa posizione dell'autovalore negativo del sistema.

Risulta:

$$\mathbf{A} = \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Verifichiamo la raggiungibilità:

$$\mathbf{K}_r = [\mathbf{B} \quad \mathbf{A}\mathbf{B}] = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}, \quad \det(\mathbf{K}_r) = -6 \neq 0$$

Il sistema è quindi completamente raggiungibile. Inoltre gli autovalori di \mathbf{A} sono in -1 e 2 . Il polinomio caratteristico di \mathbf{A} è quindi:

$$\chi_A(s) = (s+1)(s-2) = s^2 - s - 2.$$

Pertanto la forma canonica di controllo del sistema è:

$$\hat{\mathbf{A}} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

cui corrisponde la matrice di raggiungibilità:

$$\hat{\mathbf{K}}_r = [\hat{\mathbf{B}} \quad \hat{\mathbf{A}}\hat{\mathbf{B}}] = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},$$

che è non singolare per costruzione.

La matrice del cambiamento di variabili di stato che porta il sistema in forma canonica di controllo è quindi:

$$\mathbf{T} = \hat{\mathbf{K}}_r \mathbf{K}_r^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{4}{6} & \frac{5}{6} \\ \frac{2}{6} & -\frac{1}{6} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Poiché si vogliono entrambi gli autovalori nel punto -1 , il polinomio caratteristico desiderato è:

$$\chi^o(s) = (s+1)(s+1) = s^2 + 2s + 1.$$

La legge di controllo per il sistema in forma canonica di controllo è quindi costituita dalla matrice con coefficienti:

$$\begin{aligned} \hat{k}_1 &= a_1 - b_1 = -2 - 1 = -3 \\ \hat{k}_2 &= a_2 - b_2 = -1 - 2 = -3 \end{aligned} \Rightarrow \hat{\mathbf{K}} = [-3 \quad -3].$$

Infine la legge di controllo nelle originarie variabili di stato è data da:

$$\mathbf{K} = \hat{\mathbf{K}}\mathbf{T} = [-3 \quad -3] \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{3}{2} \end{bmatrix}.$$

Si può facilmente verificare che gli autovalori della matrice $\mathbf{A} + \mathbf{BK}$ sono effettivamente entrambi nel punto -1 .

Osservazione

Se il sistema ha più di un ingresso, si può comunque procedere all'assegnamento degli autovalori con stato accessibile con argomentazioni simili a quelle fin qui sviluppate.

Nel caso generale infatti la matrice \mathbf{B} ha n righe e m colonne, m essendo il numero delle variabili di ingresso. Può succedere che da uno o più degli m ingressi il sistema sia completamente raggiungibile, ossia che, detta \mathbf{B}_i la i -sima colonna di \mathbf{B} , la coppia $(\mathbf{A}, \mathbf{B}_i)$ risulti raggiungibile per qualche i .

In questo caso è possibile ricondursi al problema monovariabile appena risolto, agendo su uno di questi ingressi, senza utilizzare gli altri, ovvero progettare una matrice riga di guadagni \mathbf{K}_i in modo che gli autovalori della matrice $\mathbf{A} + \mathbf{B}\mathbf{K}_i$ siano in posizioni desiderate del piano complesso.

Se così non è, occorre dapprima assicurarsi che la coppia (\mathbf{A}, \mathbf{B}) sia completamente raggiungibile nel senso dei sistemi multivariabili, ovvero che presa la matrice:

$$\mathbf{K}_r = [\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}],$$

di n righe e $n \times m$ colonne, essa abbia *rango* n , ossia sia possibile estrarne n colonne linearmente indipendenti (tali da formare una matrice quadrata non singolare).

Se questa ipotesi è verificata, è possibile dimostrare che esiste sempre almeno una matrice \mathbf{K} di dimensioni $m \times n$ per cui gli autovalori della matrice $\mathbf{A} + \mathbf{B}\mathbf{K}$ sono in posizioni desiderate del piano complesso.

Si osservi che, contrariamente al caso di sistema a singolo ingresso, per i sistemi a più ingressi esiste una **pluralità di soluzioni** al problema dell'assegnamento degli autovalori. Ne consegue che è possibile stabilire dei criteri (delle cifre di merito) in base ai quali scegliere una soluzione piuttosto che un'altra. Uno dei possibili criteri, utilizzato per esempio dalla funzione `place` del Control Systems Toolbox di Matlab, mira a massimizzare la robustezza del posizionamento degli autovalori, a fronte di incertezze sui parametri del modello.

Stima dello stato

Poniamoci ora nella situazione in cui lo stato del sistema non sia accessibile, ma si disponga solo di misure delle uscite (oltre ovviamente a disporre degli ingressi di controllo). Ci proponiamo di progettare un sistema che, alimentato da ingressi ed uscite del sistema oggetto dello studio, fornisca una stima delle variabili di stato del sistema.

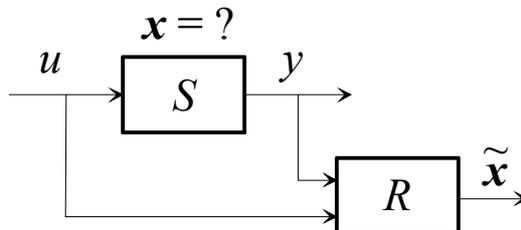


Fig. 4 : Stima dello stato

Chiameremo il sistema che fornisce la stima dello stato **ricostruttore (asintotico)** dello stato o anche, con lieve abuso di terminologia, **osservatore dello stato**.

Supporremo il sistema sotto controllo, oltre che lineare tempo invariante, anche strettamente proprio e SISO, per cui le sue equazioni si potranno scrivere come:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t)$$

$$\mathbf{x}(0) = \mathbf{x}_0$$

$$y(t) = \mathbf{C}\mathbf{x}(t)$$

con u e y variabili scalari.

Supporremo inoltre di conoscere senza incertezza le matrici \mathbf{A} , \mathbf{B} , \mathbf{C} .

Costruiamo ora una **replica** del sistema, cioè un sistema con le stesse equazioni ed alimentato dallo stesso ingresso:

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{B}u(t)$$

$$\tilde{\mathbf{x}}(0) = \tilde{\mathbf{x}}_0$$

$$\tilde{y}(t) = \mathbf{C}\tilde{\mathbf{x}}(t)$$

La replica differisce dal sistema originario solo per lo stato iniziale, che non è noto. Se lo stato iniziale fosse noto senza incertezza, l'uscita vera y e la sua replica \tilde{y} coinciderebbero. In presenza di incertezza sullo stato iniziale si forma un errore tra le due uscite. Appare allora ragionevole **correggere** le equazioni dinamiche della replica del sistema con un termine che pesi la differenza tra \tilde{y} e y :

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{B}u(t) + \mathbf{L}(\tilde{y}(t) - y(t))$$

$$\tilde{\mathbf{x}}(0) = \tilde{\mathbf{x}}_0$$

$$\tilde{y}(t) = \mathbf{C}\tilde{\mathbf{x}}(t)$$

Se, come ipotizzato, il sistema è a singola uscita, \mathbf{L} è dimensionalmente un vettore colonna di n componenti.

Uno schema a blocchi del sistema comprensivo del ricostruttore è riportato di seguito:

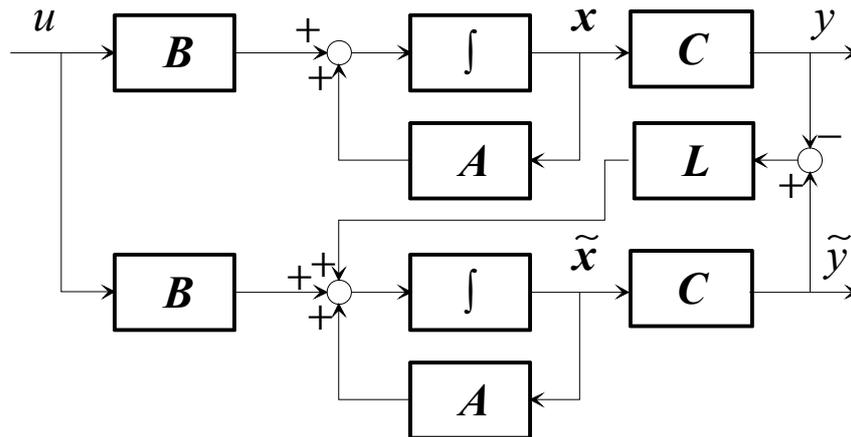


Fig. 5 : Sistema con ricostruttore

Si osservi che non abbiamo introdotto ipotesi di stabilità per il sistema dato. Dobbiamo invece trovare sotto quali condizioni il ricostruttore opera correttamente, ossia produce una stima \tilde{x} dello stato che differisce dallo stato vero x per un errore limitato nel tempo e asintoticamente nullo.

Per questo, riscriviamo le equazioni del sistema e del ricostruttore, sostituendo in queste ultime le trasformazioni di uscita:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$\dot{\tilde{x}}(t) = A\tilde{x}(t) + Bu(t) + LC(\tilde{x}(t) - x(t))$$

Sottraiamo membro a membro le equazioni:

$$\dot{x}(t) - \dot{\tilde{x}}(t) = (A + LC)(x(t) - \tilde{x}(t))$$

Introducendo quindi la variabile:

$$\epsilon(t) = x(t) - \tilde{x}(t),$$

errore nella stima dello stato, si ottiene l'equazione dinamica:

$$\dot{\epsilon}(t) = (A + LC)\epsilon(t).$$

L'errore è quindi governato da un sistema privo di ingresso di matrice dinamica $A+LC$. Se quindi fossimo in grado di scegliere la matrice L in modo tale da posizionare arbitrariamente gli autovalori della matrice $A+LC$, potremmo anzitutto assegnarli nel semipiano sinistro, in modo da rendere la dinamica dell'errore asintoticamente stabile, e anche scegliere arbitrariamente la velocità con cui l'errore di stima tende a zero.

E' già stato dimostrato che, data una coppia di matrici (A, B) raggiungibile, è possibile determinare una matrice K in modo tale che gli autovalori della matrice $A+BK$ siano in punti desiderati del piano complesso.

Se ora indichiamo con la notazione $\lambda_i[X]$ l' i -simo autovalore di una matrice quadrata X , è immediato convincersi che:

$$\lambda_i[A + LC] = \lambda_i[(A + LC)^T] = \lambda_i[A^T + C^T L^T].$$

Ne consegue che, se la coppia (A^T, C^T) è raggiungibile, saremo in grado di risolvere il problema dell'assegnamento degli autovalori, che è uguale a quello già risolto pur di porre $K = L^T$. D'altra parte la coppia (A^T, C^T) è raggiungibile se e solo se la coppia (A, C) è osservabile, come si deduce dal fatto che la matrice di raggiungibilità della prima coppia coincide con la matrice di osservabilità della seconda coppia.

Conclusione: Esiste **una e una sola** matrice di guadagni che consente di assegnare arbitrariamente gli n autovalori della dinamica dell'errore di stima dello stato se e solo se il sistema è **completamente osservabile**.

La procedura per ricavare la matrice dei guadagni L si ottiene facilmente per dualità di quella già illustrata per ricavare la matrice K dell'assegnamento degli autovalori. Si ricorda che, data una terna (A, B, C) che definisce un sistema dinamico strettamente proprio, si definisce **sistema duale** il sistema definito dalla terna (F, G, H) , con $F=A^T, G=C^T, H=B^T$. E' facile verificare che i due sistemi ammettono la stessa funzione di trasferimento.

Scritta quindi la coppia (\hat{F}, \hat{G}) del sistema duale in forma canonica di controllo ¹:

$$\hat{F} = \hat{A}^T = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 & -a_2 & \dots & \dots & -a_n \end{bmatrix}, \quad \hat{G} = \hat{C}^T = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

e posto:

$$\hat{L}^T = [\hat{l}_1 \quad \hat{l}_2 \quad \dots \quad \hat{l}_n],$$

si ricava:

$$\hat{F} + \hat{G}\hat{L}^T = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 + \hat{l}_1 & -a_2 + \hat{l}_2 & \dots & \dots & -a_n + \hat{l}_n \end{bmatrix}.$$

Se il polinomio caratteristico desiderato per la dinamica dell'errore è:

$$\chi^o(s) = s^n + b_n s^{n-1} + \dots + b_2 s + b_1,$$

si ottengono le relazioni:

$$b_i = a_i - \hat{l}_i, \quad \Rightarrow \quad \hat{l}_i = a_i - b_i \quad i = 1 \dots n.$$

Detta quindi K_o la matrice di osservabilità della coppia originaria di matrici (A, C) , coincidente come detto con la matrice di raggiungibilità della coppia (F, G) del sistema

¹ Il polinomio caratteristico di F coincide con il polinomio caratteristico di A .

duale, e $\hat{\mathbf{K}}_o$ la matrice di osservabilità della coppia $(\hat{\mathbf{A}}, \hat{\mathbf{C}})$, a sua volta coincidente con la matrice di raggiungibilità della coppia $(\hat{\mathbf{F}}, \hat{\mathbf{G}})$ del sistema duale in forma canonica di controllo, la matrice del cambio di variabili di stato (che porta il sistema duale in forma canonica di controllo) è:

$$\mathbf{T} = \hat{\mathbf{K}}_o \mathbf{K}_o^{-1}.$$

Pertanto, sempre procedendo in stretta analogia con il problema dell'assegnamento autovalori, si avrà:

$$\mathbf{L}^T = \hat{\mathbf{L}}^T \mathbf{T},$$

e quindi, in definitiva:

$$\boxed{\mathbf{L} = \mathbf{T}^T \hat{\mathbf{L}}}.$$

Esempio

Si riprenda il sistema dell'esempio precedente:

$$\begin{cases} \dot{x}_1 = -x_1 + 3x_2 + u \\ \dot{x}_2 = 2x_2 + 2u \\ y = x_1 + x_2 \end{cases}$$

Si vuole progettare un ricostruttore asintotico dello stato in modo tale che la dinamica dell'errore sia caratterizzata da due autovalori reali coincidenti nel punto -10 .

Risulta:

$$\mathbf{A} = \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 1].$$

Verifichiamo l'osservabilità:

$$\mathbf{K}_o = [\mathbf{C}^T \quad \mathbf{A}^T \mathbf{C}^T] = \begin{bmatrix} 1 & -1 \\ 1 & 5 \end{bmatrix}, \quad \det(\mathbf{K}_o) = 6 \neq 0.$$

Il sistema è quindi completamente osservabile. Il polinomio caratteristico di \mathbf{A} è:

$$\chi_{\mathbf{A}}(s) = (s+1)(s-2) = s^2 - s - 2.$$

Pertanto la forma canonica di controllo per il sistema duale è:

$$\hat{\mathbf{F}} = \hat{\mathbf{A}}^T = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \hat{\mathbf{G}} = \hat{\mathbf{C}}^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

cui corrisponde la matrice di raggiungibilità (di osservabilità per la coppia $(\hat{\mathbf{A}}, \hat{\mathbf{C}})$):

$$\hat{\mathbf{K}}_o = [\hat{\mathbf{G}} \quad \hat{\mathbf{F}}\hat{\mathbf{G}}] = [\hat{\mathbf{C}}^T \quad \hat{\mathbf{A}}^T \hat{\mathbf{C}}^T] = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},$$

che è non singolare per costruzione.

La matrice di cambiamento di variabili di stato che porta il sistema duale in forma canonica di controllo è quindi:

$$T = \hat{K}_o K_o^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} \end{bmatrix} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Poiché si vogliono entrambi gli autovalori nel punto -10 , il polinomio caratteristico desiderato è:

$$\chi^o(s) = (s + 10)(s + 10) = s^2 + 20s + 100.$$

La matrice dei guadagni per il sistema duale in forma canonica di controllo è quindi costituita dai coefficienti:

$$\begin{aligned} \hat{l}_1 &= a_1 - b_1 = -2 - 100 = -102 \\ \hat{l}_2 &= a_2 - b_2 = -1 - 20 = -21 \end{aligned} \Rightarrow \hat{L}^T = [-102 \quad -21].$$

Infine la matrice dei guadagni nelle originarie variabili di stato è data da:

$$L^T = \hat{L}^T T = [-102 \quad -21] \begin{bmatrix} -\frac{1}{6} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = [3 \quad -24],$$

e quindi ovviamente:

$$L = \begin{bmatrix} 3 \\ -24 \end{bmatrix}.$$

Si può facilmente verificare che gli autovalori della matrice $A+LC$ sono effettivamente entrambi nel punto -10 .

Osservazioni

- 1) Come il problema dell'assegnamento degli autovalori, anche il problema della stima dello stato può essere risolto con strumenti analoghi a quelli qui sviluppati anche nel caso di **sistema multivariabile** (con più uscite).
- 2) Si è supposto che il ricostruttore dello stato avesse lo stesso ordine n del sistema sotto controllo, ovvero che tutte le variabili di stato venissero ricostruite. E' evidente che in questa operazione c'è una certa ridondanza, in quanto l'uscita può coincidere essa stessa con una variabile di stato e comunque è sempre uguale ad una combinazione lineare delle variabili di stato. E' allora possibile progettare, con tecniche che non vengono affrontate in questo corso, dei **ricostruttori di ordine ridotto**, che forniscono la stima di un sottoinsieme di variabili di stato.

Assegnamento con stato non accessibile

Una volta progettato il ricostruttore dello stato, ci si chiede se sia possibile risolvere il problema dell'assegnamento degli autovalori con una legge di controllo agente sulla stima dello stato:

$$u(t) = \mathbf{K}\tilde{\mathbf{x}}(t).$$

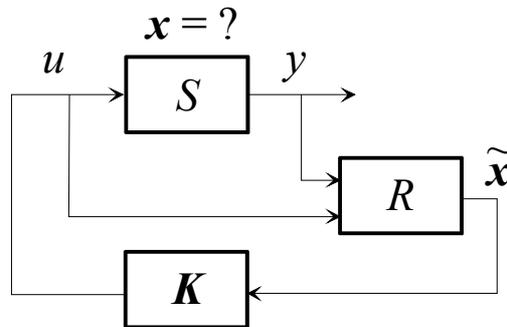


Fig. 6 : Legge di controllo sulla stima dello stato

Ci poniamo nelle ipotesi che la coppia (\mathbf{A}, \mathbf{B}) sia raggiungibile e la coppia (\mathbf{A}, \mathbf{C}) osservabile.

Scriviamo le equazioni del sistema sotto controllo, dell'osservatore e della legge di controllo:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t)$$

$$y(t) = \mathbf{C}\mathbf{x}(t)$$

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{B}u(t) + \mathbf{L}(\tilde{y}(t) - y(t))$$

$$\tilde{y}(t) = \mathbf{C}\tilde{\mathbf{x}}(t)$$

$$u(t) = \mathbf{K}\tilde{\mathbf{x}}(t)$$

Eliminando u , y e \tilde{y} otteniamo:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{K}\tilde{\mathbf{x}}(t)$$

$$\dot{\tilde{\mathbf{x}}}(t) = -\mathbf{L}\mathbf{C}\mathbf{x}(t) + (\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{L}\mathbf{C})\tilde{\mathbf{x}}(t)$$

Nulla vieta, a questo punto, di effettuare un cambiamento di variabili di stato, esprimendo il sistema nelle variabili $\mathbf{x}(t)$ e $\boldsymbol{\varepsilon}(t) = \mathbf{x}(t) - \tilde{\mathbf{x}}(t)$. Sottraendo membro a membro le equazioni si ottiene:

$$\dot{\mathbf{x}}(t) = (\mathbf{A} + \mathbf{B}\mathbf{K})\mathbf{x}(t) - \mathbf{B}\mathbf{K}\boldsymbol{\varepsilon}(t)$$

$$\dot{\boldsymbol{\varepsilon}}(t) = (\mathbf{A} + \mathbf{L}\mathbf{C})\boldsymbol{\varepsilon}(t)$$

La matrice dinamica del sistema in anello chiuso è quindi:

$$\mathbf{A}_c = \begin{bmatrix} \mathbf{A} + \mathbf{B}\mathbf{K} & -\mathbf{B}\mathbf{K} \\ \mathbf{0} & \mathbf{A} + \mathbf{L}\mathbf{C} \end{bmatrix},$$

e risulta triangolare a blocchi.

Ne consegue che gli autovalori della matrice sono la riunione degli autovalori delle due sottomatrici, $\mathbf{A} + \mathbf{B}\mathbf{K}$ e $\mathbf{A} + \mathbf{L}\mathbf{C}$, sulla diagonale. Sappiamo che se la coppia (\mathbf{A}, \mathbf{B}) è raggiungibile, siamo in grado di posizionare arbitrariamente gli autovalori della matrice

$A+BK$ e che, se la coppia (A, C) è osservabile, siamo in grado di posizionare arbitrariamente gli autovalori della matrice $A+LC$.

Conclusione: E' possibile assegnare arbitrariamente gli autovalori del sistema in anello chiuso misurando la sola uscita del sistema se e solo se il sistema è **raggiungibile e osservabile**.

Osservazioni

Di fatto se il sistema è di ordine n si perviene ad un sistema dinamico in anello chiuso di ordine $2n$ i cui autovalori sono posizionabili arbitrariamente. Naturalmente se si usassero ricostruttori di ordine ridotto si pervirebbe ad un sistema in anello chiuso di ordine ridotto.

Si osservi che vige un importante **principio di separazione**: si può progettare la legge di controllo K come se lo stato fosse misurabile e si può progettare il ricostruttore dello stato (matrice L) come se il sistema sotto controllo fosse in anello aperto.

Il fatto che gli autovalori del sistema in anello chiuso possano essere scelti in modo arbitrario è naturalmente da intendersi come un risultato di notevolissima valenza concettuale, ma al quale vanno associate anche delle **limitazioni di ordine pratico** che non emergono dalla trattazione formale dell'argomento. Apparirebbe infatti possibile rendere un sistema in anello chiuso arbitrariamente più veloce del sistema in anello aperto, spostando gli autovalori in posizioni arbitrariamente lontane. E' evidente che ciò comporta conseguenze sulla variabile di controllo, che può risultare sollecitata in modo del tutto incompatibile con i limiti fisici degli attuatori e comunque con criteri di moderazione del controllo essenziali per l'economicità complessiva del progetto. Una metodologia di controllo nello spazio di stato analoga a quella qui trattata ma in grado di tenere esplicitamente conto dello sforzo del controllo è il **controllo ottimo**, trattato in corsi più avanzati.

Proprietà del controllore

Riprendiamo le equazioni del ricostruttore e della legge di controllo:

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{B}u(t) + \mathbf{L}(\tilde{y}(t) - y(t))$$

$$\tilde{y}(t) = \mathbf{C}\tilde{\mathbf{x}}(t)$$

$$u(t) = \mathbf{K}\tilde{\mathbf{x}}(t)$$

Eliminiamo \tilde{y} :

$$\dot{\tilde{\mathbf{x}}}(t) = (\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{L}\mathbf{C})\tilde{\mathbf{x}}(t) - \mathbf{L}y(t)$$

$$u(t) = \mathbf{K}\tilde{\mathbf{x}}(t)$$

Riconosciamo in queste equazioni un sistema dinamico SISO di ordine n , con ingresso y , uscita u e stato $\tilde{\mathbf{x}}$. La matrice dinamica è $\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{L}\mathbf{C}$.

Si osservi che, poiché dalla stabilità delle matrici $\mathbf{A} + \mathbf{B}\mathbf{K}$ e $\mathbf{A} + \mathbf{L}\mathbf{C}$ non si evince la stabilità della matrice $\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{L}\mathbf{C}$, non c'è alcuna garanzia a priori che questo sistema dinamico sia asintoticamente stabile.

Definendo poi le due funzioni di trasferimento, rispettivamente del sistema sotto controllo e del controllore:

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$$

$$R(s) = \mathbf{K}(s\mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{L}\mathbf{C}))^{-1}\mathbf{L}$$

si riconosce che la connessione tra G e R è in retroazione negativa, come la connessione dei regolatori progettati con i metodi classici nel dominio della frequenza.

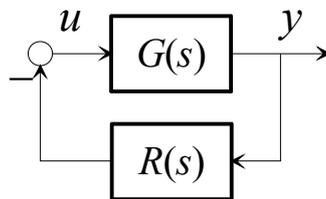


Fig. 7 : Connessione tra sistema sotto controllo e controllore

Ci si chiede a questo punto se non sia possibile progettare la funzione di trasferimento $R(s)$ direttamente nel dominio della frequenza, ossia determinare una funzione di trasferimento razionale (rapporto di polinomi) in modo tale che le radici del polinomio caratteristico in anello chiuso siano in punti desiderati del piano complesso. In effetti questa strada è percorribile (con i metodi cosiddetti polinomiali) ma, oltre a non prestarsi al caso multivariabile, comporta dei problemi di natura numerica (cattivo condizionamento di alcune matrici usate per risolvere il problema).

Esempio

Si riprenda il sistema trattato negli esempi precedenti:

$$\begin{cases} \dot{x}_1 = -x_1 + 3x_2 + u \\ \dot{x}_2 = 2x_2 + 2u \\ y = x_1 + x_2 \end{cases}$$

per il quale:

$$\mathbf{A} = \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 1]$$

Abbiamo determinato le due matrici:

$$\mathbf{K} = \begin{bmatrix} 0 & -\frac{3}{2} \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} 3 \\ -24 \end{bmatrix}.$$

La matrice dinamica del controllore è dunque:

$$\mathbf{A} + \mathbf{BK} + \mathbf{LC} = \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & -\frac{3}{2} \end{bmatrix} + \begin{bmatrix} 3 \\ -24 \end{bmatrix} [1 \quad 1] = \begin{bmatrix} 2 & \frac{9}{2} \\ -24 & -25 \end{bmatrix}.$$

Il polinomio caratteristico di questa matrice risulta:

$$\chi_{\mathbf{A}+\mathbf{BK}+\mathbf{LC}}(s) = s^2 + 23s + 58,$$

per cui il controllore in questo caso risulta asintoticamente stabile. Calcolandone la funzione di trasferimento si ottiene:

$$R(s) = \frac{36s + 36}{s^2 + 23s + 58}.$$

La funzione di trasferimento del sistema sotto controllo è invece:

$$G(s) = \frac{3s + 6}{s^2 - s - 2},$$

da cui si ricava la funzione di trasferimento d'anello:

$$L(s) = R(s)G(s) = \frac{36s + 36}{s^2 + 23s + 58} \frac{3s + 6}{s^2 - s - 2}.$$

Il polinomio caratteristico in anello chiuso (somma di numeratore e denominatore di L) è quindi:

$$\chi(s) = (36s + 36)(3s + 6) + (s^2 + 23s + 58)(s^2 - s - 2) = s^4 + 22s^3 + 141s^2 + 220s + 100,$$

ed ha una radice doppia in -1 ed una doppia in -10 (come previsto).

In figura è riportato il diagramma di Nyquist associato a L , che compie un giro in senso antiorario intorno al punto -1 dell'asse reale, coerentemente con il fatto che L ha un polo nel semipiano destro e con l'enunciato del criterio di Nyquist.

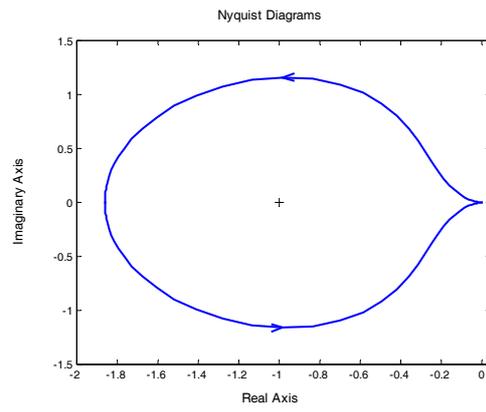


Fig. 8 : Diagramma di Nyquist associato a L

Regolazione a zero dell'errore

Il problema di assegnamento degli autovalori come lo abbiamo affrontato finora non prevede la presenza di un **segnale di riferimento** per l'uscita del sistema. Se viceversa assume rilevanza anche un problema di inseguimento del riferimento ("servo problem" in terminologia inglese) allora è noto che per garantire **precisione statica** in presenza di segnale di riferimento ed eventuale disturbo in linea d'andata costanti a regime è di norma necessario un integratore, ovvero un regolatore di tipo uno.

Uno schema per l'introduzione dell'integratore in un sistema di controllo ad assegnamento degli autovalori con stima dello stato può essere quello rappresentato in figura:

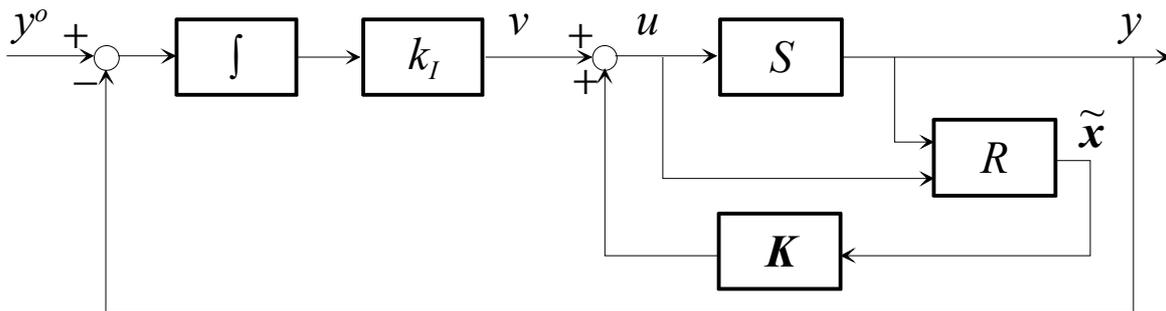


Fig. 9 : Introduzione dell'integratore

Il progetto del guadagno k_I dell'integratore può essere condotto congiuntamente a quello della matrice dei guadagni K in modo tale da allocare gli autovalori del sistema in anello chiuso comprensivo dello stato dell'integratore.

Detto x_I dello stato dell'integratore si ha:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t)$$

$$\dot{x}_I(t) = y^o(t) - y(t) = y^o(t) - \mathbf{C}\mathbf{x}(t)$$

Definiamo:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ x_I \end{bmatrix}$$

lo stato del sistema "aumentato" con l'integratore. Si ha quindi:

$$\dot{\mathbf{z}}(t) = \mathbf{F}\mathbf{z}(t) + \mathbf{G}_u u(t) + \mathbf{G}_{y^o} y^o(t),$$

con:

$$\mathbf{F} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{C} & 0 \end{bmatrix}, \quad \mathbf{G}_u = \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix}, \quad \mathbf{G}_{y^o} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}.$$

Ne consegue che sarà possibile allocare arbitrariamente gli autovalori del sistema aumentato, previo stima dello stato \mathbf{x} , agendo sulla variabile di ingresso u se e solo se la coppia $(\mathbf{F}, \mathbf{G}_u)$ è raggiungibile. Calcoliamone la matrice di raggiungibilità:

$$\begin{aligned} \mathbf{K}_r(F, \mathbf{G}_u) &= \begin{bmatrix} \mathbf{G}_u & F\mathbf{G}_u & F^2\mathbf{G}_u & \dots & F^n\mathbf{G}_u \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^n\mathbf{B} \\ 0 & -\mathbf{C}\mathbf{B} & -\mathbf{C}\mathbf{A}\mathbf{B} & \dots & -\mathbf{C}\mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} 0 & \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} 0 & \mathbf{K}_r \\ 1 & \mathbf{0} \end{bmatrix}, \end{aligned}$$

dove \mathbf{K}_r è la matrice di raggiungibilità del sistema sotto controllo. Essendo il sistema raggiungibile per ipotesi, \mathbf{K}_r è non singolare e non lo è neanche il secondo fattore del prodotto che forma la matrice di raggiungibilità di (F, \mathbf{G}_u) . Per quanto riguarda il primo fattore, osserviamo che, se \mathbf{A} è non singolare:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix},$$

per cui:

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & 0 \end{vmatrix} = \det(\mathbf{A})(\mathbf{C}\mathbf{A}^{-1}\mathbf{B}).$$

Nell'espressione $\mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ riconosciamo, a meno del segno, il **guadagno statico** del sistema, ossia:

$$\mathbf{C}\mathbf{A}^{-1}\mathbf{B} = -G(0),$$

con $G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$ funzione di trasferimento del sistema.

Ne consegue che il determinante è diverso da zero, e quindi il sistema aumentato è raggiungibile, se e solo se il sistema sotto controllo non ha zeri in $s=0$.

Se \mathbf{A} è singolare si può dimostrare, con argomentazioni lievemente più avanzate, che se la coppia (\mathbf{A}, \mathbf{B}) è raggiungibile e la coppia (\mathbf{A}, \mathbf{C}) è osservabile, la matrice:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & 0 \end{bmatrix}$$

è sempre non singolare.

Conclusione: Il problema dell'assegnamento degli autovalori con retroazione dall'uscita e regolazione a zero dell'errore è risolvibile se e solo se il sistema sotto controllo è **raggiungibile e osservabile e non presenta zeri in $s = 0$** .

Elementi in feedforward

Per attribuire anche precisione dinamica al sistema di controllo progettato con l'assegnamento degli autovalori e regolazione a zero dell'errore si utilizzano elementi del sistema di controllo che agiscono solo sul riferimento e non sulla linea di retroazione. Questi elementi prendono il nome di elementi in **feedforward**.

Un possibile schema complessivo degli elementi in feedforward è il seguente:

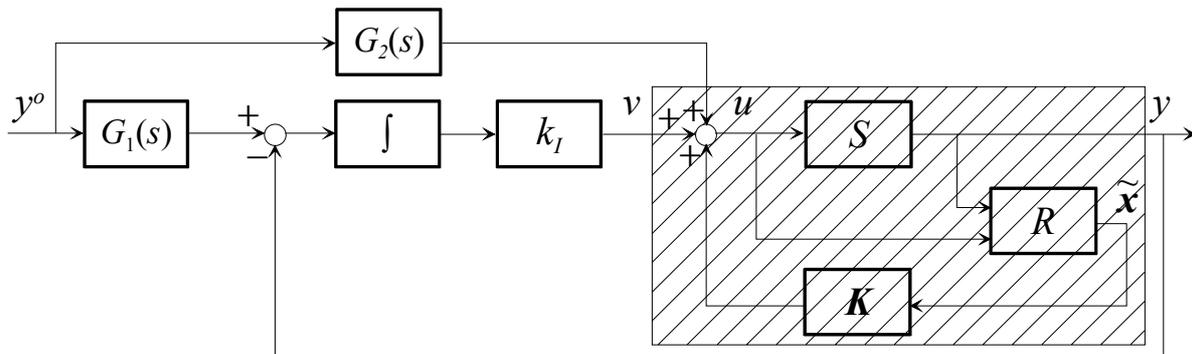


Fig. 10 : Introduzione degli elementi in feedforward

Sia:

$$G(s) = \frac{N(s)}{D(s)}$$

la funzione di trasferimento del sistema sotto controllo (dall'ingresso u all'uscita y). La funzione di trasferimento del sistema tratteggiato, comprensivo del sistema sotto controllo e della retroazione con guadagno K sulla stima dello stato è:

$$\frac{Y(s)}{V(s)} = G_k(s) = \frac{N(s)}{D_k(s)},$$

con:

$$D_k(s) = \det(sI - (A + BK))$$

(è immediato verificare che la dinamica del ricostruttore non compare in questa funzione di trasferimento).

La funzione di trasferimento dal riferimento y^o alla variabile controllata y è:

$$\frac{Y(s)}{Y^o(s)} = \frac{\left[G_1(s) \frac{k_I}{s} + G_2(s) \right] G_k(s)}{1 + \frac{k_I}{s} G_k(s)}.$$

Posto ora:

$$G_2(s) = G_1(s) G_k(s)^{-1}$$

risulta:

$$\frac{Y(s)}{Y^o(s)} = G_1(s).$$

G_1 costituisce quindi la funzione di trasferimento nominale dal riferimento alla variabile controllata: la si progetterà quindi a guadagno unitario e con una banda compatibile con le specifiche di precisione dinamica imposte.

Si osservi tuttavia che il progetto così come descritto non si può condurre se il sistema sotto controllo ha zeri (radici di N) a parte reale positiva, che implicherebbero una G_2 instabile. Inoltre la realizzabilità di G_2 richiede che G_1 abbia grado relativo almeno uguale a quello di G_k e quindi a quello del sistema sotto controllo G .

Lezione 3

Luogo delle radici

Introduzione

Si consideri il sistema retroazionato di Fig. 1, la cui funzione di trasferimento d'anello è $L(s)$.

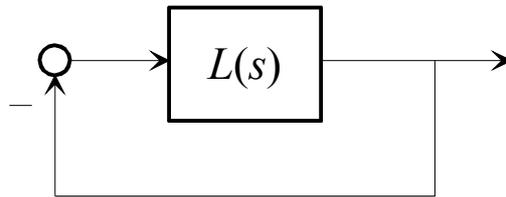


Fig. 1 : Sistema retroazionato

Scritta $L(s)$ nella forma:

$$L(s) = \rho \frac{N(s)}{D(s)} = \rho \frac{\prod_i (s + z_i)}{\prod_k (s + p_k)},$$

l'equazione caratteristica del sistema in anello chiuso assume la seguente espressione:

$$\chi(s) = D(s) + \rho N(s) = 0.$$

Si definisce **luogo delle radici** il luogo descritto nel piano complesso dalle radici di $\chi(s)$ (poli del sistema in anello chiuso) al variare della costante di trasferimento ρ da $-\infty$ a $+\infty$. Per la precisione, si parla di luogo delle radici **diretto** quando ρ varia da 0 a $+\infty$, di luogo delle radici **inverso** quando ρ varia da 0 a $-\infty$.

Esempio

Consideriamo il sistema con funzione di trasferimento d'anello:

$$L(s) = \frac{\rho}{(s+1)(s+2)}$$

Il polinomio caratteristico in anello chiuso è quindi:

$$\chi(s) = (s+1)(s+2) + \rho = s^2 + 3s + 2 + \rho = 0$$

ed ha le radici:

$$s = \frac{-3 \pm \sqrt{1 - 4\rho}}{2}$$

Al variare di ρ si ha:

- $\rho = 0 \quad \Rightarrow \quad$ Due poli in -1 e -2
- $0 < \rho < 1/4 \quad \Rightarrow \quad$ Due poli reali negativi
- $\rho = 1/4 \quad \Rightarrow \quad$ Due poli reali coincidenti in -3/2
- $\rho > 1/4 \quad \Rightarrow \quad$ Due poli complessi con parte reale -3/2
- $\rho < 0 \quad \Rightarrow \quad$ Due poli reali

Sulla base di questa discussione, si possono tracciare i luoghi:

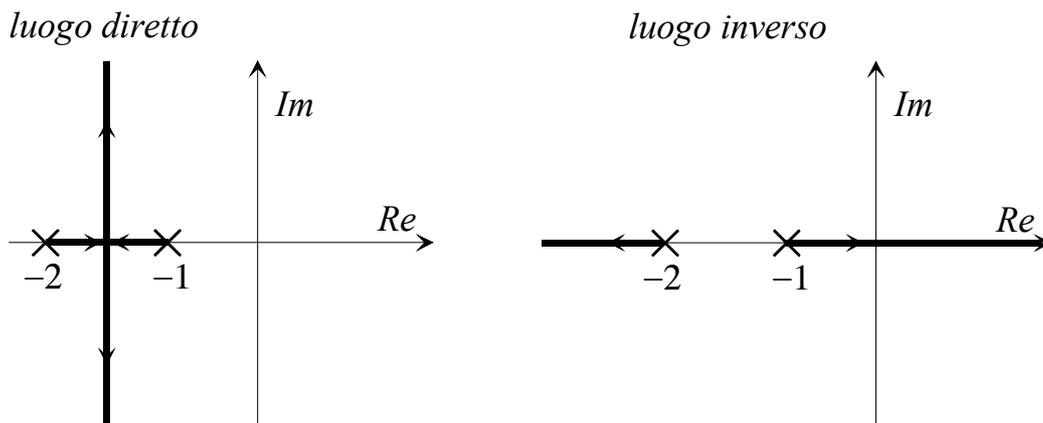


Fig. 2 : Luoghi delle radici per l'esempio introduttivo

È evidente che in questo caso il tracciamento dei luoghi è stato possibile grazie al fatto che le radici si possono calcolare esplicitamente.

Più in generale, l'equazione caratteristica può essere riscritta nella forma:

$$\frac{D(s)}{N(s)} = -\rho$$

ed equivale alle due equazioni nel campo reale:

$$\frac{|D(s)|}{|N(s)|} = |\rho|$$

$$\angle D(s) - \angle N(s) = \angle(-\rho)$$

La seconda equazione determina compiutamente la **forma del luogo**, mentre la prima consente di attribuire ad ogni punto del luogo il corrispondente valore di ρ , ossia, come si usa dire, di **"punteggiare"** il luogo rispetto a ρ .

Regole di tracciamento

Il tracciamento del luogo è fortemente agevolato dall'uso di alcune **semplici regole**. Ci limiteremo di seguito ad enunciare le principali (immediatamente verificabili sull'esempio introduttivo trattato sopra).

Regola 1: Detto m il grado del polinomio $N(s)$ (numero di zeri di $L(s)$) e n il grado del polinomio $D(s)$ (numero di poli di $L(s)$), il luogo diretto e il luogo inverso sono costituiti da n rami.

Regola 2: Il luogo è simmetrico rispetto all'asse reale.

Regola 3: Ogni ramo parte (per $\rho=0$) da un polo di L .

Regola 4: I rami terminano (per $|\rho|\rightarrow\infty$) in uno zero di L , oppure tendono all'infinito, secondo un asintoto. Tanto il luogo diretto quanto il luogo inverso presentano un numero di asintoti pari alla differenza $n-m$ (grado relativo di L).

Regola 5: Tutti gli asintoti si incontrano in un punto dell'asse reale, individuato dall'ascissa:

$$x_a = \frac{\sum_i z_i - \sum_k p_k}{n - m}$$

Regola 6: Gli $n - m$ asintoti formano con il semiasse reale positivo i seguenti angoli:

$$\vartheta_{ah} = \begin{cases} \frac{180^\circ + h360^\circ}{n - m}, & \text{per il luogo diretto} \\ \frac{h360^\circ}{n - m}, & \text{per il luogo inverso} \end{cases}$$

dove h assume i valori $0, 1, 2, \dots, n - m - 1$.

Regola 7: Tutti i punti dell'asse reale appartengono al luogo. Precisamente, appartengono al luogo diretto i punti dell'asse reale che hanno alla loro destra un numero dispari di poli e zeri di L , mentre appartengono al luogo inverso i punti dell'asse reale che hanno alla loro destra un numero pari (o nullo) di poli e zeri di L .

Regola 8: Se $n - m \geq 2$, la somma delle parti reali dei poli del sistema in anello chiuso si conserva al variare di ρ (regola del baricentro).

Regola 9: Il valore di $|\rho|$ corrispondente ad un punto \bar{s} del luogo si ottiene eseguendo il rapporto tra il prodotto delle distanze di \bar{s} dai poli di L ed il prodotto delle distanze di \bar{s} dagli zeri di L .

$$|\rho| = \frac{\prod_k |\bar{s} + p_k|}{\prod_i |\bar{s} + z_i|}$$

Regola 10: I punti di diramazione dall'asse reale e di ricongiunzione con esso (cioè i punti in cui il luogo abbandona l'asse reale per entrare nel piano complesso o vi ritorna) sono individuati da valori dell'ascissa x_d che risolvono la seguente equazione implicita (oltre all'equazione del luogo):

$$\sum_k \frac{1}{x_d + p_k} - \sum_i \frac{1}{x_d + z_i} = 0$$

E' utile osservare che l'innesto o il disinnesto di *due* rami in un punto (diramazione) dell'asse reale avviene, di norma, con tangente verticale.

Esempio 1

Consideriamo il sistema di funzione di trasferimento d'anello:

$$L(s) = \frac{\rho}{(s+1)(s+2)(s+3)}$$

Si ha $p_1 = 1, p_2 = 2, p_3 = 3, m = 0, n = 3$. Poiché $n-m = 3$, sia il luogo diretto sia quello inverso presentano 3 asintoti, che si incontrano sull'asse reale nel punto:

$$x_a = \frac{-p_1 - p_2 - p_3}{n - m} = \frac{-1 - 2 - 3}{3} = -2$$

Gli asintoti formano con l'asse reale gli angoli:

$$\vartheta_{ah} = \begin{cases} \frac{180^\circ + h360^\circ}{3} = \begin{cases} 60^\circ \\ 180^\circ \\ 300^\circ \end{cases} & \text{per il luogo diretto} \\ \frac{h360^\circ}{3} = \begin{cases} 0^\circ \\ 120^\circ \\ 240^\circ \end{cases} & \text{per il luogo inverso} \end{cases}$$

Tenendo conto della regola sull'appartenenza dei punti dell'asse reale, si possono facilmente tracciare i luoghi:

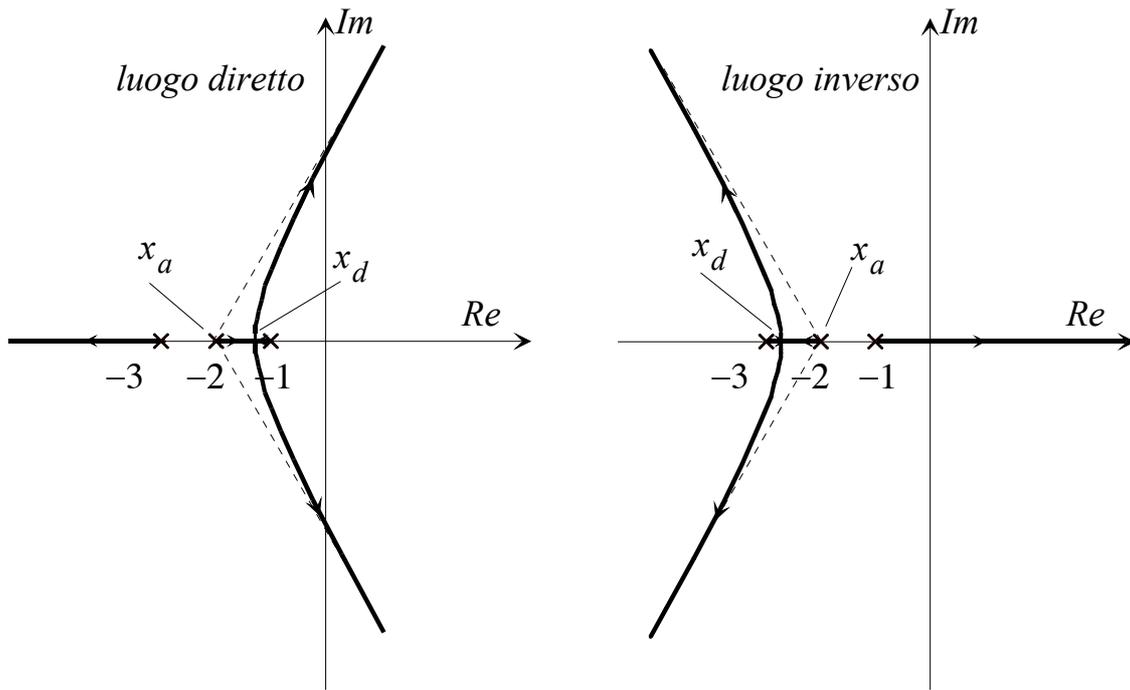


Fig. 3 : Luoghi delle radici per l'esempio 1

I punti di diramazione dell'asse reale si possono calcolare con la formula:

$$\frac{1}{x_d + 1} + \frac{1}{x_d + 2} + \frac{1}{x_d + 3} = 0 \Rightarrow 3x_d^2 + 12x_d + 11 = 0 \Rightarrow x_d = \begin{cases} -1.42 \text{ (L.D.)} \\ -2.58 \text{ (L.I.)} \end{cases}$$

E' interessante determinare per quali valori della costante di trasferimento ρ il sistema in anello chiuso risulta asintoticamente stabile (cioè tutti i suoi poli si trovano nel semipiano sinistro aperto).

Luogo diretto

Occorre determinare il valore ρ_M di ρ che porta due rami nel semipiano destro. Tuttavia i punti di intersezione con l'asse immaginario sono difficili da valutare. Si può però osservare che, essendo $n-m=3 (>2)$, si conserva la somma delle parti reali dei poli in anello chiuso. Tale somma, valutata per $\rho=0$, risulta -6 (somma dei poli di L). Ne consegue che quando due dei poli hanno parte reale nulla, il terzo avrà parte reale -6 . Eseguendo la punteggiatura in $s = -6$:

$$\rho_M = |-6 + p_1| |-6 + p_2| |-6 + p_3| = 5 \times 4 \times 3 = 60$$

Luogo inverso

Occorre determinare il valore ρ_m di ρ per cui uno dei rami entra nel semipiano destro. Eseguendo la punteggiatura in $s=0$, e tenendo conto che nel luogo inverso ρ assume valori negativi:

$$\rho_m = -|p_1| |p_2| |p_3| = -1 \times 2 \times 3 = -6$$

Possiamo pertanto concludere:

$$\boxed{\text{Sistema asintoticamente stabile} \Leftrightarrow -6 < \rho < 60}$$

Esempio 2

Consideriamo il sistema di funzione di trasferimento d'anello:

$$L(s) = \rho \frac{(s+4)(s+5)}{(s+1)(s+2)(s+3)}$$

Si ha $z_1 = 4, z_2 = 5, p_1 = 1, p_2 = 2, p_3 = 3, m = 2, n = 3$. Poiché $n-m = 1$, sia il luogo diretto sia quello inverso presentano un asintoto, che si sovrappone all'asse reale:

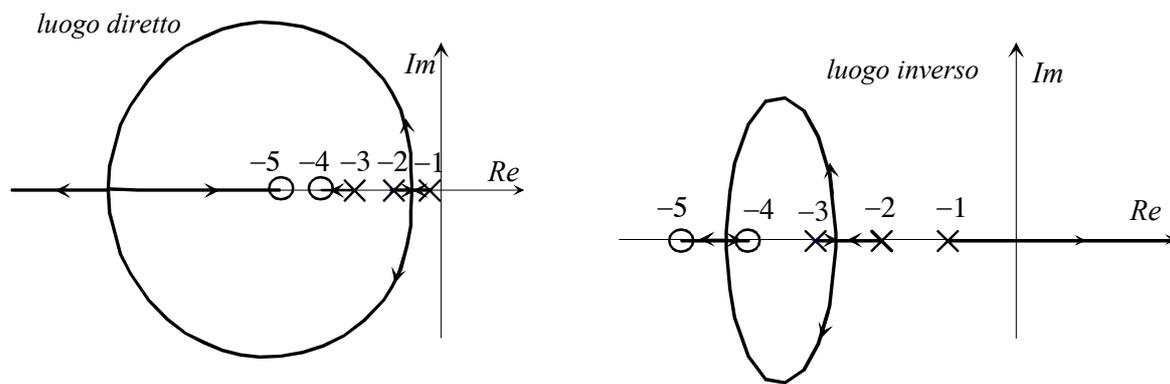


Fig. 4 : Luoghi delle radici per l'esempio 2

Analizziamo la stabilità del sistema in anello chiuso:

Luogo diretto

Tutto il luogo è compreso nel semipiano sinistro: il sistema è asintoticamente stabile per tutti i valori di $\rho > 0$.

Luogo inverso

Occorre determinare il valore di ρ per cui uno dei rami entra nel semipiano destro. Eseguendo la punteggiatura in $s=0$:

$$\rho_m = -\frac{|p_1||p_2||p_3|}{|z_1||z_2|} = -\frac{1 \times 2 \times 3}{4 \times 5} = -\frac{6}{20} = -0.3$$

Possiamo pertanto concludere:

Sistema asintoticamente stabile $\Leftrightarrow \rho > -0.3$

Progetto del regolatore

Oltre che per l'analisi della stabilità dei sistemi in anello chiuso, il metodo del luogo delle radici costituisce un valido ausilio anche per la sintesi del regolatore.

E' infatti possibile affrontare in maniera particolarmente intuitiva il problema della **stabilizzazione** in anello chiuso di sistemi instabili in anello aperto, nonché condurre a termine progetti che presentano specifiche di un certo dettaglio sulla forma dei transitori in anello chiuso.

Le caratteristiche di tali transitori sono infatti legate strettamente alla posizione nel piano complesso dei poli del sistema in anello chiuso, sui quali si ha completo controllo con il metodo del luogo delle radici.

Esempio

Si consideri il sistema di controllo in anello chiuso:

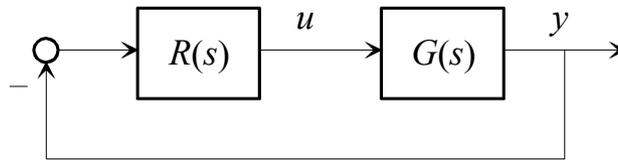


Fig. 5 : Sistema di controllo in anello chiuso

Sia:

$$G(s) = \frac{1}{(s+1)(s-2)}$$

Si vuole progettare $R(s)$ in modo tale che il sistema in anello chiuso abbia due poli in -2 .

E' opportuno esplorare la possibilità che la stabilizzazione sia ottenuta con un regolatore proporzionale. Si pone quindi:

$$R(s) = \rho_R \Rightarrow L(s) = R(s)G(s) = \frac{\rho_R}{(s+1)(s-2)}$$

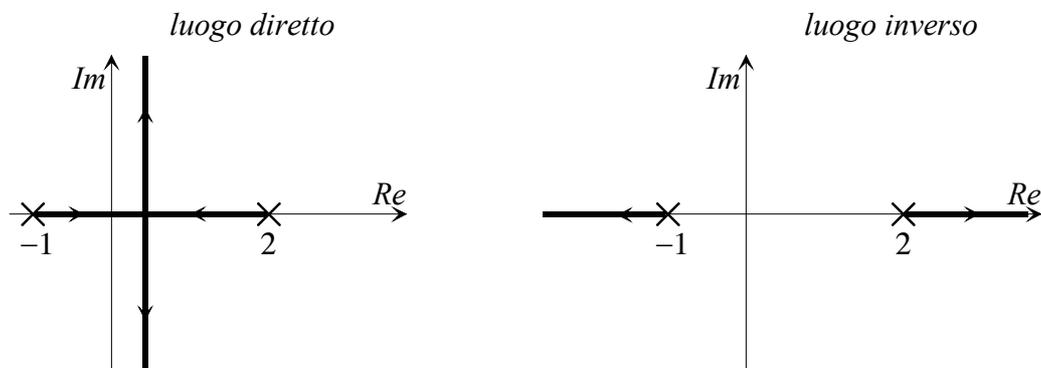


Fig. 6 : Luoghi delle radici con regolatore proporzionale

Per nessun valore di ρ_R il sistema in anello chiuso risulta asintoticamente stabile (c'è sempre un ramo nel semipiano destro). Occorre quindi complicare la struttura del regolatore, in modo da modificare la forma del luogo delle radici. In particolare si può cercare di portare l'asse verticale del luogo diretto nel semipiano sinistro.

Con riferimento al luogo diretto prima tracciato, si supponga di cancellare il polo in -1 con uno zero del regolatore, e di sostituirlo con un polo messo in modo tale che l'asintoto del luogo passi per il punto -2 : è immediato dedurre che il nuovo polo va posto nel punto -6 .

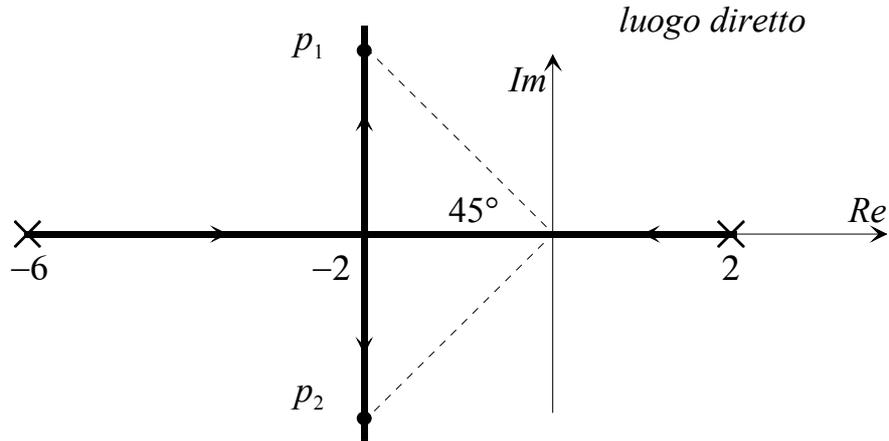


Fig. 7 : Luogo delle radici modificato

Il regolatore è quindi:

$$R(s) = \rho_R \frac{s+1}{s+6} \Rightarrow L(s) = \frac{\rho_R}{(s+6)(s-2)}$$

Punteggiando nel punto -2 si ottiene:

$$\rho_R = |-2-2| |-2+6| = 4 \times 4 = 16$$

Pertanto il regolatore che risolve il problema ha funzione di trasferimento:

$$R(s) = 16 \frac{s+1}{s+6}$$

Se viceversa si volessero posizionare i poli in anello chiuso nei punti p_1 e p_2 (sull'asse del luogo, a smorzamento $\zeta = 1/\sqrt{2}$), si eseguirebbe la punteggiatura del luogo in uno dei due punti, ottenendo (valutando le distanze dai poli di L con il teorema di Pitagora):

$$\rho_R = \sqrt{2^2 + 4^2} = 20.$$

Lezione 4

Schemi avanzati di controllo

Compensazione di disturbi misurabili

Nei sistemi di controllo è frequente il caso in cui il disturbo in linea d'andata sia misurabile. E' allora possibile sfruttare l'informazione data dalla misura e agire direttamente sulla variabile di controllo, anticipando l'effetto del disturbo sull'uscita, senza attendere che questo si manifesti in errore. Si usa parlare di **compensazione diretta** del disturbo quando:

- il disturbo è misurabile
- si esercita un'azione di controllo dipendente dalla sua misura.

Consideriamo dunque un sistema sotto controllo su cui agisce un disturbo:

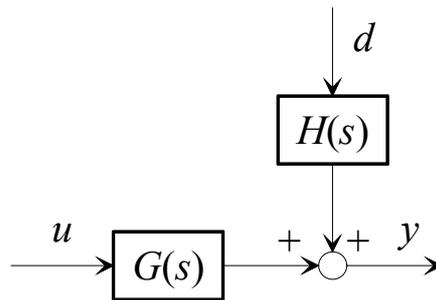


Fig. 1 : Sistema sotto controllo

Uno schema di compensazione potrebbe essere il seguente:

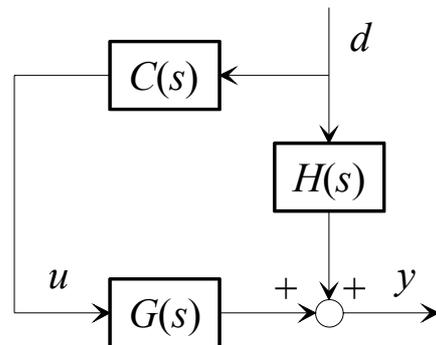


Fig. 2 : Compensazione del disturbo

La funzione di trasferimento $C(s)$ descrive il comportamento dinamico del **compensatore** e comprende anche l'effetto dinamico del trasduttore del disturbo, qualora esso non risulti trascurabile.

Per il progetto di $C(s)$ si può imporre che la funzione di trasferimento da d a y sia nulla:

$$\frac{Y(s)}{D(s)} = H(s) + C(s)G(s) = 0 ,$$

da cui:

$$C(s) = -\frac{H(s)}{G(s)} .$$

Il risultato di questa operazione non è tuttavia direttamente utilizzabile ogniqualvolta $G(s)$ ha zeri a parte reale positiva o ritardi (si otterrebbe un compensatore instabile o con azione predittiva) oppure quando il grado relativo di G è superiore a quello di H (si otterrebbe un compensatore con più zeri che poli). In ogni caso il compensatore derivante da questa formula può essere di ordine eccessivamente elevato.

Si può tuttavia utilizzare la formula precedente per ricavare la funzione di trasferimento di compensatori in grado di far fronte a tipologie di disturbi specifiche. In particolare sono di interesse le seguenti situazioni:

- **Disturbo costante** a regime. Per annullarne asintoticamente l'effetto è sufficiente approssimare la formula generale per $s=0$:

$$C(s) = \mu_C = -\frac{H(0)}{G(0)}.$$

- **Disturbo sinusoidale** alla pulsazione $\bar{\omega}$. Per annullarne asintoticamente l'effetto è sufficiente approssimare la risposta in frequenza solo ad una determinata pulsazione $\bar{\omega}$:

$$C(j\bar{\omega}) = -\frac{H(j\bar{\omega})}{G(j\bar{\omega})}$$

Si osservi che questa relazione nel campo complesso equivale a due relazioni nel campo reale (su modulo e fase). E' quindi opportuno parametrizzare la funzione di trasferimento $C(s)$ con due parametri, ricavabili dalla due condizioni. Ad esempio, si può porre:

$$C(s) = \frac{\mu}{(1+sT)^n}, \quad T > 0,$$

con n intero scelto in modo tale che la fase della risposta in frequenza possa assumere il valore imposto dalla condizione alla pulsazione $\bar{\omega}$.

- **Disturbo con trasformata di Fourier limitata alla pulsazione $\bar{\omega}$** . Per annullarne asintoticamente l'effetto, almeno approssimativamente, è sufficiente approssimare la risposta in frequenza fino alla pulsazione $\bar{\omega}$:

$$C(j\omega) = -\frac{H(j\omega)}{G(j\omega)}, \quad \omega < \bar{\omega}.$$

Di norma uno schema di compensazione viene associato ad uno schema di **retroazione** per rendere la reiezione del disturbo più robusta rispetto ad incertezze di modello e per soddisfare le altre specifiche. Si perviene allora allo schema di controllo di figura:

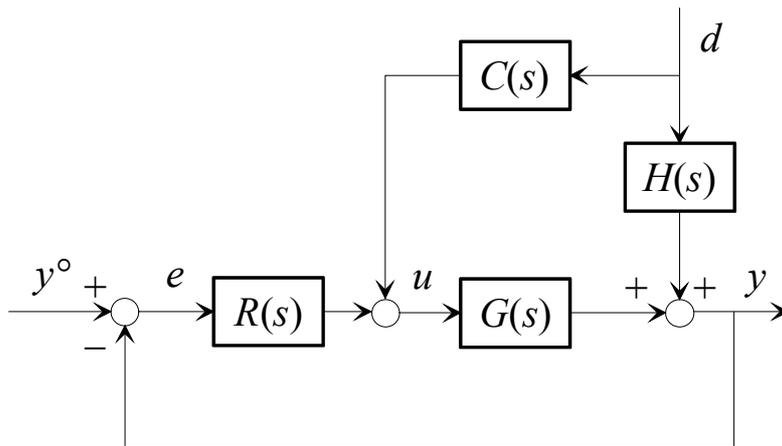


Fig. 3 : Controllo in retroazione con compensazione del disturbo

La funzione di trasferimento da d a y è ora:

$$\frac{Y(s)}{D(s)} = \frac{H(s) + C(s)G(s)}{1 + R(s)G(s)},$$

e si annulla ancora per:

$$C(s) = -\frac{H(s)}{G(s)}.$$

Pertanto, assunto naturalmente che il regolatore $R(s)$ stabilizzi l'anello di controllo, il progetto del compensatore C si può condurre esattamente come in anello aperto. In altre parole, il progetto del regolatore in retroazione e del compensatore in anello aperto sono disaccoppiati (indipendenti l'uno dall'altro).

Compensazione e prefiltraggio del riferimento

Come la misura del disturbo, anche il segnale di riferimento della variabile controllata può essere elaborato con un sistema dinamico per favorire un'azione di controllo più pronta.

Si consideri il seguente schema di **compensazione del segnale di riferimento**:

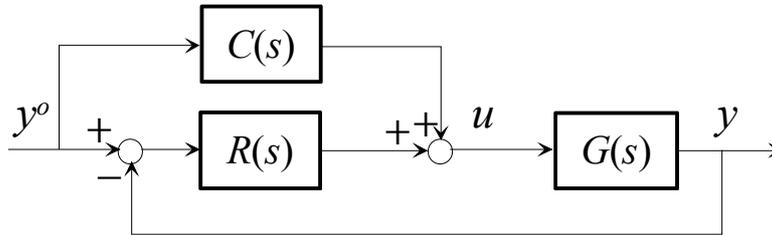


Fig. 4 : Compensazione del riferimento

In virtù della presenza del compensatore in anello aperto di funzione di trasferimento $C(s)$, la funzione di trasferimento dal riferimento alla variabile controllata diventa:

$$\frac{Y(s)}{Y^o(s)} = \frac{R(s)G(s) + C(s)G(s)}{1 + R(s)G(s)}$$

Ponendo quindi:

$$C(s) = \frac{1}{G(s)},$$

si otterrebbe il risultato ideale $Y(s)/Y^o(s) \equiv 1$. Naturalmente si tratta di una scelta non realizzabile a causa di molteplici ragioni (causalità del compensatore, eventuale presenza di zeri nel semipiano destro o di ritardi di tempo in G). Tuttavia rimane un'utile indicazione di progetto, nel senso che si cercherà di rendere la risposta in frequenza di C simile a quella di $1/G$ nella banda di interesse.

Alternativamente, si può **prefiltrare il segnale di riferimento**, come nello schema seguente:

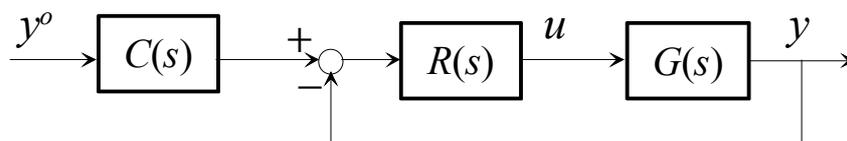


Fig. 5 : Prefiltraggio del riferimento

In questo modo la funzione di trasferimento dal riferimento alla variabile controllata diventa:

$$\frac{Y(s)}{Y^o(s)} = C(s)F(s),$$

con:

$$F(s) = \frac{R(s)G(s)}{1 + R(s)G(s)}.$$

Il prefiltro di funzione di trasferimento $C(s)$ può quindi essere scelto in modo da modificare la funzione di trasferimento $Y(s)/Y^o(s)$, senza modificare la funzione di trasferimento d'anello, e quindi senza modificare le proprietà dell'anello (margine di fase, pulsazione critica, reiezione dei disturbi). Se per esempio scegliamo:

$$C(s) = \mu_C = \frac{1}{F(0)},$$

otterremo guadagno unitario della funzione di trasferimento $Y(s)/Y^o(s)$, anche in assenza di azione integrale nell'anello. Si osservi tuttavia che mentre la precisione statica ottenuta introducendo un integratore nel regolatore è robusta, ossia garantita anche in presenza di incertezza sul guadagno del sistema, questo non è più vero se la precisione statica viene forzata con il prefiltro.

Si supponga invece che, per effetto di dinamiche non modellate in alta frequenza o di disturbi sulla linea di retroazione sia consigliabile limitare la pulsazione critica dell'anello di controllo al valore $\bar{\omega}_c$. Se fosse prescritta una banda nell'inseguimento del segnale di riferimento di valore superiore $\bar{\omega}_b > \bar{\omega}_c$, si potrebbe progettare un prefiltro avente una funzione di trasferimento di tipo **passaalto**, di espressione:

$$C(s) = \frac{1 + s/\bar{\omega}_c}{1 + s/\bar{\omega}_b}.$$

Naturalmente, la migliore prontezza nell'inseguimento del segnale di riferimento va a discapito della moderazione del controllo, in quanto la variabile di controllo viene maggiormente sollecitata.

Gli schemi di **compensazione e prefiltraggio** possono poi essere combinati in uno schema più generale:

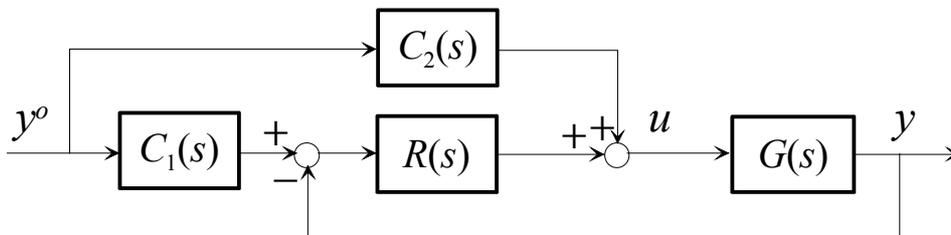


Fig. 6 : Compensazione e prefiltraggio del riferimento

Si ottiene:

$$\frac{Y(s)}{Y^o(s)} = \frac{[C_1(s)R(s) + C_2(s)]G(s)}{1 + R(s)G(s)}.$$

Ponendo:

$$C_1(s) = F^o(s), \quad C_2(s) = F^o(s)G(s)^{-1},$$

si ottiene:

$$\frac{Y(s)}{Y^o(s)} = F^o(s).$$

La funzione di trasferimento $F^o(s)$ deve quindi essere scelta come il modello di riferimento per il sistema in anello chiuso. Per la causalità del compensatore $C_2(s)$ deve avere grado relativo (differenza tra il numero dei poli ed il numero degli zeri) almeno pari a quello di $G(s)$. Inoltre, deve essere caratterizzata degli stessi eventuali zeri a parte reale positiva o ritardi di tempo presenti in $G(s)$ affinché essi non vengano cancellati dal compensatore.

Si osservi, in conclusione, che questa soluzione complica in modo considerevole la realizzazione del sistema di controllo.

Regolatori in cascata

In molte applicazioni avviene che il sistema sotto controllo sia strutturabile nella connessione in serie di due funzioni di trasferimento. Si assuma che all'uscita del primo sottosistema si sommi un disturbo d e che la variabile intermedia v tra i due sottosistemi sia misurabile.

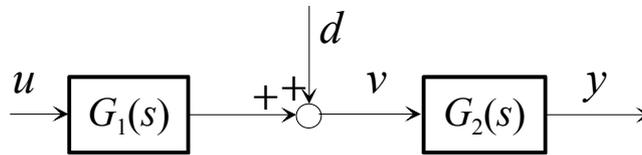


Fig. 7 : Sistemi in cascata

Spesso avviene che la dinamica di G_1 sia molto più favorevole, in termini di progetto del controllore, di quella di G_2 . Per esempio, G_1 potrebbe essere a fase minima e G_2 no oppure, pur essendo entrambe le funzioni di trasferimento a fase minima, G_1 potrebbe avere dinamica molto più veloce rispetto a G_2 e quindi potrebbe dar luogo ad una dinamica in anello chiuso veloce con sforzo del controllo molto più limitato di quello richiesto per accelerare G_2 .

Tutte queste circostanze si manifestano in particolare quando G_1 costituisce l'**attuatore** per il sistema sotto controllo. Se per esempio $G_1(s)$ fosse la funzione di trasferimento di un motore elettrico a corrente continua, u sarebbe la tensione del circuito d'armatura, i la corrente, mentre il disturbo d potrebbe essere associato alla forza controelettrica. Se viceversa l'attuatore fosse una valvola idraulica, u sarebbe l'area di apertura della valvola, v la portata di liquido, mentre il disturbo d sarebbe associabile alla dipendenza della portata dal livello di liquido a monte della valvola.

In tutte queste situazioni può risultare molto opportuno strutturare il sistema di controllo con due anelli di controllo innestati, secondo la metodologia del **controllo in cascata**.

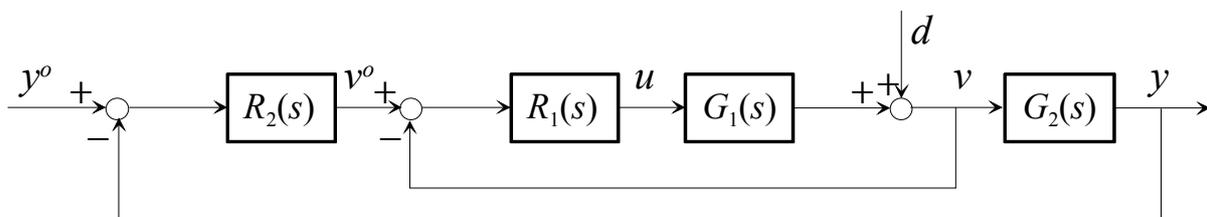


Fig. 8 : Schema di controllo in cascata

Il progetto del regolatore interno R_1 si basa esclusivamente sulla funzione di trasferimento $G_1(s)$ e mira a far sì che v inseguia il suo riferimento v^o su un'ampia banda e che sulla stessa banda il disturbo d sia quanto più possibile attenuato. Nel progetto del regolatore esterno R_2 potremo limitare la banda a valori decisamente inferiori rispetto alla banda dell'anello interno: questo consente di approssimare, su questa banda, la dinamica dell'anello interno come infinitamente veloce ($v \approx v^o$) e quindi basare il progetto di R_2 esclusivamente sulla funzione di trasferimento $G_2(s)$.

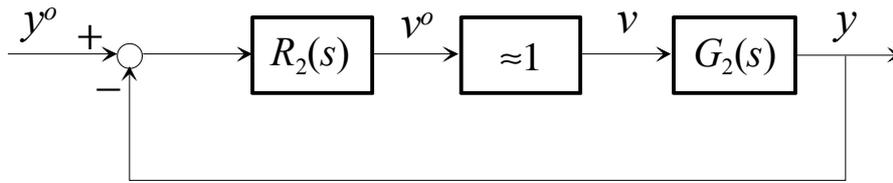


Fig. 9 : Schema per il progetto del regolatore esterno

I due regolatori R_1 e R_2 vengono quindi progettati secondo un **disaccoppiamento in frequenza**, che da un lato semplifica il progetto, suddividendolo in due sottoprogetti più semplici, dall'altra conferisce al sistema delle prestazioni di norma molto superiori rispetto a quelle ottenibili con un unico controllore chiuso su y .

Esempio

Consideriamo un motore a corrente continua:

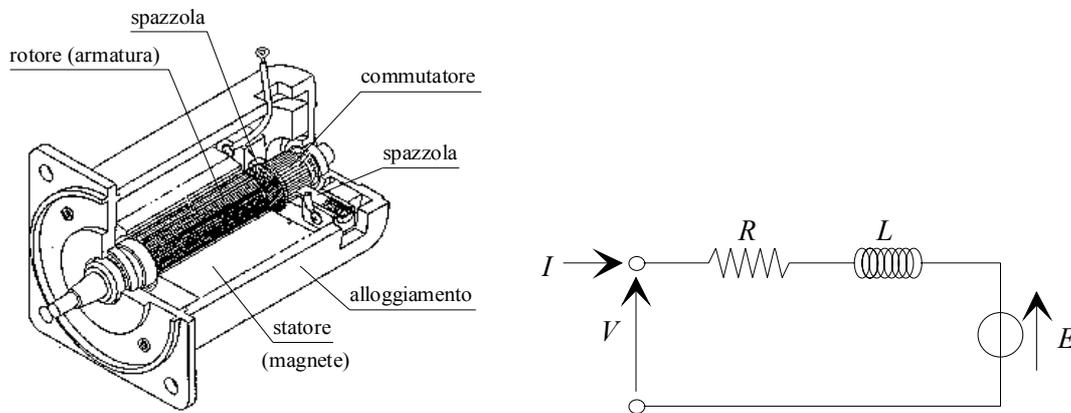


Fig. 10 : Motore a corrente continua e schema del circuito d'armatura

Il motore, caratterizzato meccanicamente da un momento d'inerzia J , eroga una coppia proporzionale alla corrente I nel circuito d'armatura. Su questo circuito, avente resistenza R ed induttanza L , agisce anche la forza controelettromotrice E , proporzionale alla velocità angolare ω del motore. Sono misurabili la posizione ϑ del motore e la corrente I del circuito d'armatura, e l'obiettivo ultimo è il controllo del moto del sistema, ossia della posizione ϑ . Il sistema è retto dalle seguenti equazioni:

$$V(t) = RI(t) + L\dot{I}(t) + E(t)$$

$$E(t) = K\omega(t)$$

$$\tau(t) = KI(t)$$

$$\tau(t) = J\dot{\omega}(t)$$

$$\dot{\vartheta}(t) = \omega(t)$$

che trasformate secondo Laplace danno luogo al seguente sistema:

$$I(s) = \frac{1}{sL + R}(V(s) - E(s))$$

$$E(s) = K\omega(s)$$

$$\tau(s) = KI(s)$$

$$\omega(s) = \frac{1}{Js}\tau(s)$$

$$\vartheta(s) = \frac{1}{s}\omega(s)$$

Lo schema a blocchi del sistema è quindi il seguente:

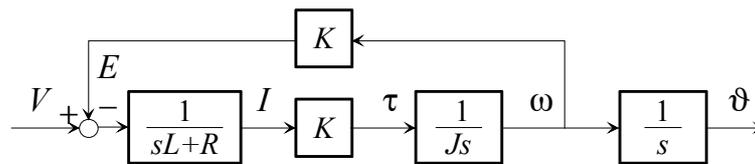


Fig. 11 : Schema a blocchi del motore a corrente continua

Uno schema di controllo in cascata per questo sistema è il seguente:

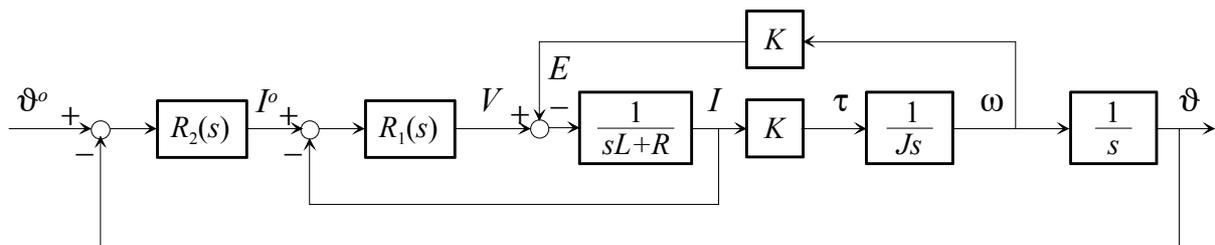


Fig. 12 : Controllo in cascata del motore a corrente continua

Nel progetto del regolatore più interno, R_1 , si assumerà la forza contro elettromotrice come un disturbo di bassa frequenza. Vista la dinamica veloce legata ai transistori elettrici (la costante di tempo elettrica L/R è di norma dell'ordine dei ms) si potrà progettare $R_1(s)$ per ottenere bande dell'ordine del kHz . Una volta chiuso l'anello di controllo della corrente, questo potrà ritenersi istantaneo ai fini del progetto del controllore di posizione esterno:

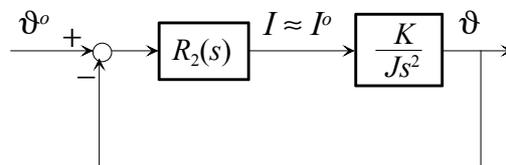


Fig. 13 : Schema per il progetto del controllore esterno (di posizione)

Il controllore dell'anello esterno R_2 sarà progettato in modo da ottenere una banda decisamente inferiore a quella dell'anello di corrente (qualche decina di Hz).

Predittore di Smith

E' noto che i sistemi a fase non minima (che contengono zeri nel semipiano destro e/o ritardi di tempo) comportano problemi nel controllo: l'impossibilità di cancellare tali termini nel regolatore lascia come unica possibilità, nel progetto classico del controllore, la limitazione della pulsazione critica dell'anello di controllo, in modo da ottenere accettabili margini di stabilità.

Esiste in realtà uno schema di controllo più avanzato, che prende il nome di schema a **predittore di Smith** che consente di affrontare il problema in modo diverso.

Si supponga che il sistema sotto controllo sia **asintoticamente stabile** e che la sua funzione di trasferimento sia fattorizzabile nel prodotto:

$$G(s) = G_1(s)G_2(s), \quad G_2(0) = 1,$$

con $G_2(s)$ a fase non minima (in particolare può essere $G_2(s) = e^{-s\tau}$).

Lo schema a predittore di Smith è il seguente:

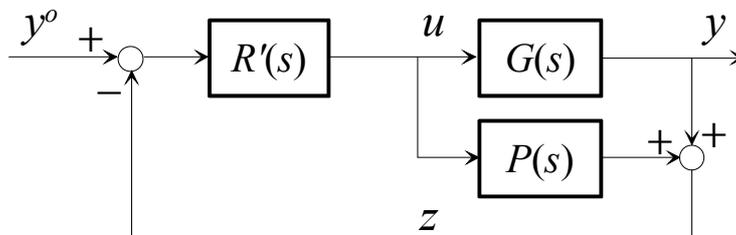


Fig. 14 : Schema di controllo con predittore di Smith

Il controllore prevede quindi, oltre alla funzione di trasferimento $R'(s)$, anche una funzione di trasferimento, $P(s)$, posta in parallelo a $G(s)$. A questa funzione di trasferimento, che costituisce il **predittore**, viene attribuita la funzione di trasferimento:

$$P(s) = [1 - G_2(s)]G_1(s) = G_1(s) - G(s).$$

Si osservi a questo punto che, tagliando l'anello in corrispondenza dell'ingresso o dell'uscita di $R'(s)$, si ottiene come funzione di trasferimento d'anello:

$$L(s) = R'(s)[G(s) + P(s)] = R'(s)G_1(s).$$

Pertanto nella funzione di trasferimento d'anello **non compare più il termine a fase non minima** $G_2(s)$ o, in altre parole, il progetto di $R'(s)$ si può condurre sulla base della parte a fase minima $G_1(s)$. Non si hanno più quindi i limiti di banda associati alla fase non minima del sistema, anche se rimarranno limiti legati ad altri fattori (moderazione del controllo, eventuali dinamiche di alta frequenza non modellate, disturbi in linea di retroazione).

Il nome predittore risulta chiaro se si pone $G_2(s) = e^{-s\tau}$, ossia se la parte a fase non minima è costituita da un ritardo puro, come del resto originariamente previsto nella concezione del metodo. In questo caso, infatti, considerando la variabile z che viene retroazionata al regolatore $R'(s)$ (si veda la Fig. 14) si ha:

$$Z(s) = [G(s) + P(s)]U(s) = G_1(s)U(s) = G_2(s)^{-1}G_1(s)G_2(s)U(s) = e^{s\tau}Y(s).$$

Nel dominio del tempo questa relazione si traduce in:

$$z(t) = y(t + \tau).$$

La situazione è quindi analoga a quella che si avrebbe se nello schema di controllo venisse **retroazionata** non y , ma la sua **predizione** $z(t) = y(t + \tau)$: si osservi che questa circostanza è piuttosto ovvia se si considera che in un ritardo puro l'ingresso ad un dato istante (utilizzato dal predittore) non è altro che il valore che assumerà l'uscita esaurito il ritardo stesso.

Lo schema a predittore di Smith si lascia poi interpretare in termini di uno schema classico di retroazione, come segue:

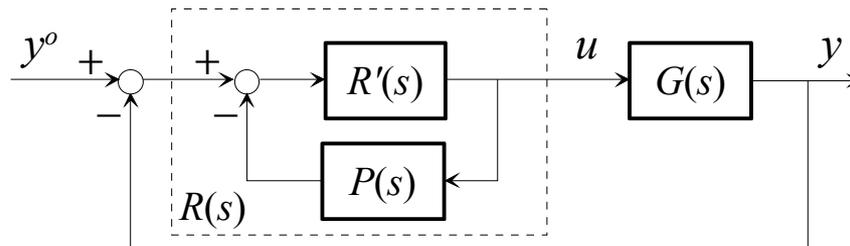


Fig. 15 : Interpretazione dello schema di controllo con predittore di Smith

Il sistema retroazionato tratteggiato è interpretabile come un regolatore classico in retroazione a $G(s)$, ed ha funzione di trasferimento:

$$R(s) = \frac{R'(s)}{1 + R'(s)P(s)} = \frac{R'(s)}{1 + R'(s)[1 - G_2(s)]G_1(s)}.$$

Esprese le funzioni di trasferimento in termini di rapporti di polinomi e tenuto conto dell'eventuale termine di ritardo puro:

$$R'(s) = \frac{N_R(s)}{D_R(s)}, \quad G_1(s) = \frac{N_1(s)}{D_1(s)}, \quad G_2(s) = \frac{N_2(s)}{D_2(s)} e^{-s\tau},$$

si ha:

$$R(s) = \frac{\frac{N_R(s)}{D_R(s)}}{1 + \frac{N_R(s)}{D_R(s)} \left[1 - \frac{N_2(s)}{D_2(s)} e^{-s\tau} \right] \frac{N_1(s)}{D_1(s)}} = \frac{N_R(s)D_1(s)D_2(s)}{D_R(s)D_1(s)D_2(s) + N_1(s)N_R(s)[D_2(s) - N_2(s)e^{-s\tau}]}$$

Da questa espressione scaturiscono alcune osservazioni:

- Il fatto che $R(s)$ abbia a numeratore il denominatore di G ($D_1(s)D_2(s)$), giustifica l'ipotesi di asintotica stabilità introdotta all'inizio della trattazione;
- Se il sistema presenta un ritardo di tempo ($\tau \neq 0$) il regolatore ha funzione di trasferimento trascendente;
- Poiché $G_2(0)=1$, e quindi $D_2(0)=N_2(0)$, risulta:

$$\lim_{s \rightarrow 0} R(s) = \lim_{s \rightarrow 0} \frac{N_R(s)}{D_R(s)} = \lim_{s \rightarrow 0} R'(s).$$

Pertanto $R(s)$ ha lo stesso comportamento di $R'(s)$ in bassa frequenza, il che comporta che se il progetto statico impone azione integrale nel regolatore, questa deve essere presente nella funzione di trasferimento $R'(s)$.

Le prestazioni del predittore di Smith per sistemi con ritardo si deteriorano considerevolmente nel caso di **valutazione imprecisa del ritardo** stesso, il che ne limita in qualche misura l'utilità pratica. Inoltre la realizzazione del ritardo di tempo nella funzione di trasferimento $P(s)$ comporta problemi di implementazione.

Regolatori anti-windup

Il comportamento degli attuatori è stato finora sempre ipotizzato lineare. In realtà tutti gli attuatori presentano una saturazione, che può essere rappresentata nel sistema di controllo come nel seguente schema:

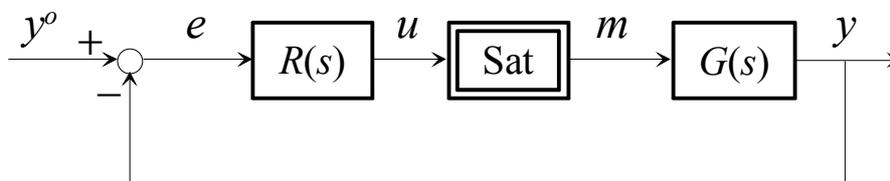


Fig. 16 : Sistema di controllo con saturazione

Il blocco non lineare “Sat” rappresenta una caratteristica di saturazione, che per semplicità supporremo simmetrica, come in figura:

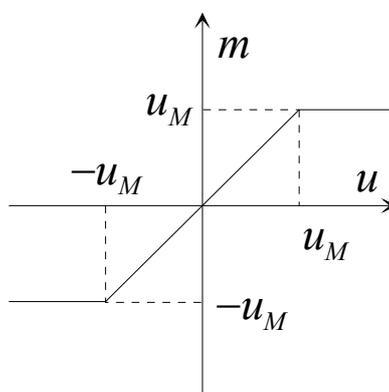


Fig. 17 : Caratteristica di saturazione

Dal punto di vista matematico, la caratteristica di saturazione si esprime quindi nei seguenti termini:

$$m(t) = \begin{cases} -u_M, & u(t) < -u_M \\ u(t) & -u_M \leq u(t) \leq u_M \\ u_M & u(t) > u_M \end{cases}$$

Se il regolatore presenta azione integrale, la combinazione di questa con la saturazione dell'attuatore può dar luogo ad un effetto indesiderato, noto come **carica integrale** (o **integral windup**), che può pregiudicare le prestazioni del sistema di controllo. Per comprendere il fenomeno, si faccia riferimento al caso in cui il regolatore sia puramente integrale ($R(s) = k_I/s$): le considerazioni che si svolgeranno in questo caso sono peraltro estendibili al caso generale. Con riferimento alla Fig.16, se nel corso dell'operato del sistema di controllo accade che l'errore e resta a lungo dello stesso segno, ad esempio positivo, la variabile di stato dell'integratore, e quindi la variabile u , cresce indefinitamente e può senz'altro superare, anche di molto, il valore di saturazione u_M , facendo operare l'attuatore in regime di saturazione ($m \equiv u_M$). Quando, per effetto di questa azione di controllo perdurante al valore massimo, la variabile di controllo y avrà superato il valore del suo riferimento y^o ,

rendendo quindi l'errore e negativo, sarebbe auspicabile che la variabile m lasciasse rapidamente il valore massimo u_M ed assumesse valori inferiori. Affinché ciò avvenga, tuttavia, è necessario attendere che la variabile u rientri dai valori elevati raggiunti nella fase precedente, fino ad assumere valori inferiori al limite di saturazione u_M : questa fase di scarica, o **desaturazione**, dell'azione integrale, può essere di durata rilevante e dà luogo ad un comportamento anomalo del sistema di controllo.

Per ovviare a questo inconveniente, occorre realizzare il regolatore in modo differente, senza ovviamente alterarne la funzione di trasferimento. Si ponga dunque:

$$R(s) = \frac{N_R(s)}{D_R(s)},$$

con N_R e D_R polinomi primi tra loro e $D_R(0)=0$ per effetto della presenza dell'azione integrale. Si considerino quindi le seguenti elaborazioni formali sulla funzione di trasferimento:

$$R(s) = \frac{N_R(s)}{D_R(s)} = \frac{N_R(s)}{\varphi(s)} \frac{\varphi(s)}{D_R(s)} = \frac{N_R(s)}{\varphi(s)} \frac{1}{\frac{D_R(s)}{\varphi(s)}} = \frac{N_R(s)}{\varphi(s)} \frac{1}{1 - \left(1 - \frac{D_R(s)}{\varphi(s)}\right)} = K(s) \frac{1}{1 - H(s)},$$

dove $\varphi(s)$ è un polinomio, le cui proprietà saranno analizzate in seguito, mentre $K(s)$ e $H(s)$ sono le seguenti funzioni di trasferimento:

$$K(s) = \frac{N_R(s)}{\varphi(s)}, \quad H(s) = \frac{\varphi(s) - D_R(s)}{\varphi(s)}.$$

Si osservi che $H(s)$ ha per costruzione guadagno unitario. La funzione di trasferimento $R(s)$ viene quindi ottenuta come serie di $K(s)$ e di un sistema retroazionato positivamente con linea d'andata unitaria e linea di retroazione $H(s)$.

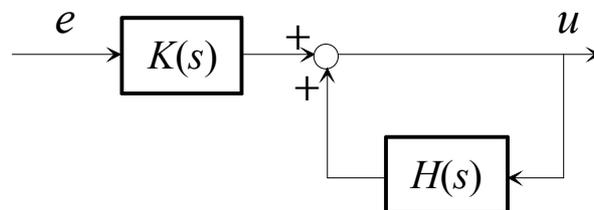


Fig. 18 : Realizzazione alternativa di $R(s)$

Da quanto precede, risulta chiaro che il polinomio $\varphi(s)$ deve soddisfare le seguenti condizioni:

- deve avere tutte le radici a parte reale negativa (per l'asintotica stabilità di $K(s)$)
- deve avere grado non inferiore a quello di $N_R(s)$ (per la causalità di $K(s)$)
- deve rendere $H(s)$ strettamente proprio, per evitare un anello algebrico

Si osservi che nel caso, frequente, in cui $R(s)$ sia un sistema proprio, non strettamente, a fase minima, scritti i polinomi N_R e D_R nella forma:

$$\begin{aligned} N_R(s) &= \rho(s^n + b_n s^{n-1} + \dots + b_2 s + b_1) \\ D_R(s) &= s^n + a_n s^{n-1} + \dots + a_2 s \end{aligned}$$

tutte le condizioni sopra riportate sono soddisfatte ponendo $\varphi(s)=N_R(s)/\rho$, dove ρ è la costante di trasferimento di $R(s)$, da cui:

$$K(s) = \rho, \quad H(s) = \frac{(b_n - a_n)s^{n-1} + \dots + (b_2 - a_2)s + b_1}{s^n + b_n s^{n-1} + \dots + b_2 s + b_1}.$$

La realizzazione anti carica integrale (o **anti windup**) del regolatore si ottiene con lo schema seguente, in cui si suppone misurabile la variabile manipolabile m , a valle della saturazione dell'attuatore:

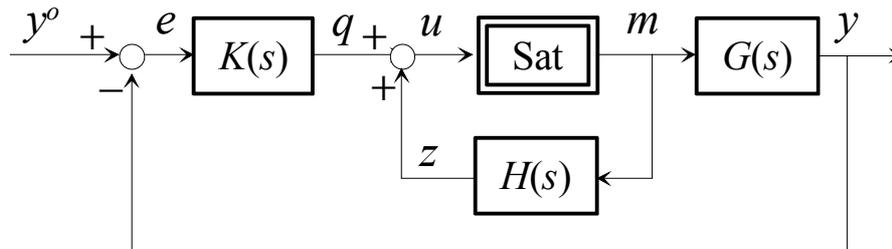


Fig. 19 : Realizzazione anti windup del regolatore

Se invece la variabile m non è accessibile, si dovrebbe riprodurre all'interno del regolatore, una caratteristica, indicata con \hat{Sat} in Fig. 20, uguale a quella fisica dell'attuatore, e porla sulla linea d'andata dello schema a blocchi che realizza il regolatore, secondo lo schema seguente:

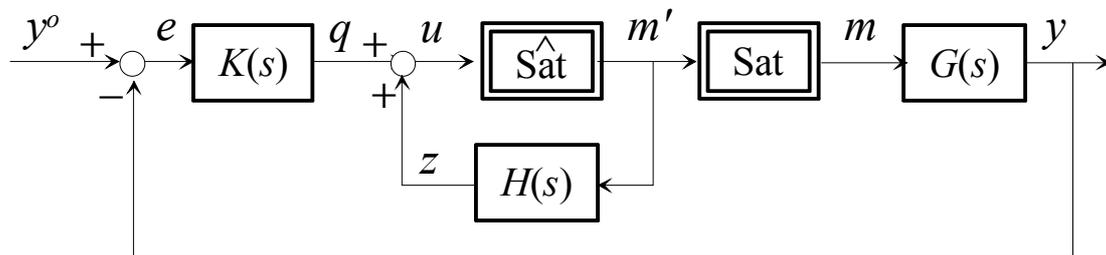


Fig. 20 : Realizzazione anti windup del regolatore senza misura di m

Facendo riferimento alla Fig. 19 (ma le stesse considerazioni valgono per lo schema di Fig. 20), si supponga nuovamente che l'errore e rimanga dello stesso segno, ad esempio positivo, a lungo. Se il regolatore ha guadagno positivo, anche q assumerà valori positivi. Si supponga quindi l'attuatore in saturazione con $m=u_M$. Poiché $H(0)=1$, anche la variabile z si assesterà sul valore u_M . Quando, per effetto della prolungata azione dell'attuatore, l'errore cambia segno, anche la variabile q cambierà segno, istantaneamente se $K(s)$ è algebrica (come avviene con la scelta sopra proposta per $\varphi(s)$) o comunque con la dinamica legata alle radici di $\varphi(s)$, arbitrariamente assegnabili. Il fatto che q assuma valori negativi, unito al fatto sopra ricordato che $z=u_M$, implica poi che u assuma immediatamente valori inferiori a u_M , facendo subito uscire l'attuatore dalla saturazione, ripristinandone quindi il comportamento lineare.

Esempio

Si consideri un regolatore PI:

$$R(s) = K_P \left(1 + \frac{1}{sT_I} \right) = K_P \frac{s + 1/T_I}{s} = \frac{N_R(s)}{D_R(s)},$$

con:

$$N_R(s) = K_P(s + 1/T_I), \quad D_R(s) = s.$$

Posto:

$$\varphi(s) = \frac{N_R(s)}{\rho} = \frac{N_R(s)}{K_P} = s + \frac{1}{T_I},$$

si ha:

$$K(s) = \rho = K_P, \quad H(s) = \frac{s + 1/T_I - s}{s + 1/T_I} = \frac{1}{1 + sT_I},$$

da cui segue lo schema anti windup:

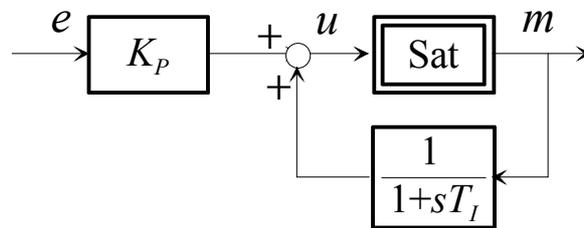


Fig. 21 : Realizzazione anti windup del regolatore PI

Controllo di sistemi multivariabili

Le tecniche di controllo nel dominio della frequenza finora studiate fanno unicamente riferimento a sistemi SISO. Il controllo di sistemi con più ingressi e/o più uscite è argomento complesso e richiede nozioni di teoria del controllo avanzate. Si osservi, d'altra parte, che le tecniche di controllo nel dominio del tempo per l'assegnamento degli autovalori si prestano naturalmente all'applicazione su sistemi multivariabili.

Rimanendo nel dominio della frequenza, una strada percorribile con sforzo ragionevolmente contenuto consiste nel ricondurre il problema di controllo a **tanti problemi di controllo monovariabili**, risolvibili secondo le tecniche note.

Nel seguito si daranno cenni a due procedure utilizzabili per ricondurre il problema MIMO a una serie di problemi SISO: il progetto di **controllori di disaccoppiamento** ed il progetto di **controllori decentralizzati**.

Controllori di disaccoppiamento

Si consideri un sistema con m ingressi e m uscite ($m > 1$), con funzione di trasferimento matriciale $\mathbf{G}(s)$.

Siano y_i le uscite del sistema e y_i^o i corrispondenti segnali di riferimento. Si dice che un controllore realizza il **disaccoppiamento** (o la **non interazione**) se è in grado di fare in modo che risulti:

y_i indipendente da $y_k^o \forall k \neq i$.

Il controllore deve quindi essere in grado di rendere il sistema in anello chiuso completamente disaccoppiato. Si consideri a questo scopo il seguente sistema di controllo:

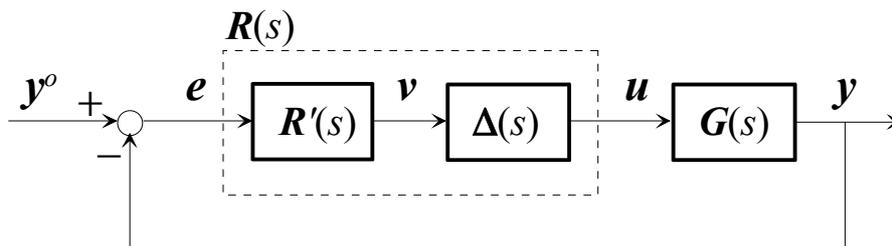


Fig. 22 : Sistema di controllo con disaccoppiatore

Il controllore, di ingresso (vettoriale) e , uscita (vettoriale) u e di funzione di trasferimento $\mathbf{R}(s)$, è costituito dalla serie di due sistemi a m ingressi e m uscite. E' infatti descritto dalle relazioni:

$$\mathbf{U}(s) = \mathbf{\Delta}(s)\mathbf{V}(s), \quad \mathbf{V}(s) = \mathbf{R}'(s)\mathbf{E}(s) .$$

Il sistema di funzione di trasferimento $\mathbf{\Delta}(s)$, che prende il nome di **disaccoppiatore**, deve essere progettato in modo tale che il sistema di ingresso v ed uscita y abbia funzione di trasferimento $\mathbf{G}(s)\mathbf{\Delta}(s)$ **diagonale**.

A questo punto, progettando la restante parte del controllore con funzione di trasferimento $\mathbf{R}'(s)$ diagonale, si rende la funzione di trasferimento d'anello $\mathbf{G}(s)\mathbf{\Delta}(s)\mathbf{R}'(s)$ diagonale, dal che consegue che anche la funzione di trasferimento del sistema in anello chiuso (da y^o a y) è diagonale, e quindi che viene realizzata la condizione di disaccoppiamento prima enunciata. Esistono condizioni sufficienti per l'esistenza del disaccoppiatore $\mathbf{\Delta}(s)$.

Vediamo come si conduce il progetto nel caso semplice di sistema a due ingressi e due uscite ($m=2$). Sia dunque:

$$\mathbf{G}(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix}$$

la funzione di trasferimento del sistema sotto controllo e:

$$\mathbf{\Delta}(s) = \begin{bmatrix} \Delta_{11}(s) & \Delta_{12}(s) \\ \Delta_{21}(s) & \Delta_{22}(s) \end{bmatrix}$$

quella del disaccoppiatore. Imponiamo che il prodotto $\mathbf{G}(s)\mathbf{\Delta}(s)$ sia diagonale, ovvero che i termini fuori diagonale siano nulli:

$$\begin{cases} G_{11}(s)\Delta_{12}(s) + G_{12}(s)\Delta_{22}(s) = 0 \\ G_{21}(s)\Delta_{11}(s) + G_{22}(s)\Delta_{21}(s) = 0 \end{cases}$$

Il problema è indeterminato (vi sono due equazioni per le quattro incognite $\Delta_{ij}(s)$). Una soluzione semplice si ottiene ponendo:

$$\begin{cases} \Delta_{11}(s) = \Delta_{22}(s) = 1 \\ \Delta_{12}(s) = -\frac{G_{12}(s)}{G_{11}(s)} \\ \Delta_{21}(s) = -\frac{G_{21}(s)}{G_{22}(s)} \end{cases}$$

Utilizzando questo disaccoppiatore, il sistema disaccoppiato sotto controllo ha funzione di trasferimento diagonale, espressa da:

$$\mathbf{G}(s)\mathbf{\Delta}(s) = \begin{bmatrix} G_{d1}(s) & 0 \\ 0 & G_{d2}(s) \end{bmatrix},$$

con:

$$\begin{cases} G_{d1}(s) = G_{11}(s)\Delta_{11}(s) + G_{12}(s)\Delta_{21}(s) = G_{11}(s) - \frac{G_{12}(s)G_{21}(s)}{G_{22}(s)} = \frac{\det(\mathbf{G}(s))}{G_{22}(s)} \\ G_{d2}(s) = G_{21}(s)\Delta_{12}(s) + G_{22}(s)\Delta_{22}(s) = G_{22}(s) - \frac{G_{21}(s)G_{12}(s)}{G_{11}(s)} = \frac{\det(\mathbf{G}(s))}{G_{11}(s)} \end{cases}$$

Il problema di controllo multivariabile è stato quindi ricondotto a **due problemi di controllo monovariabili**, in ciascuno dei quali occorre progettare un regolatore $R'_i(s)$ ($i=1,2$) sulla base della funzione di trasferimento $G_{di}(s)$.

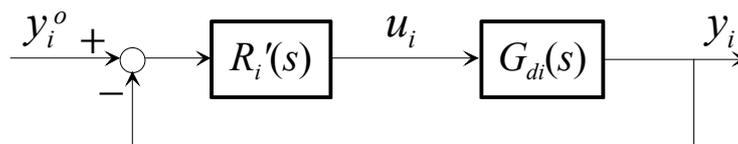


Fig. 23 : Problema di controllo monovariabile

Si osservi che le funzioni di trasferimento $G_{di}(s)$ possono risultare di ordine elevato, rendendo problematico il progetto del controllore. Inoltre le funzioni di trasferimento $\Delta_{12}(s)$ e $\Delta_{21}(s)$ possono risultare non proprie: in questo caso si può limitare l'azione di disaccoppiamento ad una certa banda di frequenza, prevedendo funzioni di trasferimento che approssimino la risposta in frequenza ideale nella banda d'interesse.

Controllo decentralizzato

Se nella struttura di controllo vista precedentemente si pone $\Delta(s)=I$, ovvero si rinuncia all'azione del disaccoppiatore, si ottiene un controllore di funzione di trasferimento $R(s)=R'(s)$ diagonale, che prende il nome di controllore (totalmente) **decentralizzato**. Il comportamento dinamico del controllore è quindi retto dalle equazioni:

$$U_i(s) = R'_i(s)E_i(s), \quad i = 1, \dots, m.$$

Per $m=2$, la struttura del sistema di controllo è la seguente:

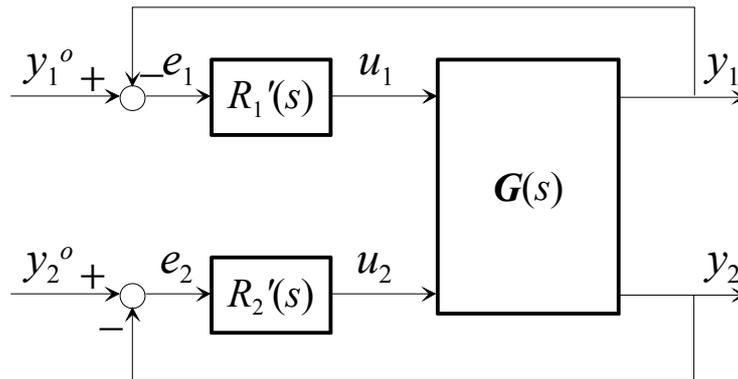


Fig. 24 : Sistema di controllo decentralizzato

A questo punto, si potrebbe condurre il progetto di $R'_i(s)$ sulla base dell' i -simo elemento diagonale $G_{ii}(s)$ della matrice $G(s)$. Tuttavia questa strada non offre particolari garanzie in termini del comportamento del sistema multivariabile in anello chiuso, in quanto le interazioni possono alterarne sensibilmente il comportamento, anche in termini di stabilità.

Un modo migliore di procedere consiste nel progettare le funzioni di trasferimento $R'_i(s)$ **sequenzialmente**, tenendo conto, in ogni progetto, dell'effetto dei regolatori precedentemente inseriti. L'ordine da seguire in questa sequenza si basa di norma sul tempo di risposta ammissibile in anello chiuso: si progettano prima i regolatori destinati alla chiusura degli anelli più veloci.

Nel caso $m=2$, si procederà quindi progettando $R'_1(s)$ sulla base di $G_{11}(s)$ e progettando $R'_2(s)$ sulla base della funzione di trasferimento, che denominiamo $G'_{22}(s)$, che si manifesta da u_2 a y_2 quando è inserito il primo regolatore. Per calcolare questa funzione di trasferimento, si faccia riferimento allo schema a blocchi di Fig. 25, in cui non si è tenuto conto del riferimento y^o_1 in quanto ininfluenza per la discussione. Lo schema corrisponde alle equazioni:

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} -R'_1(s)Y_1(s) \\ U_2(s) \end{bmatrix}.$$

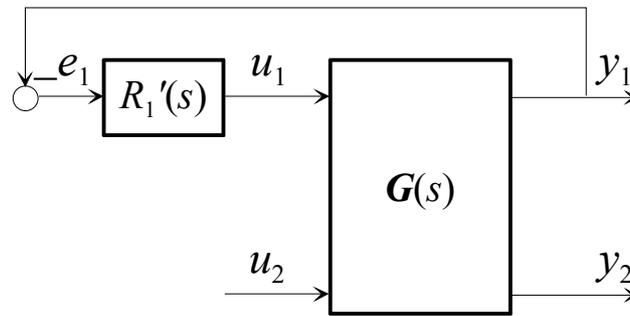


Fig. 25 : Sistema sotto controllo per il secondo regolatore

Eliminando $Y_1(s)$ da queste equazioni, si ottiene il legame tra $U_2(s)$ e $Y_2(s)$, espresso dalla funzione di trasferimento:

$$G'_{22}(s) = G_{22}(s) - \frac{G_{12}(s)G_{21}(s)R'_1(s)}{1 + R'_1(s)G_{11}(s)}.$$

Anche questo metodo, tuttavia, è di natura euristica e non basato su solidi fondamenti teorici.

Si osservi che nel progetto di un controllore decentralizzato si pone il problema di quale sia il migliore **accoppiamento** tra gli ingressi u_i del sistema e le uscite y_j del sistema. Stabilito un certo accoppiamento, e quindi determinata la funzione di trasferimento $\mathbf{G}(s)$, si passa ad un altro accoppiamento permutando opportunamente righe e colonne di $\mathbf{G}(s)$. Ci si chiede come scegliere tra quelli possibili, l'accoppiamento migliore. Un metodo, anch'esso euristico, che guida nella scelta si basa sulla costruzione della cosiddetta **matrice dei guadagni relativi**, avente espressione:

$$\mathbf{\Lambda} = \mathbf{G}(0) \otimes \left(\mathbf{G}(0)^{-1} \right)^T,$$

dove il simbolo \otimes indica il prodotto, elemento per elemento, delle due matrici. Si può dimostrare che l'accoppiamento migliore è quello che rende la matrice $\mathbf{\Lambda}$ più diagonale, con elementi sulla diagonale prossimi a 1.

Per $m=2$, posto:

$$\mathbf{G}(0) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

si ottiene facilmente:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda & 1-\lambda \\ 1-\lambda & \lambda \end{bmatrix},$$

con:

$$\lambda = \frac{g_{11}g_{22}}{g_{11}g_{22} - g_{12}g_{21}}.$$

Si osservi che se $\lambda=1$, $\mathbf{\Lambda}$ è la matrice identità, e risulta $g_{12}=0$ e/o $g_{21}=0$, ovvero il sistema è già diagonale o al più triangolare: l'accoppiamento è quindi corretto.

Se $\lambda=0$, Λ ha elementi nulli sulla diagonale ed unitari fuori dalla diagonale, e risulta $g_{11}=0$ e/o $g_{22}=0$: l'accoppiamento è errato. Accoppiando u_1 con y_2 e u_2 con y_1 si ha l'accoppiamento corretto, che ha struttura al più triangolare.

Argomenti simili si possono usare nei casi intermedi.

Lezione 5

Sistemi a tempo discreto

Introduzione

Un sistema dinamico a tempo discreto è caratterizzato dal fatto che tutte le variabili del sistema sono funzioni di una variabile temporale k che assume solo valori interi.

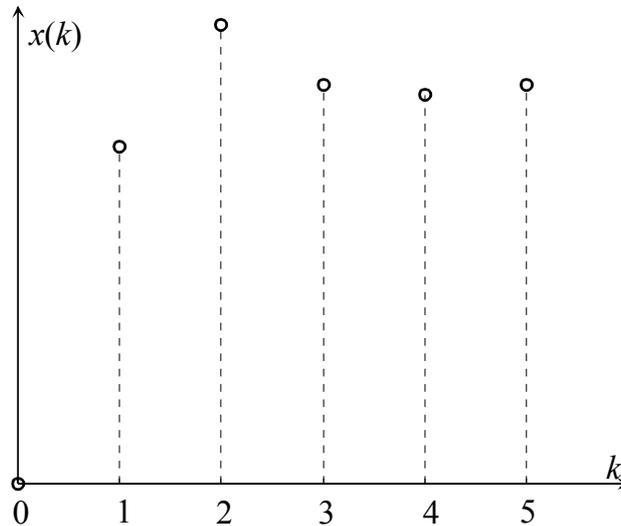


Fig. 1 : Segnale a tempo discreto

La motivazione dello studio dei sistemi a tempo discreto è duplice: da un lato questo studio è utile per la comprensione di alcuni aspetti del controllo digitale (eseguito al calcolatore), dall'altro vi sono sistemi (economici, ecologici, sociologici, ecc.) che si lasciano naturalmente descrivere come sistemi a tempo discreto. Ciò avviene in particolare in tutti i casi in cui i dati disponibili sono nella forma di serie temporali.

Nel seguito ripercorreremo rapidamente l'analisi dei sistemi già svolta a tempo continuo, soffermandoci in particolare sui punti in cui l'analisi dei sistemi a tempo discreto differisce.

Il sistema dinamico

Il sistema dinamico a tempo discreto è caratterizzato da un certo numero (m) di ingressi e un certo numero (p) di uscite.

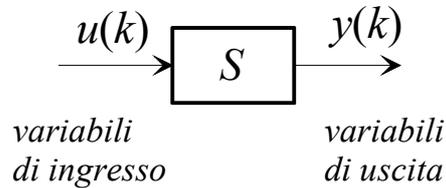


Fig. 2 : Sistema a tempo discreto

Il numero minimo di condizioni iniziali che occorre assegnare per determinare tutte le uscite del sistema, noti gli andamenti degli ingressi a partire dall'istante iniziale, prende il nome di *ordine* del sistema: lo si indica con n .

Il sistema si lascia descrivere per mezzo di n **equazioni alle differenze**, cui si aggiungono p equazioni algebriche per determinare le uscite:

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k), k)$$

$$\mathbf{y}(k) = \mathbf{g}(\mathbf{x}(k), \mathbf{u}(k), k)$$

Si usano le stesse classificazioni viste per i sistemi a tempo continuo: sistemi SISO e MIMO, tempo varianti e invarianti, strettamente propri e no, lineari e non lineari.

In particolare, un sistema lineare tempo invariante (LTI) potrà essere descritto per mezzo di quattro matrici, nella forma:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)$$

Assegnata una condizione iniziale all'istante k_0 ed un ingresso a partire da k_0 , definiamo **movimento** dello stato la soluzione delle equazioni di stato corredate dalla condizione iniziale assegnata e movimento dell'uscita la conseguente uscita, ricavabile dalla trasformazione d'uscita.

L'**equilibrio** è un particolare movimento costante nel tempo a seguito di un ingresso costante nel tempo. Occorre però prestare attenzione al fatto che, nei sistemi tempo invarianti, per determinare gli stati di equilibrio corrispondenti ad un ingresso $\bar{\mathbf{u}}$ si deve imporre che lo stato sia uguale a se stesso in tutti gli istanti, ossia che:

$$\mathbf{x}(k+1) = \mathbf{x}(k) = \bar{\mathbf{x}}$$

Pertanto gli stati di equilibrio sono le soluzioni dell'equazione implicita:

$$\bar{\mathbf{x}} = \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}})$$

In corrispondenza di ogni soluzione di questa equazione, si ha la corrispondente uscita di equilibrio:

$$\bar{\mathbf{y}} = \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}})$$

Si danno infine le stesse definizioni, viste a tempo continuo, di movimento stabile, instabile, asintoticamente stabile.

Esempio 1

Si supponga di dover risolvere numericamente l'equazione scalare:

$$z = f(z),$$

con f generica funzione non lineare. Un metodo per risolvere numericamente l'equazione può consistere nel partire da una certa soluzione iniziale di tentativo x_0 ed iterare secondo la formula:

$$\begin{aligned}x(k+1) &= f(x(k)) \\ x(0) &= x_0\end{aligned}$$

Questa formula costituisce un sistema dinamico non lineare tempo invariante, dove l'indice temporale k scandisce le successive iterazioni dell'algoritmo. Si osservi che gli stati di equilibrio del sistema dinamico sono soluzioni dell'equazione implicita:

$$\bar{x} = f(\bar{x}),$$

ossia sono soluzioni dell'equazione non lineare data. Naturalmente non è detto che si converga alla soluzione dell'equazione partendo da una generica condizione iniziale.

Se ad esempio $f(z) = -z^3$, per cui l'unica soluzione reale dell'equazione è $z=0$, l'algoritmo iterativo è il seguente:

$$x(k+1) = -x(k)^3.$$

Partendo da $x_0 = 1/2$, le prime iterazioni sono:

$$x(1) = -(1/2)^3 = -1/8$$

$$x(2) = -(-1/8)^3 = 1/512$$

$$x(3) = -(1/512)^3 = -1/134217728$$

Esempio 2

Consideriamo un sistema economico in cui definiamo le variabili:

$y(k)$: reddito nazionale nell'anno k ;

$c(k)$: consumi nell'anno k ;

$i(k)$: investimenti privati nell'anno k ;

$u(k)$: spesa pubblica nell'anno k .

Il sistema può essere descritto dalle equazioni:

$$y(k) = c(k) + i(k) + u(k)$$

$$c(k) = \alpha y(k-1)$$

$$i(k) = \beta(c(k) - c(k-1))$$

Possiamo rappresentare queste equazioni in termini di sistema dinamico, introducendo le variabili di stato:

$$x_1(k) = c(k)$$

$$x_2(k) = i(k)$$

Traslando le ultime due equazioni di un passo in avanti, si ottiene:

$$\begin{aligned}x_1(k+1) &= c(k+1) = \alpha y(k) = \alpha(c(k) + i(k) + u(k)) = \\ &= \alpha x_1(k) + \alpha x_2(k) + \alpha u(k) \\ x_2(k+1) &= i(k+1) = \beta(c(k+1) - c(k)) = \beta(\alpha(c(k) + i(k) + u(k)) - c(k)) = \\ &= \beta(\alpha - 1)x_1(k) + \beta\alpha x_2(k) + \beta\alpha u(k) \\ y(k) &= x_1(k) + x_2(k) + u(k)\end{aligned}$$

Il sistema è LTI, con matrici:

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} \alpha & \alpha \\ \beta(\alpha - 1) & \beta\alpha \end{bmatrix}, & \mathbf{B} &= \begin{bmatrix} \alpha \\ \beta\alpha \end{bmatrix} \\ \mathbf{C} &= [1 \quad 1], & D &= 1\end{aligned}$$

Sistemi lineari tempo invarianti

Consideriamo un sistema LTI:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)$$

Proponiamoci di calcolare il movimento dello stato a partire dall'istante $k=0$, in cui $\mathbf{x}(0)=\mathbf{x}_0$. Si ottiene, iterativamente:

$$\mathbf{x}(1) = \mathbf{A}\mathbf{x}(0) + \mathbf{B}\mathbf{u}(0) = \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{u}(0)$$

$$\mathbf{x}(2) = \mathbf{A}\mathbf{x}(1) + \mathbf{B}\mathbf{u}(1) = \mathbf{A}^2\mathbf{x}_0 + \mathbf{A}\mathbf{B}\mathbf{u}(0) + \mathbf{B}\mathbf{u}(1)$$

$$\mathbf{x}(3) = \mathbf{A}\mathbf{x}(2) + \mathbf{B}\mathbf{u}(2) = \mathbf{A}^3\mathbf{x}_0 + \mathbf{A}^2\mathbf{B}\mathbf{u}(0) + \mathbf{A}\mathbf{B}\mathbf{u}(1) + \mathbf{B}\mathbf{u}(2)$$

⋮

Per induzione, possiamo ricavare la formula generale del movimento dello stato:

$$\mathbf{x}(k) = \mathbf{A}^k \mathbf{x}_0 + \sum_{i=0}^{k-1} [\mathbf{A}^{k-i-1} \mathbf{B}\mathbf{u}(i)],$$

cui corrisponde la formula per il movimento dell'uscita:

$$\mathbf{y}(k) = \mathbf{C}\mathbf{A}^k \mathbf{x}_0 + \sum_{i=0}^{k-1} [\mathbf{C}\mathbf{A}^{k-i-1} \mathbf{B}\mathbf{u}(i)] + \mathbf{D}\mathbf{u}(k).$$

Come a tempo continuo, il movimento può quindi essere scomposto in **movimento libero e movimento forzato**: il primo dipende solo dallo stato iniziale, il secondo solo dall'ingresso.

$$\left. \begin{aligned} \mathbf{x}_l(k) &= \mathbf{A}^k \mathbf{x}_0 \\ \mathbf{y}_l(k) &= \mathbf{C}\mathbf{A}^k \mathbf{x}_0 \end{aligned} \right\} \text{ moto libero}$$

$$\left. \begin{aligned} \mathbf{x}_f(k) &= \sum_{i=0}^{k-1} [\mathbf{A}^{k-i-1} \mathbf{B}\mathbf{u}(i)] \\ \mathbf{y}_f(k) &= \sum_{i=0}^{k-1} [\mathbf{C}\mathbf{A}^{k-i-1} \mathbf{B}\mathbf{u}(i)] + \mathbf{D}\mathbf{u}(k) \end{aligned} \right\} \text{ moto forzato}$$

Poiché il moto libero è lineare nello stato iniziale, ed il moto forzato lo è nell'ingresso, vale il **principio di sovrapposizione degli effetti** (di fatto valido anche per sistema tempo variante, purché lineare).

Per quanto riguarda gli equilibri, per quanto già osservato, essi risolvono l'equazione implicita:

$$\bar{\mathbf{x}} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{B}\bar{\mathbf{u}}$$

dove $\bar{\mathbf{u}}$ è un ingresso costante. A ciascun eventuale stato di equilibrio è associata l'uscita di equilibrio:

$$\bar{\mathbf{y}} = \mathbf{C}\bar{\mathbf{x}} + \mathbf{D}\bar{\mathbf{u}}.$$

Se la matrice $\mathbf{I}-\mathbf{A}$ è invertibile, il che avviene se \mathbf{A} non ha autovalori in $s=1$, esiste un solo stato di equilibrio, dato dall'espressione:

$$\bar{\mathbf{x}} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}\bar{\mathbf{u}}$$

Inoltre risulta:

$$\bar{y} = \mu \bar{u}$$

con:

$$\mu = \mathbf{C}(\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}$$

guadagno statico del sistema.

E' utile ricordare che, effettuando un cambiamento di variabili di stato:

$$\hat{\mathbf{x}}(k) = \mathbf{T}\mathbf{x}(k), \quad \det \mathbf{T} \neq 0$$

le matrici del sistema si trasformano esattamente come a tempo continuo:

$$\hat{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \quad \hat{\mathbf{B}} = \mathbf{T}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{T}^{-1}, \quad \hat{\mathbf{D}} = \mathbf{D}.$$

Inoltre le proprietà di raggiungibilità ed osservabilità si definiscono come a tempo continuo, ed i test per verificarle sono identici.

Sistemi LTI: stabilità

Per i sistemi LTI a tempo discreto valgono considerazioni sulla stabilità del tutto analoghe a quelle fatte a tempo continuo. In particolare quindi la stabilità è una proprietà del sistema (tutti i movimenti sono asintoticamente stabili, stabili o instabili). Inoltre la stabilità si può valutare studiando le soluzioni dell'equazione libera in $\delta\mathbf{x}$ (differenza tra movimento perturbato e movimento nominale):

$$\delta\mathbf{x}(k+1) = \mathbf{A}\delta\mathbf{x}(k)$$

al variare della condizione iniziale $\delta\mathbf{x}(0)$. Risulta:

$$\delta\mathbf{x}(k) = \mathbf{A}^k \delta\mathbf{x}(0).$$

Se tutte le componenti del moto libero sono limitate, qualunque sia lo stato iniziale, il sistema è stabile; se inoltre decadono tutte a zero, il sistema è asintoticamente stabile; se per almeno uno stato iniziale almeno una componente del moto libero non decade a zero il sistema è instabile.

Se la matrice \mathbf{A} è diagonalizzabile, cioè se:

$$\exists \mathbf{T} : \mathbf{T}\mathbf{A}\mathbf{T}^{-1} = \hat{\mathbf{A}} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

dove λ_i sono gli autovalori di \mathbf{A} , risulta:

$$\delta\mathbf{x}(k) = (\mathbf{T}^{-1} \hat{\mathbf{A}} \mathbf{T})^k \delta\mathbf{x}(0) = \mathbf{T}^{-1} \hat{\mathbf{A}}^k \mathbf{T} \delta\mathbf{x}(0) = \mathbf{T}^{-1} \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^k \end{bmatrix} \mathbf{T} \delta\mathbf{x}(0).$$

Le componenti del moto libero del sistema sono quindi combinazioni lineari degli esponenziali degli autovalori. Di seguito sono riportati gli andamenti di λ^k al variare di λ reale:

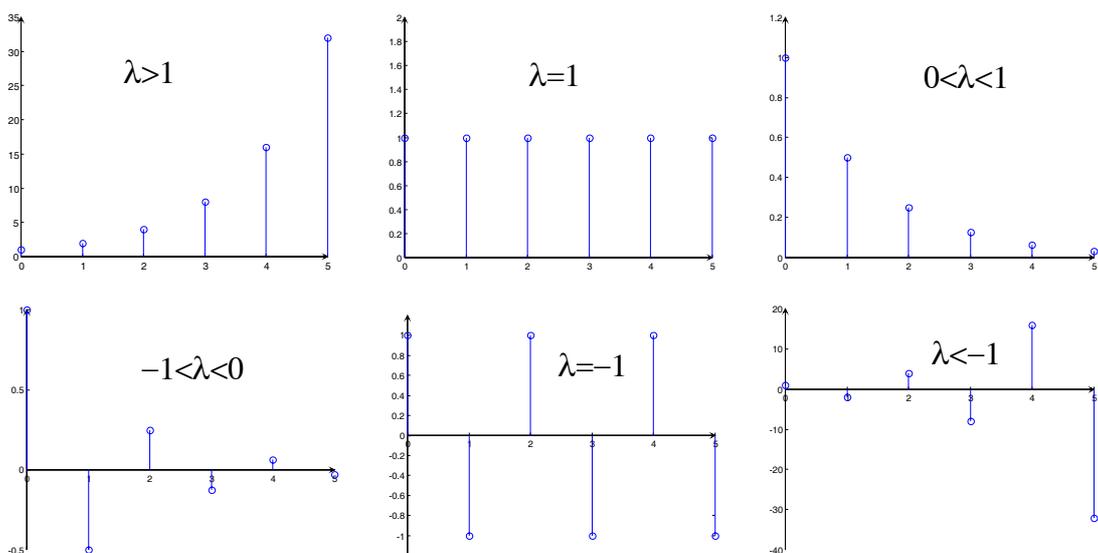


Fig. 3 : Esponenziale a tempo discreto

Naturalmente, accanto ad un autovalore complesso $\lambda_i = \rho_i e^{j\theta_i}$ ci sarà anche il coniugato e la combinazione dei due termini darà luogo ad un termine reale del tipo $\rho_i^k \cos(\theta_i k + \varphi_i)$, con φ_i fase opportuna.

E' facile allora rendersi conto che se tutti gli autovalori hanno modulo minore di 1, tutti i moti liberi sono limitati e decadono a zero; se non ci sono autovalori a modulo maggiore di 1, ma ce ne sono a modulo unitario, nessun moto libero diverge, ma vi sono moti liberi che non decadono a zero; se c'è almeno un autovalore a modulo maggiore di 1, almeno un moto libero non è limitato.

Possiamo pertanto concludere che un sistema dinamico lineare tempo invariante a tempo discreto è:

- asintoticamente stabile:** se e solo se tutti gli autovalori di A hanno modulo minore di 1;
- stabile:** se e solo se tutti gli autovalori di A hanno modulo minore o uguale a 1 e ne esistono a modulo uguale a 1;
- instabile:** se e solo se esistono autovalori di A a modulo maggiore di 1.

Osservazioni

- 1) In pratica si tratta di parafrasare le condizioni di stabilità valide per un sistema a tempo continuo, sostituendo al semipiano sinistro del piano complesso il cerchio di centro l'origine e raggio unitario (insieme dei numeri complessi che hanno modulo minore di 1).
- 2) Ricordiamo che a seguito di un cambiamento di variabili di stato descritto da una matrice di trasformazione T , la matrice A del sistema si trasforma secondo una relazione di similitudine ($\hat{A} = TAT^{-1}$). Poiché matrici simili hanno gli stessi autovalori, l'analisi della stabilità è del tutto indipendente dalla scelta delle variabili di stato. In altre parole, la proprietà di stabilità è una **proprietà strutturale** del sistema dinamico.
- 3) Se la matrice A non è diagonalizzabile, può essere messa in relazione di similitudine con una forma canonica (forma di Jordan). Seguendo questa strada si giunge alla conclusione che il teorema precedentemente enunciato va corredato dalla precisazione che se vi sono autovalori multipli a modulo unitario (e non vi sono autovalori a modulo maggiore di 1), il sistema è instabile se per almeno uno degli autovalori a modulo unitario la cosiddetta *molteplicità geometrica* (numero degli autovettori linearmente indipendenti associati all'autovalore) è inferiore alla *molteplicità algebrica* (molteplicità con cui l'autovalore è radice del polinomio caratteristico).

Esempio

Si consideri nuovamente il sistema economico, in cui:

$$A = \begin{bmatrix} \alpha & \alpha \\ \beta(\alpha - 1) & \beta\alpha \end{bmatrix}.$$

Posto $\alpha=0.5$, $\beta=1$, si ha:

$$A = \begin{bmatrix} 0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix}$$

Gli autovalori sono le radici del polinomio caratteristico:

$$\chi_A(\lambda) = \begin{vmatrix} \lambda - 0.5 & -0.5 \\ -0.5 & \lambda - 0.5 \end{vmatrix} = (\lambda - 0.5)^2 + 0.25 = \lambda^2 - \lambda + 0.5$$
$$\lambda = \frac{1 \pm \sqrt{1-2}}{2} = \frac{1 \pm j}{2}$$

e sono disposti nel piano complesso come in figura:

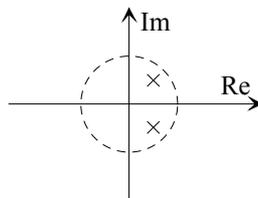


Fig. 4 : Autovalori per l'esempio economico

Poiché gli autovalori hanno entrambi modulo minore di 1, il sistema è asintoticamente stabile.

Trasformazione bilineare

Anche per i sistemi a tempo discreto è possibile studiare la stabilità evitando il calcolo diretto degli autovalori, ma studiando i coefficienti del polinomio caratteristico della matrice A :

$$\varphi(z) = \varphi_0 z^n + \varphi_1 z^{n-1} + \varphi_2 z^{n-2} + \dots + \varphi_n.$$

Esiste un criterio, dovuto a Jury, per determinare le condizioni sui coefficienti φ_i necessarie e sufficienti perché il polinomio ammetta tutte radici a modulo minore di 1. Tuttavia è anche possibile ricondursi all'impiego del criterio noto a tempo continuo, il criterio di Routh, per verificare se un polinomio ha tutte le radici a parte reale negativa. A questo scopo, occorre individuare una trasformazione di variabili $z = \psi(s)$, tale che, sostituita nell'equazione $\varphi(z) = 0$, dia luogo ancora ad un'equazione polinomiale, in s , le cui radici abbiano tutte parte reale negativa se e solo se l'equazione $\varphi(z) = 0$ ha tutte le radici a modulo minore di 1.

Una trasformazione utilizzabile a questo proposito è la **trasformazione bilineare**:

$$z = \frac{1+s}{1-s}.$$

Ovviamente questa trasformazione, essendo razionale, dà luogo ad un'equazione polinomiale. Inoltre è facile verificare che il modulo di z è minore di 1 se e solo se la parte reale di s è negativa. Posto infatti $s = x + jy$, si ha:

$$|z|^2 = \frac{|1+x+jy|^2}{|1-x-jy|^2} = \frac{x^2 + y^2 + 2x + 1}{x^2 + y^2 - 2x + 1} < 1 \Leftrightarrow x < 0.$$

Una volta determinato il polinomio in s , si procede con il criterio di Routh per l'analisi di stabilità.

Esempio

Si consideri il polinomio di terzo grado:

$$\varphi(z) = 8z^3 - 12z^2 + 6z - 1.$$

Applicando la trasformazione bilineare ed uguagliando a zero si ottiene:

$$8\left(\frac{1+s}{1-s}\right)^3 - 12\left(\frac{1+s}{1-s}\right)^2 + 6\left(\frac{1+s}{1-s}\right) - 1 = 0.$$

Prendendo il denominatore comune ed uguagliando a zero il polinomio a numeratore:

$$8(1+s)^3 - 12(1+s)^2(1-s) + 6(1+s)(1-s)^2 - (1-s)^3 = 0,$$

ossia:

$$27s^3 + 27s^2 + 9s + 1 = 0.$$

Costruiamo la tabella di Routh:

27 9 0

27 1 0

8 0

1

Poiché tutti gli elementi della prima colonna sono positivi, il polinomio in s ha tutte le radici a parte reale negativa. Questo comporta che il polinomio in z ha tutte le radici a modulo minore di 1 e quindi che il sistema è asintoticamente stabile. Si osservi che, in effetti:

$$\varphi(z) = 8(z - 0.5)^3,$$

ovvero che le tre radici del polinomio in z coincidono nel punto 0.5.

Stabilità degli stati di equilibrio

Anche a tempo discreto lo studio dei sistemi lineari risulta utile per lo studio locale dei sistemi non lineari nell'intorno di stati di equilibrio.

Si consideri un sistema non lineare tempo invariante:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k)) \\ \mathbf{y}(k) &= \mathbf{g}(\mathbf{x}(k), \mathbf{u}(k)) \end{aligned}$$

e si supponga che, in corrispondenza ad un ingresso costante $\bar{\mathbf{u}}$ si abbia lo stato di equilibrio $\bar{\mathbf{x}}$ e l'uscita di equilibrio $\bar{\mathbf{y}}$, per cui:

$$\bar{\mathbf{x}} = \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}})$$

$$\bar{\mathbf{y}} = \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}}).$$

Introducendo le variabili:

$$\delta \mathbf{u}(k) = \mathbf{u}(k) - \bar{\mathbf{u}}, \quad \delta \mathbf{x}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}, \quad \delta \mathbf{y}(k) = \mathbf{y}(k) - \bar{\mathbf{y}},$$

si formula il sistema linearizzato:

$$\begin{aligned} \delta \mathbf{x}(k+1) &= \mathbf{A} \delta \mathbf{x}(k) + \mathbf{B} \delta \mathbf{u}(k) \\ \delta \mathbf{y}(k) &= \mathbf{C} \delta \mathbf{x}(k) + \mathbf{D} \delta \mathbf{u}(k) \end{aligned}$$

con:

$$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}}, \quad \mathbf{B} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}}, \quad \mathbf{C} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}}, \quad \mathbf{D} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}}$$

Il sistema linearizzato risulta in particolare utile per lo studio della stabilità dello stato di equilibrio del sistema non lineare Valgono infatti i seguenti risultati:

- 1) Se la matrice \mathbf{A} ha tutti autovalori a modulo minore di 1 (ossia se il sistema linearizzato è asintoticamente stabile) lo stato di equilibrio $\bar{\mathbf{x}}$ è **asintoticamente stabile**.
- 2) Se la matrice \mathbf{A} ha almeno un autovalore a modulo maggiore di 1 lo stato di equilibrio $\bar{\mathbf{x}}$ è **instabile**.

Resta come caso indecidibile sulla base dell'analisi del sistema linearizzato quello in cui la matrice \mathbf{A} non ha autovalori a modulo maggiore di 1 ma ne ha a modulo uguale a 1. In questo caso occorrono approssimazioni del sistema originario non lineare estese a termini di ordine superiore al primo per decidere circa la stabilità dello stato di equilibrio.

Esempio

Consideriamo il sistema a tempo discreto del primo ordine:

$$x(k+1) = f(x(k), u(k)) = \cos(x(k)) + u(k).$$

Siamo interessati allo studio di eventuali equilibri che si hanno quando l'ingresso è nullo: $u(k) = \bar{u} = 0$. Per trovare gli stati di equilibrio, imponiamo la condizione:

$$\bar{x} = f(\bar{x}, \bar{u}) = f(\bar{x}, 0) = \cos(\bar{x}).$$

Graficamente si trova una sola soluzione:

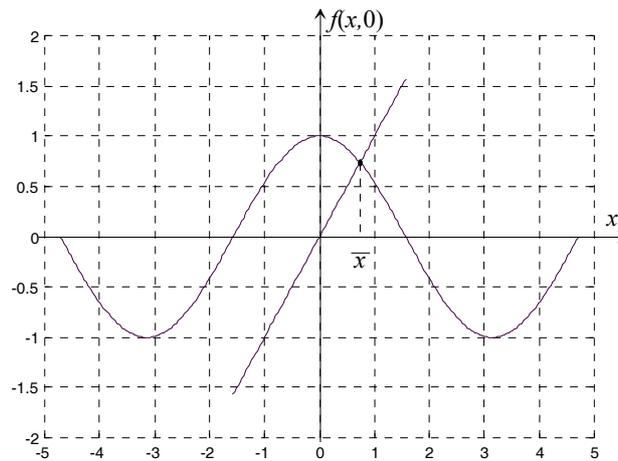


Fig. 5 : Ricerca grafica dello stato di equilibrio

Risulta $\bar{x} \approx 0.74$. L'unico autovalore del sistema linearizzato si ottiene come:

$$A = \left. \frac{\partial f}{\partial x} \right|_{\bar{x}, \bar{u}} = -\sin(\bar{x}) = -\sqrt{1 - \cos(\bar{x})^2} = -\sqrt{1 - \bar{x}^2} = -0.67.$$

Poiché l'autovalore è in modulo minore di 1, lo stato di equilibrio è asintoticamente stabile.

Trasformata Zeta

Dato un sistema LTI, in analogia con il procedimento seguito a tempo continuo, considereremo una rappresentazione alternativa del sistema, ottenuta introducendo i vettori $U(z)$ e $Y(z)$, rispettivamente vettori delle trasformate Zeta degli ingressi e delle uscite del sistema dinamico. Anche nel caso discreto il pregio dell'approccio nelle trasformate consiste nel fatto che, come si dimostrerà, il legame tra la trasformata dell'ingresso e la trasformata dell'uscita è espresso da equazioni algebriche.

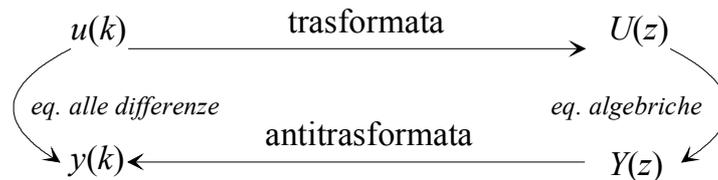


Fig. 6 : Approccio nel dominio delle trasformate

Si consideri dunque una generica funzione reale $v(k)$, definita per k intero ≥ 0 , tale che, per almeno un $r \in \mathfrak{R}$, $r > 0$, risulti:

$$\sum_{k=0}^{\infty} |v(k)| r^{-k} < \infty .$$

La serie:

$$V(z) = \sum_{k=0}^{\infty} v(k) z^{-k}$$

converge per valori della variabile complessa z tali che $|z| > r$, ovvero nel cosiddetto co-cerchio di convergenza. Definiamo **trasformata Zeta** di v l'unica funzione di z , definita ed analitica quasi ovunque nel piano complesso, che nel co-cerchio di convergenza della serie, coincide con la somma $V(z)$ della serie stessa.

Esempi di calcolo diretto della trasformata

Impulso

Consideriamo l'impulso unitario a tempo discreto (delta di Kronecker):

$$v(k) = \text{imp}(k) = \delta_0(k) = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

risulta:

$$V(z) = \sum_{k=0}^{\infty} v(k) z^{-k} = v(0) + v(1)z^{-1} + v(2)z^{-2} + \dots = 1 + 0 + 0 + \dots = 1$$

Esponenziale

Consideriamo l'esponenziale a tempo discreto $v(k) = a^k$. Risulta:

$$V(z) = \sum_{k=0}^{\infty} a^k z^{-k} = \sum_{k=0}^{\infty} (az^{-1})^k = \frac{1}{1-az^{-1}} = \frac{z}{z-a}, \quad \text{per } |az^{-1}| < 1 \quad (|z| > |a|)$$

Per $a=1$ si ha lo scalino a tempo discreto:

$$v(k) = \text{sca}(k) \Rightarrow V(z) = \frac{z}{z-1}$$

Proprietà della trasformata Zeta

Linearità

$$v(k) = \alpha v_1(k) + \beta v_2(k) \Rightarrow V(z) = \alpha V_1(z) + \beta V_2(z)$$

Anticipi e ritardi

$$v_2(k) = v_1(k+1) \Rightarrow V_2(z) = z(V_1(z) - v_1(0))$$

$$v_2(k) = v_1(k-1) \Rightarrow V_2(z) = z^{-1}V_1(z)$$

Derivazione in z

$$v_2(k) = kv_1(k) \Rightarrow V_2(z) = -z \frac{dV_1(z)}{dz}$$

Valore iniziale

$$v(0) = \lim_{z \rightarrow \infty} V(z)$$

Valore finale (applicabile se i poli di V sono a modulo < 1 o in $z=1$)

$$v(\infty) = \lim_{z \rightarrow 1} [(z-1)V(z)]$$

Esempi

1) Consideriamo la rampa a tempo discreto:

$$\text{ram}(k) = k, \quad k \geq 0$$

Poiché $\text{ram}(k) = k \text{sca}(k)$, si ha:

$$Z[\text{ram}(k)] = -z \frac{d}{dz} Z[\text{sca}(k)] = -z \frac{d}{dz} \left[\frac{z}{z-1} \right] = \frac{z}{(z-1)^2}$$

2) Consideriamo un segnale di trasformata

$$V(z) = \frac{z}{z-a}$$

Dai teoremi del valore iniziale e finale:

$$v(0) = \lim_{z \rightarrow \infty} V(z) = 1$$

$$v(\infty) = \lim_{z \rightarrow 1} [(z-1)V(z)] = \lim_{z \rightarrow 1} \left[(z-1) \frac{z}{z-a} \right] = 0 \Leftrightarrow |a| < 1$$

Se $a=1$, il teorema del valore finale è ancora applicabile e risulta:

$$v(\infty) = \lim_{z \rightarrow 1} [(z-1)V(z)] = \lim_{z \rightarrow 1} \left[(z-1) \frac{z}{z-1} \right] = 1.$$

Tutti questi risultati sono coerenti con il fatto che $v(k) = a^k$

Trasformate notevoli

Utilizzando le proprietà della trasformata, si può compilare la seguente tabella di trasformate notevoli:

$v(k)$	$V(z)$
imp(k)	1
sca(k)	$\frac{z}{z-1}$
ram(k)	$\frac{z}{(z-1)^2}$
par(k)	$\frac{z}{(z-1)^3}$
a^k	$\frac{z}{z-a}$
ka^k	$\frac{az}{(z-a)^2}$

dove par(k) = $k(k-1)/2$, $k \geq 0$.

Antitrasformata Zeta

Come la trasformata di Laplace, anche la trasformata Zeta è biunivoca: data una trasformata, è sempre possibile risalire univocamente alla funzione del tempo che la genera.

Per trasformate Zeta razionali (rapporti di polinomi), si può utilizzare per l'antitrasformata il metodo di **Heaviside**, ossia di scomposizione in frazioni semplici. Di fatto conviene scomporre $V(z)/z$, secondo il seguente schema (per poli semplici):

$$\frac{V(z)}{z} = \frac{\alpha_0}{z} + \frac{\alpha_1}{z-p_1} + \dots + \frac{\alpha_n}{z-p_n}$$

$$V(z) = \alpha_0 + \alpha_1 \frac{z}{z-p_1} + \dots + \alpha_n \frac{z}{z-p_n}$$

$$v(k) = \alpha_0 \text{imp}(k) + \alpha_1 p_1^k + \dots + \alpha_n p_n^k, \quad k \geq 0$$

In alternativa si può usare il metodo della **lunga divisione**, che consiste nel dividere il polinomio a numeratore e quello a denominatore, in modo da trovare, per confronto tra il risultato parziale della divisione e l'espansione della serie che costituisce la trasformata Zeta, i primi campioni dell'antitrasformata:

$$V(z) = \frac{N(z)}{D(z)} = \beta_0 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots \Rightarrow \begin{cases} v(0) = \beta_0 \\ v(1) = \beta_1 \\ v(2) = \beta_2 \end{cases}$$

Esempio

Consideriamo la trasformata Zeta:

$$V(z) = \frac{3z+12}{z^2+5z+6}$$

Con il metodo di Heaviside, si ottiene:

$$\frac{V(z)}{z} = \frac{3z+12}{z(z+2)(z+3)} = \frac{\alpha_0}{z} + \frac{\alpha_1}{z+2} + \frac{\alpha_2}{z+3} = \frac{\alpha_0(z+2)(z+3) + \alpha_1 z(z+3) + \alpha_2 z(z+2)}{z(z+2)(z+3)}$$

Valutando il polinomio a numeratore in $z=0$, $z=-2$, $z=-3$, si ottiene:

$$\begin{cases} 6\alpha_0 = 12 \\ -2\alpha_1 = 6 \\ 3\alpha_2 = 3 \end{cases} \Rightarrow \begin{cases} \alpha_0 = 2 \\ \alpha_1 = -3 \\ \alpha_2 = 1 \end{cases}$$

Pertanto la trasformata Zeta data si scompone nella somma:

$$V(z) = 2 - 3 \frac{z}{z+2} + \frac{z}{z+3}$$

immediatamente antitrasformabile in:

$$v(k) = 2\text{imp}(k) - 3(-2)^k + (-3)^k, \quad k \geq 0.$$

Con il metodo della lunga divisione si ottiene invece:

$$\begin{array}{r|l}
 3z + 12 & z^2 + 5z + 6 \\
 3z + 15 + 18z^{-1} & 3z^{-1} - 3z^{-2} - 3z^{-3} \\
 \hline
 -3 - 18z^{-1} & \\
 -3 - 15z^{-1} - 18z^{-2} & \\
 \hline
 -3z^{-1} + 18z^{-2} &
 \end{array}
 \Rightarrow
 \begin{array}{l}
 v(0) = 0 \\
 v(1) = 3 \\
 v(2) = -3 \\
 v(3) = -3
 \end{array}$$

Funzione di trasferimento

Si consideri un sistema LTI :

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k)$$

$$y(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k)$$

Applichiamo la trasformata Zeta ad entrambi i membri delle equazioni, supponendo lo stato iniziale nullo ($\mathbf{x}(0)=0$):

$$\begin{aligned} z\mathbf{X}(z) &= \mathbf{A}\mathbf{X}(z) + \mathbf{B}U(z) & \Rightarrow & \quad \mathbf{X}(z) = (z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}U(z) \\ Y(z) &= \mathbf{C}\mathbf{X}(z) + \mathbf{D}U(z) & & \quad Y(z) = [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}]U(z) \end{aligned}$$

Si è ottenuto:

$$Y(z) = \mathbf{G}(z)U(z), \quad \mathbf{G}(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}$$

La matrice a p righe e m colonne $\mathbf{G}(z)$ prende il nome di **funzione di trasferimento** del sistema e dà la trasformata dell'uscita forzata dall'ingresso. Nel caso SISO possiamo dunque scrivere:

$$G(z) = \frac{Y(z)}{U(z)}$$

La funzione di trasferimento a tempo discreto ha formalmente la stessa espressione di quella a tempo continuo. Pertanto gode delle stesse proprietà:

- La f.d.t. è invariante rispetto a cambiamenti di variabili di stato.
- Nel caso SISO la f.d.t. è rapporto di due polinomi:

$$G(z) = \frac{N(z)}{D(z)}$$

- Le radici del polinomio a numeratore si chiamano **zeri**, le radici del polinomio a denominatore **poli**
- A meno di cancellazioni, i poli coincidono con gli autovalori della matrice \mathbf{A} , sicché l'analisi di stabilità del sistema può essere condotta anche analizzando i poli della funzione di trasferimento.

Facendo ora riferimento ad un sistema SISO, si introduce il concetto di **tipo g** della funzione di trasferimento, numero intero positivo, nullo o negativo: se positivo, la funzione di trasferimento presenta g poli in $z=1$, se negativo, vi sono $(-g)$ zeri in $z=1$, se nullo non vi sono né zeri, né poli in $z=1$.

Consideriamo una funzione di trasferimento priva di poli o zeri in $z=1$ (ovvero di tipo nullo). Definiamo **guadagno** della funzione di trasferimento il valore che assume per $z=1$:

$$\mu = G(1) = \mathbf{C}(\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}$$

Il guadagno della funzione di trasferimento coincide quindi con il guadagno statico del sistema, ossia con il rapporto uscita/ingresso all'equilibrio.

Si supponga inoltre il sistema **asintoticamente stabile** e lo si solleciti con un ingresso a scalino:

$$u(k) = \text{sca}(k) \Rightarrow U(z) = \frac{z}{z-1}$$

Risulta:

$$\lim_{k \rightarrow \infty} y(k) = \lim_{z \rightarrow 1} [(z-1)Y(z)] = \lim_{z \rightarrow 1} \left[(z-1)G(z) \frac{z}{z-1} \right] = G(1) = \mu$$

Pertanto il guadagno della funzione di trasferimento è il valore di regime della risposta allo scalino del sistema.

In presenza di poli o zeri in $z=1$ (funzione di trasferimento di tipo non nullo), la nozione di guadagno si generalizza come:

$$\mu = \lim_{z \rightarrow 1} [(z-1)^g G(z)].$$

A differenza di quanto avviene a tempo continuo, anche il **ritardo di tempo** ha funzione di trasferimento razionale. Se infatti risulta:

$$y(k) = u(k-h),$$

con h intero positivo, applicando iterativamente la regola sul ritardo di un passo della trasformata Zeta, si ha:

$$Y(z) = z^{-h}U(z).$$

Pertanto la funzione di trasferimento è:

$$G(z) = \frac{Y(z)}{U(z)} = z^{-h} = \frac{1}{z^h}.$$

Si tratta quindi di un sistema di guadagno 1, con h poli nell'origine.

Risposte temporali

Le risposte temporali si calcolano con lo stesso procedimento utilizzato a tempo continuo:

0. Calcolo della funzione di trasferimento del sistema
1. Trasformazione Zeta dell'ingresso
2. Calcolo della trasformata Zeta dell'uscita ($Y(z)=G(z)U(z)$)
3. Antitrasformazione Zeta dell'uscita

A titolo d'esempio si consideri il sistema del primo ordine:

$$G(z) = \mu \frac{1-p}{z-p}$$

Calcoliamo la risposta allo scalino:

$$u(k) = \text{sca}(k) \Rightarrow U(z) = \frac{z}{z-1}$$

La trasformata dell'uscita è:

$$Y(z) = \mu \frac{1-p}{z-p} \frac{z}{z-1} = \mu \left(\frac{z}{z-1} - \frac{z}{z-p} \right)$$

Antitrasformando:

$$y(k) = \mu(1-p^k), \quad k \geq 0.$$

Se $|p| < 1$ il sistema è asintoticamente stabile e la risposta allo scalino converge al valore $\mu=G(1)$. Tuttavia, se $p > 0$, la risposta è monotona, se $p < 0$ è oscillante:

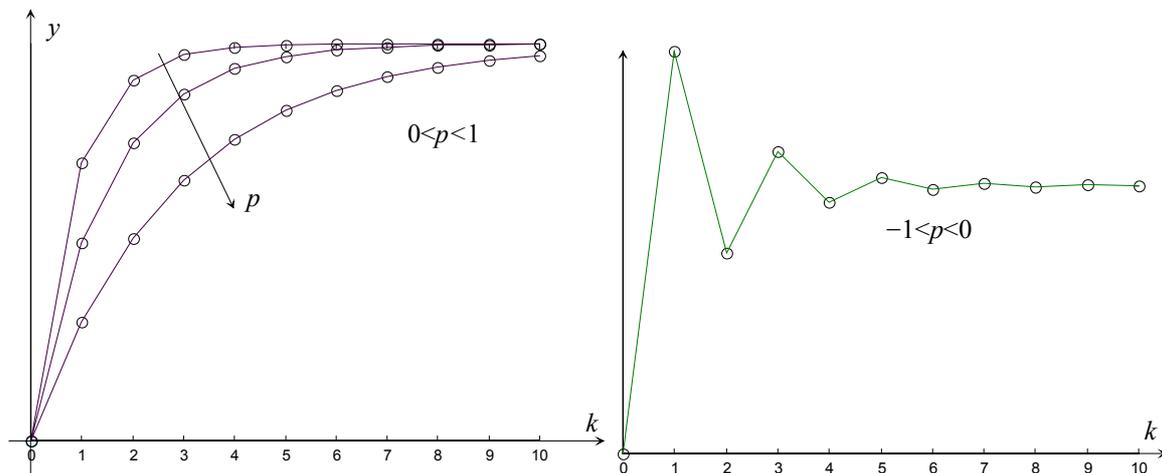


Fig. 7 : Risposta allo scalino di un sistema del primo ordine

Si osservi che, contrariamente ai sistemi a tempo continuo, anche un sistema del primo ordine, con polo compreso tra -1 e 0 , può dare luogo ad una risposta allo scalino oscillante.

Risposta in frequenza

In un sistema LTI **asintoticamente stabile**, sollecitato dall'ingresso:

$$u(k) = U \sin(\bar{\theta}k + \phi)$$

esaurito un transitorio iniziale, l'uscita assume l'espressione:

$$y(k) = Y \sin(\bar{\theta}k + \psi)$$

con:

$$Y = U \left| G(e^{j\bar{\theta}}) \right|$$

$$\psi = \phi + \angle G(e^{j\bar{\theta}})$$

Anche a tempo discreto è possibile estendere questo risultato a segnali in ingresso più generali, purché trasformabili con Fourier.

La funzione complessa della variabile reale θ definita da:

$$G(e^{j\theta}), \quad \theta \in [0, \pi]$$

prende il nome di **risposta in frequenza** del sistema, e si definisce per qualsiasi sistema lineare tempo invariante, indipendentemente dalla sua stabilità.

La risposta in frequenza si ottiene quindi valutando la funzione di trasferimento sulla semicirconferenza superiore di centro l'origine e raggio unitario:

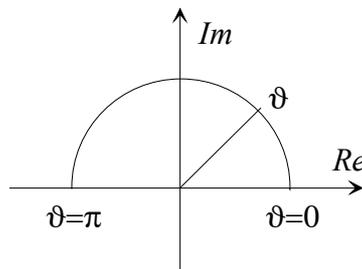


Fig. 8 : Luogo dei punti per il calcolo della risposta in frequenza

La risposta in frequenza costituisce uno strumento completo per l'analisi del sistema dinamico. Tuttavia il suo utilizzo è considerevolmente limitato dal fatto che il tracciamento dei diagrammi della risposta in frequenza (di Bode) non è agevole e non se ne danno approssimazioni asintotiche.

Lezione 6

Controllo digitale

Realizzazione dei regolatori

Come è noto, il progetto del regolatore si conclude con la determinazione di una funzione di trasferimento $R(s)$. Tuttavia la funzione di trasferimento costituisce soltanto un formalismo matematico per descrivere il comportamento dinamico di un sistema, a partire dal quale occorre **realizzare** un dispositivo avente quel comportamento dinamico, ed atto alla regolazione del processo sotto controllo.

Il modo più naturale di procedere sembra quello di realizzare, in una qualunque tecnologia (elettronica, pneumatica, fluidica) un sistema che presenti la stessa funzione di trasferimento ottenuta dal progetto, ed interfacciarlo con il trasduttore della misura della variabile controllata, da un lato, e con l'attuatore, dall'altro. Questo è il principio che sta alla base dei **sistemi di controllo analogici**.

Tuttavia, negli ultimi decenni, l'avvento ed il successivo sviluppo delle tecnologie digitali hanno indotto i progettisti dei sistemi di controllo ad un crescente interesse verso l'utilizzo dei calcolatori, in particolare dei microprocessori, all'interno dell'anello di controllo. Nei **sistemi di controllo digitale** il calcolatore esegue un algoritmo che elabora le informazioni provenienti dai trasduttori e produce un conseguente comando all'attuatore. Dopo le iniziali difficoltà legate principalmente alla, giustificata, riluttanza del mondo industriale ad abbandonare soluzioni tecnologicamente assestate a favore di altre basate su tecnologie emergenti, la tecnologia di controllo digitale ha via via soppiantato i sistemi di controllo analogici, che oggi sopravvivono soprattutto in applicazioni in cui le bande richieste agli anelli di controllo renderebbero antieconomico il passaggio alla tecnologia digitale (esempi si hanno nel campo del controllo dei motori elettrici).

Inserimento del calcolatore in un anello di controllo

Un calcolatore opera, in istanti di tempo discreti, su variabili rappresentate da numeri con precisione finita (dipendente dal numero di bit), ossia su cosiddetti **segnali digitali**. Le variabili di ingresso ed uscita del sistema sotto controllo sono invece funzioni del tempo (continuo), sono cioè cosiddetti **segnali analogici**. Sono segnali analogici anche le misure delle grandezze fornite dai trasduttori, come pure i segnali di comando degli attuatori. L'inserimento del calcolatore nell'anello di controllo comporta quindi l'adozione di dispositivi per la conversione dei segnali analogici in segnali digitali e viceversa.

Tali dispositivi prendono il nome, rispettivamente, di:

convertitori A/D (analogico/digitale)

convertitori D/A (digitale/analogico).

Uno schema di massima di un sistema di controllo digitale sarà quindi il seguente (dove C indica il calcolatore, A , S e T indicano rispettivamente l'attuatore, il sistema sotto controllo ed il trasduttore, mentre l'asterisco e la freccia a tratto più spesso distinguono i segnali digitali):

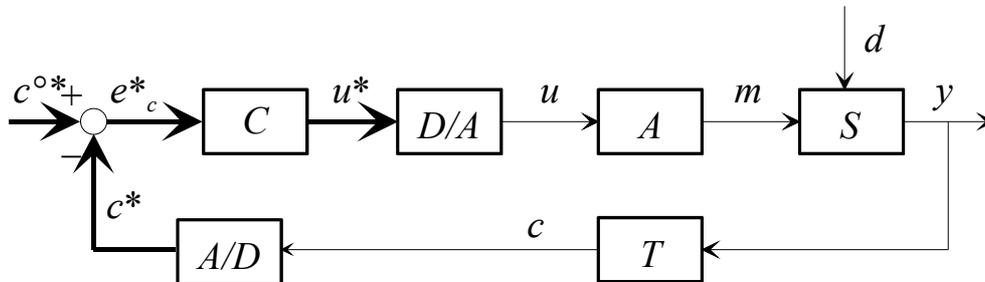


Fig. 1 : Sistema di controllo digitale

Uno schema alternativo in cui si converte in digitale l'errore è il seguente:

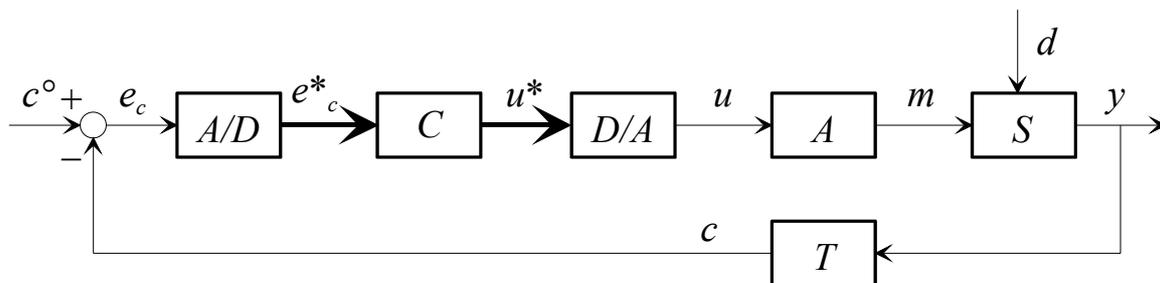


Fig. 2 : Sistema di controllo digitale con conversione in digitale dell'errore

Supporteremo tutti i sistemi oggetto della presente trattazione monovariabili (ad un ingresso ed una uscita).

Conversione A/D: il campionamento

Si consideri un generico segnale analogico $v(t)$, $t \in \mathfrak{R}$. Si fissi un'origine per l'asse dei tempi e, a partire da tale istante ($t=0$), si considerino istanti di tempo distanziati l'uno dall'altro da un intervallo T_C . Si valuti quindi il segnale $v(t)$ in corrispondenza di tutti questi istanti:

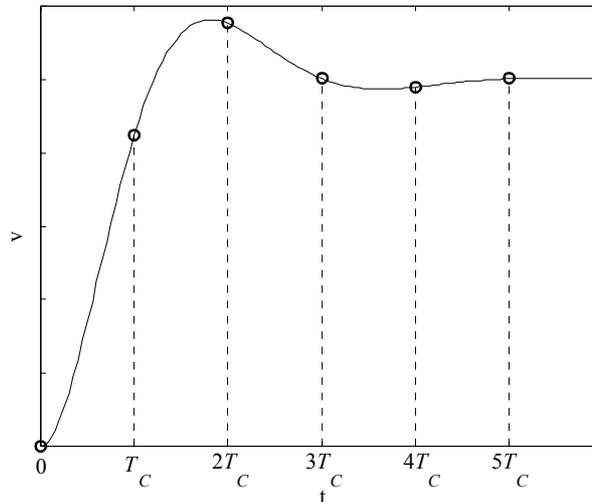


Fig. 3 : Campionamento

Si otterrà una sequenza (o successione) di numeri che indichiamo con:

$$v^*(k) = v(kT_C), \quad k = 0, 1, 2, \dots$$

Adotteremo la seguente terminologia:

- **Segnale campionato:** la sequenza di numeri $v^*(k)$.
- **Campionamento:** l'operazione che fa passare dal segnale a tempo continuo v al segnale a tempo discreto v^* .
- **Periodo (o tempo, o intervallo, o passo) di campionamento:** l'intervallo di tempo T_C .
- **Frequenza di campionamento:** l'inverso del periodo di campionamento, $f_C = 1/T_C$.
- **Pulsazione di campionamento:** la pulsazione corrispondente a f_C , $\Omega_C = 2\pi/T_C$.
- **Pulsazione di Nyquist:** la pulsazione $\Omega_N = \Omega_C/2 = \pi/T_C$.

Nella conversione analogico/digitale è anche inevitabile una *quantizzazione* del segnale, vale a dire la suddivisione dell'insieme dei valori che può assumere il segnale a tempo continuo v in un numero finito di intervalli. Tutti i valori di v interni allo stesso intervallo danno luogo allo stesso campione espresso in bit. La trattazione analitica dell'errore di quantizzazione è decisamente complessa e la sua rilevanza è in qualche misura ridimensionata dall'utilizzo di convertitori ad elevato numero di bit. Nel seguito non ci occuperemo degli effetti della quantizzazione, ritenendo l'insieme in cui varia v^* identico a quello in cui prende valori v .

Si intuisce che all'operazione di campionamento è generalmente associata una perdita di informazione, rispetto al segnale originario analogico. Non sembra infatti possibile compiere l'operazione inversa, ossia risalire univocamente dal segnale campionato al segnale originario, dal momento che diversi segnali analogici, anche molto differenti tra loro, possono dar luogo allo stesso segnale campionato:

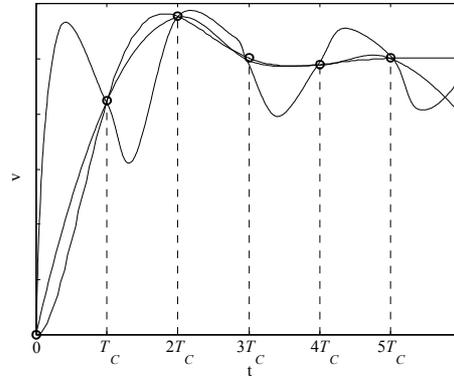


Fig. 4 : Segnali analogici che danno luogo allo stesso segnale campionato

In particolare, campionando un segnale sinusoidale con determinati periodi di campionamento, si può ottenere un segnale campionato che esibisce una oscillazione di lungo periodo, del tutto assente nel segnale originario:

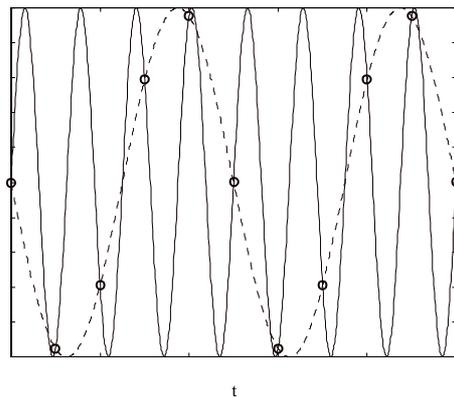


Fig. 5 : Aliasing

Questo fenomeno, detto **aliasing**, impone che il periodo di campionamento sia adeguato alle caratteristiche del segnale soggetto al campionamento. Il campionamento deve essere sufficientemente fitto (e quindi il periodo di campionamento sufficientemente piccolo) da cogliere anche le variazioni più rapide del segnale.

A questo proposito, è facile convincersi che per campionare correttamente un segnale sinusoidale occorrono almeno due campioni per periodo. Detto \bar{T} il periodo della sinusoide si avrà quindi:

$$\bar{T} > 2T_C \Rightarrow 2\pi/\bar{T} < \pi/T_C,$$

ossia, detta $\bar{\omega} = 2\pi/\bar{T}$ la pulsazione della sinusoide,:

$$\boxed{\Omega_N > \bar{\omega}}.$$

La pulsazione di Nyquist deve quindi essere superiore alla pulsazione della sinusoidale. Se così non è, si genera un'armonica di "alias" di bassa frequenza, la cui pulsazione è data dalla formula (come si può dimostrare):

$$\omega_{al} = |\Omega_C - \bar{\omega}|, \quad \text{per } \bar{\omega} < 3\Omega_N.$$

Per valori superiori di pulsazione valgono formule analoghe.

Si consideri a titolo d'esempio il campionamento di una sinusoidale di periodo \bar{T} con tempo di campionamento $T_C = 3/4\bar{T}$:

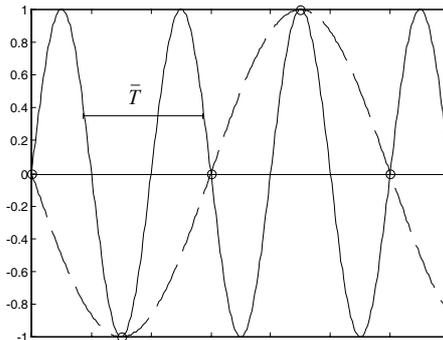


Fig. 6 : Campionamento di una sinusoidale

Risulta, in questo caso:

$$T_C = \frac{3}{4}\bar{T} \Rightarrow \frac{2\pi}{T_C} = \frac{4}{3} \frac{2\pi}{\bar{T}} \Rightarrow \Omega_C = \frac{4}{3}\bar{\omega}.$$

Pertanto l'armonica di alias ha pulsazione:

$$\omega_{al} = \Omega_C - \bar{\omega} = \frac{4}{3}\bar{\omega} - \bar{\omega} = \frac{1}{3}\bar{\omega},$$

e quindi periodo 3 volte il periodo della sinusoidale originaria (come si rileva graficamente).

Volendo estendere il risultato precedente a segnali generici, è utile ricordare che qualsiasi segnale a tempo continuo si può scomporre, sotto ipotesi molto blande, in serie o integrale (a seconda che sia periodico o aperiodico) di infinite sinusoidi, o componenti armoniche, attraverso la trasformata di Fourier. Un segnale si dice poi a **banda limitata** se ha trasformata di Fourier nulla, o sostanzialmente nulla, per pulsazioni maggiori di una pulsazione Ω_v . E' allora evidente che se tutte le armoniche del segnale hanno pulsazione inferiore alla pulsazione di Nyquist Ω_N , non si genera aliasing nel campionamento di alcuna di tali armoniche, e quindi del segnale stesso. Ne consegue il seguente importante risultato:

Teorema di Shannon (o del campionamento)

Se il segnale soggetto a campionamento $v(t)$ è a banda limitata con estremo superiore della banda Ω_v inferiore alla pulsazione di Nyquist $\Omega_N = \pi/T_C$:

$$\boxed{\Omega_v < \Omega_N}$$

allora la conoscenza del segnale campionato $v^*(k)$, con periodo di campionamento T_C , consente di ricostruire esattamente il segnale originario $v(t)$, ossia segnale campionato e segnale soggetto al campionamento sono informativamente equivalenti.

Se è rispettata la condizione del teorema del campionamento, deve essere possibile ricostruire, a partire dalla sequenza completa dei campioni del segnale campionato $v^*(k)$, il segnale originario $v(t)$ ad ogni istante. La formula che risolve il problema è la **formula di Shannon** (o del **decampionatore di Shannon**):

$$v(t) = \sum_{k=-\infty}^{+\infty} \left[v^*(k) \frac{\sin(\Omega_N t - k\pi)}{\Omega_N t - k\pi} \right].$$

Si osservi che, come è ovvio, vi sono comunque infiniti segnali a tempo continuo che generano tramite campionamento la stessa sequenza $v^*(k)$, ma di questi uno solo soddisfa la condizione del teorema di Shannon, e la sua espressione è data dalla formula del decampionatore.

Scelto il periodo di campionamento del convertitore A/D, è anche possibile forzare il segnale a tempo continuo al soddisfacimento della condizione del teorema del campionamento, filtrando il segnale stesso con un filtro passabasso. Tale filtro, che va sotto il nome di **filtro anti-aliasing**, avrà guadagno unitario e pulsazione di taglio inferiore alla pulsazione di Nyquist, in modo da tagliare le componenti ad alta frequenza del segnale.

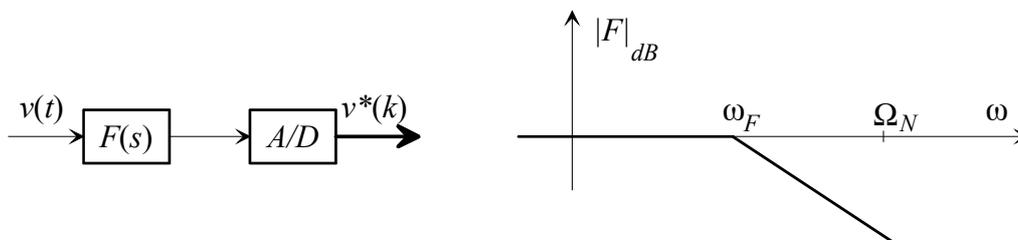


Fig. 7 : Filtro anti-aliasing

Conversione D/A

La conversione digitale/analogico consiste nel ricavare da una sequenza di valori $v^*(k)$ cui è associata una base dei tempi, un segnale a tempo continuo, che negli istanti associati ai valori $v^*(k)$, assuma gli stessi valori della sequenza data.

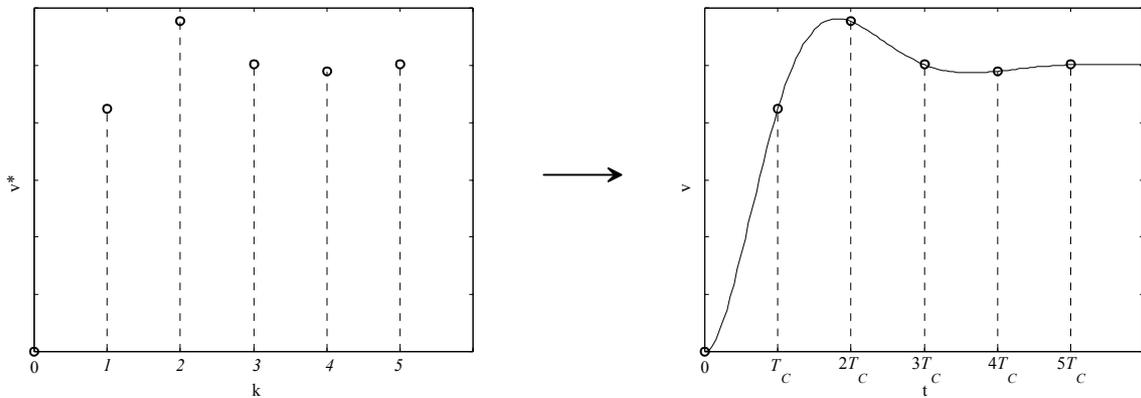


Fig. 8 : Conversione D/A

La formula di Shannon risolverebbe il problema generando un segnale a banda limitata: tuttavia non è utilizzabile negli schemi di controllo digitale in quanto **non causale**. Il valore di $v(t)$, in ciascun istante di tempo, dipende infatti da tutti i valori di $v^*(k)$, in particolare anche da quelli associati ad istanti successivi a t . Il calcolatore impiegato nel controllo digitale, operando in tempo reale, non può disporre che dei valori passati della variabile di controllo da convertire. Per risolvere il problema si usano allora degli **estrapolatori**, ossia dei dispositivi che, sulla base di un certo numero di campioni (i più recenti elaborati dal dispositivo di controllo digitale), determinano il valore che dovrà assumere l'uscita del convertitore fino al successivo campione.

In particolare la soluzione più comunemente utilizzata consiste semplicemente nel mantenere costante l'ultimo campione in tutto l'intervallo di campionamento:

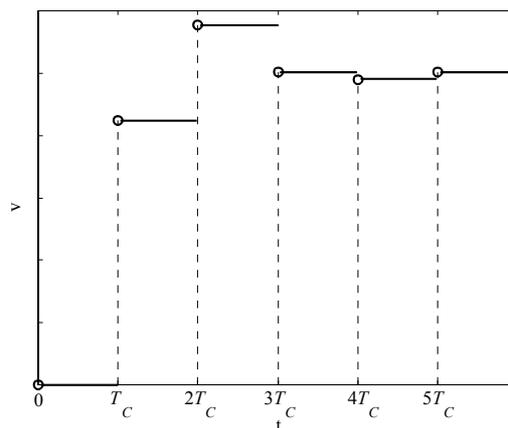


Fig. 9 : Mantentore di ordine zero

Il dispositivo che realizza questa operazione si chiama **Mantentore di ordine zero** o **ZOH** (**Z**ero **O**rdine **H**old).

Consideriamo ora la sequenza di un campionatore ed uno ZOH:

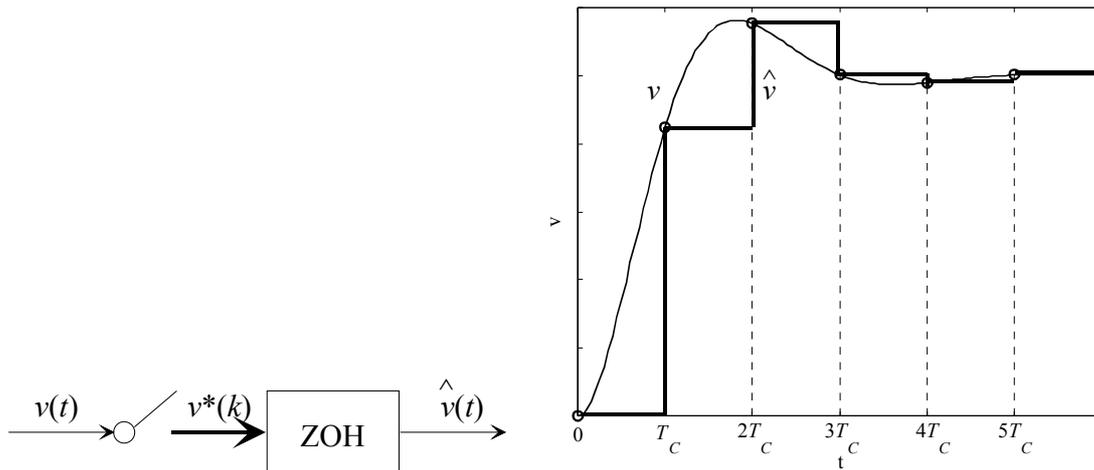


Fig. 10 : Campionatore e mantenitore

Tra i segnali a tempo continuo in ingresso ed uscita, v e \hat{v} si evidenzia un certo ritardo. Tale ritardo è quantificabile, con un'analisi non del tutto banale nel dominio della frequenza, in circa metà del periodo di campionamento, $\tau = T_c/2$, e prende il nome di **ritardo intrinseco di conversione**.

Regolatore digitale e temporizzazione

Il regolatore digitale elabora in linea la sequenza di valori ottenuti dal campionamento dell'errore e produce la sequenza di valori da attribuire alla variabile di controllo:

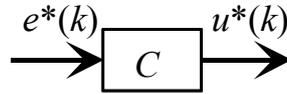


Fig. 11 : Regolatore digitale

Il regolatore può implementare, via software, un algoritmo arbitrariamente complesso, sfruttando tutte le opportunità date dalla programmazione. Noi in realtà ci limiteremo a regolatori il cui algoritmo è costituito da un insieme di equazioni alle differenze, interpretabile come un sistema dinamico a tempo discreto lineare tempo invariante, di funzione di trasferimento:

$$R(z) = \frac{U^*(z)}{E^*(z)}$$

Il problema di progetto in un sistema di controllo digitale consisterà nel ricavare la funzione di trasferimento $R(z)$.

Supporremo inoltre campionatore e mantentore sincroni e in fase (cioè operanti con lo stesso periodo e con la stessa origine dei tempi). Si osservi che tuttavia sono inevitabili ritardi nelle operazioni di conversione e soprattutto nell'elaborazione da parte del calcolatore: chiamiamo τ_c la somma di questi ritardi. Naturalmente deve essere $\tau_c < T_c$, altrimenti il controllore digitale perderebbe il passo (non avrebbe ancora terminato l'elaborazione all'arrivo del nuovo campione).

Trascurando il ritardo dovuto alla conversione, al generico istante di campionamento kT_c il calcolatore acquisisce il campione dell'errore $e^*(k)$. A questo punto, a seconda che il regolatore digitale sia strettamente proprio o proprio non strettamente si possono seguire due strategie diverse per quanto riguarda l'istante in cui aggiornare l'ingresso del convertitore digitale/analogico (uscita del calcolatore).

Se $R(z)$ è **strettamente proprio** $u^*(k)$ non dipende da $e^*(k)$, per cui la conoscenza del nuovo campione dell'errore non è richiesta per il calcolo del nuovo campione della variabile di controllo $u^*(k)$. Questo comporta che al termine dell'intervallo di campionamento si potrà aggiornare l'ingresso del convertitore D/A con il campione $u^*(k+1)$, che invece può dipendere da $e^*(k)$:

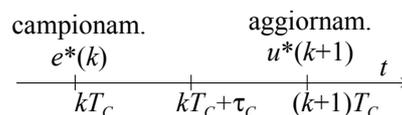


Fig. 12 : Temporizzazione nel caso di regolatore strettamente proprio

In altre parole, il ritardo di elaborazione non comporta ulteriori ritardi nell'anello (il regolatore stesso ha già in sé almeno un passo di ritardo).

Se $R(z)$ è **proprio non strettamente** $u^*(k)$ dipende da $e^*(k)$, per cui la conoscenza del nuovo campione dell'errore è richiesta per il calcolo del nuovo campione della variabile di controllo

$u^*(k)$. Questo comporta che, in linea di principio, non appena il calcolatore ha terminato l'elaborazione, la sua uscita andrebbe aggiornata con il nuovo campione $u^*(k)$ calcolato:

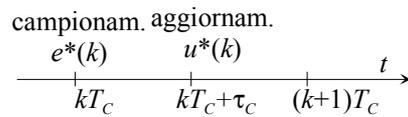


Fig. 13 : Temporizzazione nel caso di regolatore proprio non strettamente

In questo modo, tuttavia, si genererebbe un ritardo variabile (a seconda del tempo richiesto per eseguire le operazioni in ciascun intervallo di campionamento) e si incorrerebbe in difficoltà tecnologiche legate ad un'operazione di aggiornamento del convertitore D/A asincrona con la temporizzazione degli altri organi. Una soluzione alternativa consiste nell'aggiornare l'uscita del calcolatore comunque al termine dell'intervallo di campionamento, accettando di introdurre nell'anello un ritardo fisso pari a T_c :

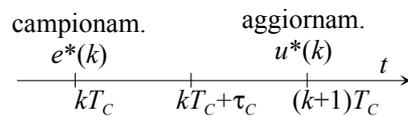


Fig. 14 : Temporizzazione alternativa nel caso di regolatore proprio non strettamente

Punti di vista analogico e digitale

Nell'analisi di un sistema di controllo digitale si possono adottare due punti di vista:

Punto di vista analogico

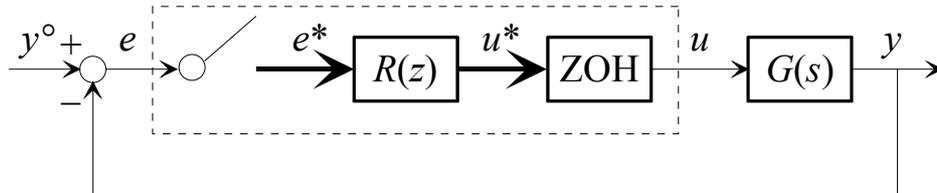


Fig. 15 : Punto di vista analogico

La parte tratteggiata è un sistema esternamente a tempo continuo. Dato un regolatore analogico ci proponiamo di realizzarlo in tecnologia digitale, con un'opportuna scelta di $R(z)$ e del periodo di campionamento.

Punto di vista digitale

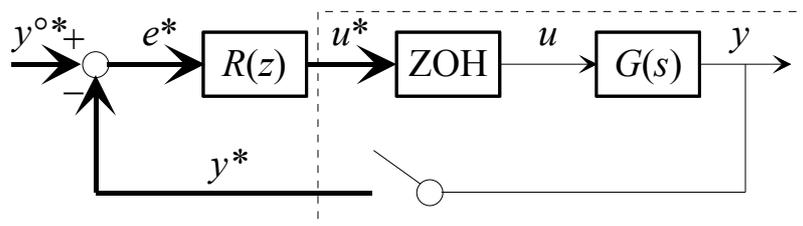


Fig. 16 : Punto di vista digitale

La parte tratteggiata è un sistema esternamente a tempo discreto, di cui si determinerà la funzione di trasferimento. Si progetterà quindi il regolatore $R(z)$ direttamente con la teoria dei sistemi a tempo discreto.

Realizzazione digitale di un controllore analogico

Supponiamo di avere già progettato un controllore analogico $R^\circ(s)$ e di volerlo realizzare mediante un regolatore a tempo discreto $R(z)$ con un tempo di campionamento T .

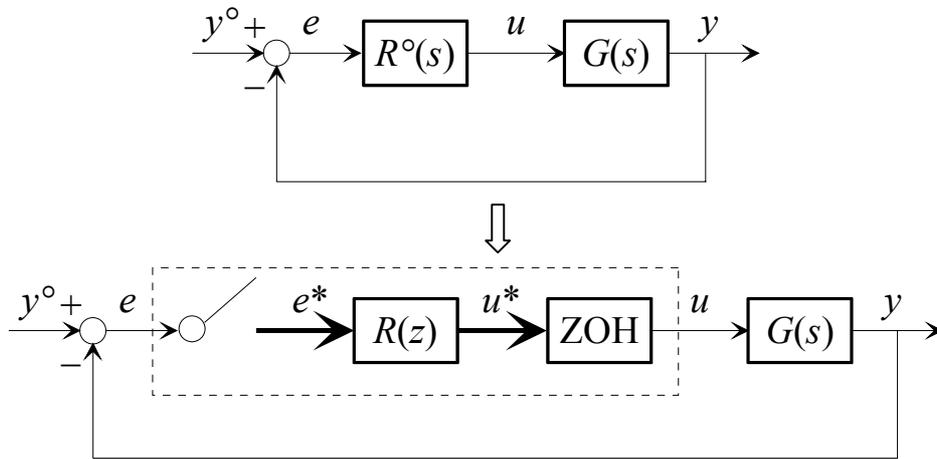


Fig. 17 : Realizzazione digitale del controllore

Poniamoci dapprima il problema della **scelta del periodo di campionamento**. Il segnale soggetto a campionamento, l'errore e , deve soddisfare la condizione del teorema di Shannon. Poiché però non si tratta di un segnale esogeno, le sue caratteristiche spettrali non sono determinabili a priori. E' noto tuttavia che in un sistema di controllo le componenti armoniche significative dei segnali risiedono nella banda passante del sistema di controllo stesso, il cui estremo superiore è di norma ben approssimato dalla pulsazione critica ω_c . Per rispettare la condizione del teorema di Shannon occorrerà allora che tutto il contenuto informativo utile dell'errore sia in una banda inferiore alla pulsazione di Nyquist, ossia che la pulsazione critica sia decisamente inferiore alla pulsazione di Nyquist:

$$\omega_c \ll \Omega_N.$$

Un buon criterio è che le due pulsazioni siano separate da una decade ($\Omega_N/\omega_c = 10$). Valori eccessivamente alti della pulsazione di Nyquist (ossia valori molto piccoli del tempo di campionamento) non sono peraltro raccomandabili sia per ragioni di costo dei dispositivi, sia per il fatto che enfatizzano gli errori legati alla quantizzazione.

In presenza di disturbi con componenti armoniche di alta frequenza o di componenti in alta frequenza del segnale di riferimento (che non interessa riprodurre in uscita) è opportuno anteporre al convertitore A/D un **filtro antialiasing**:

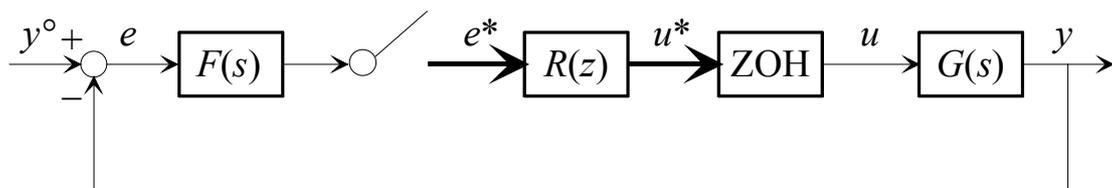


Fig. 18 : Inserimento di un filtro antialiasing nell'anello

La pulsazione di taglio del filtro dovrà essere superiore a ω_c (per non tagliare componenti armoniche significative del segnale da campionare) ma inferiore a Ω_N (per rispettare la condizione del teorema del campionamento):

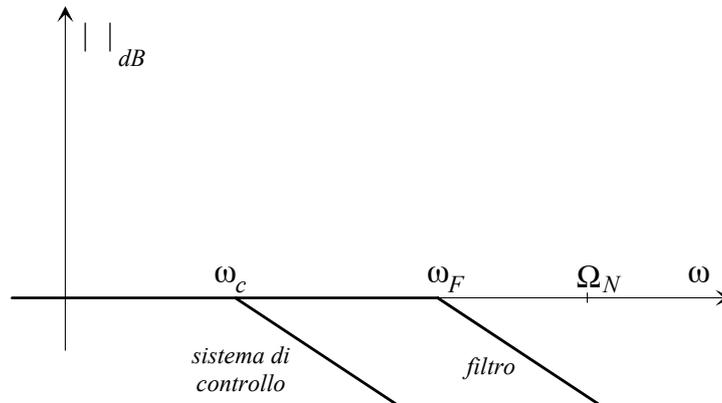


Fig. 19 : Progetto del filtro antialiasing

Si osservi a questo punto che il fatto stesso di realizzare il regolatore in tecnologia digitale con campionatori e mantentori introduce il **ritardo intrinseco di conversione** (pari a metà del periodo di campionamento). A questo si sommano i ritardi di elaborazione, di conversione e lo sfasamento introdotto dall'eventuale filtro antialiasing.

E' bene quindi che il regolatore $R^\circ(s)$ progettato a tempo continuo sia dotato di **un'eccedenza di margine di fase** tale da coprire gli sfasamenti introdotti dalla realizzazione digitale.

Se per esempio $\Omega_N=10\omega_c$, tenendo conto del solo ritardo intrinseco di conversione, si deve prevedere la seguente eccedenza di margine di fase:

$$\Delta\varphi_m = \frac{T_C}{2} \omega_c \frac{180^\circ}{\pi} = 90^\circ \frac{\omega_c}{\Omega_N} = 9^\circ.$$

Per quanto riguarda **la scelta del regolatore** $R(z)$ occorre fare in modo che la serie di campionatore, sistema di funzione di trasferimento $R(z)$ e ZOH si comporti agli effetti esterni come $R^\circ(s)$. Si osservi a questo proposito che la funzione di trasferimento non è altro che un formalismo utile per rappresentare un sistema di equazioni differenziali lineari: ne consegue che i metodi noti per **l'integrazione numerica** delle equazioni differenziali possono tornare utili anche per la soluzione del nostro problema.

Consideriamo a questo proposito un semplice integratore a tempo continuo:

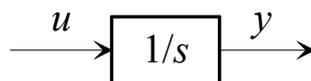


Fig. 20 : Sistema da integrare numericamente

Il sistema è descritto dall'equazione differenziale:

$$\dot{y}(t) = u(t).$$

Detti $u^*(k)=u(kT)$ e $y^*(k)=y(kT)$, con T intervallo di integrazione, risulta:

$$y^*(k) = y^*(k-1) + u_m(k)T,$$

dove $u_m(k)$ è il valor medio di u nell'intervallo di integrazione (per il cui calcolo occorre conoscere u in tutto l'intervallo d'integrazione). Approssimiamo u_m con una combinazione lineare convessa (tale che la somma dei coefficienti sia unitaria) dei valori assunti da u agli estremi dell'intervallo di integrazione:

$$y^*(k) = y^*(k-1) + T[(1-\alpha)u^*(k-1) + \alpha u^*(k)], \quad \alpha \in [0,1].$$

Si è ottenuta un'equazione alle differenze lineare, ad entrambi i membri della quale possiamo applicare la trasformata Zeta:

$$Y^*(z) = z^{-1}Y^*(z) + T[(1-\alpha)z^{-1} + \alpha]U^*(z)$$

e quindi:

$$\frac{Y^*(z)}{U^*(z)} = T \frac{(1-\alpha) + \alpha z}{z-1},$$

versione a tempo discreto dell'integratore $1/s$.

Da queste considerazioni discende un metodo per ricavare $R(z)$ a partire da $R^\circ(s)$. Si tratta semplicemente di sostituire a s l'espressione (ottenuta invertendo la precedente funzione di trasferimento):

$$s = \frac{1}{T} \frac{z-1}{\alpha z + (1-\alpha)},$$

che prende il nome di **trasformazione bilineare**.

Risulta pertanto:

$$R(z) = R^\circ\left(\frac{1}{T} \frac{z-1}{\alpha z + (1-\alpha)}\right).$$

Se $R^\circ(s)$ è razionale, di ordine n e propria, lo è anche $R(z)$, purché $R^\circ(s)$ non abbia poli in $s=1/(\alpha T)$, il che è escluso se R° è asintoticamente stabile. La trasformazione bilineare si specializza poi nelle seguenti formule di uso comune:

$$\alpha = 0 \quad s = \frac{z-1}{T} \quad \text{Eulero in avanti (o Eulero esplicito)}$$

$$\alpha = 1 \quad s = \frac{z-1}{Tz} \quad \text{Eulero all'indietro (o Eulero implicito)}$$

$$\alpha = \frac{1}{2} \quad s = \frac{2}{T} \frac{z-1}{z+1} \quad \text{Trapezi o Tustin}$$

L'applicazione della formula di Eulero esplicita ad un sistema asintoticamente stabile a tempo continuo può dar luogo a un sistema a tempo discreto instabile se non si sceglie un tempo di campionamento sufficientemente piccolo.

Esempio

Consideriamo un **regolatore PI** analogico:

$$R^\circ(s) = k_p \left(1 + \frac{1}{sT_I}\right).$$

Utilizzando la trasformazione di Tustin:

$$\begin{aligned} R(z) &= R^o\left(\frac{2}{T} \frac{z-1}{z+1}\right) = k_p \left(1 + \frac{T}{2T_I} \frac{z+1}{z-1}\right) = \frac{k_p}{2T_I} \left(\frac{(2T_I + T)z + T - 2T_I}{z-1}\right) = \\ &= \Gamma_p \frac{z-b}{z-1} \end{aligned}$$

con:

$$\Gamma_p = k_p \left(1 + \frac{T}{2T_I}\right), \quad b = \frac{2T_I - T}{2T_I + T}.$$

Pertanto:

$$\frac{U^*(z)}{E^*(z)} = \Gamma_p \frac{z-b}{z-1} = \Gamma_p \frac{1-bz^{-1}}{1-z^{-1}}$$

ossia:

$$(1-z^{-1})U^*(z) = \Gamma_p(1-bz^{-1})E^*(z)$$

Passando nel dominio del tempo:

$$u^*(k) = u^*(k-1) + \Gamma_p e^*(k) - \Gamma_p b e^*(k-1).$$

Questa equazione alle differenze può essere tradotta in un **programma di calcolo**, da eseguire ad ogni istante di campionamento:

```
input yref, y;
e = yref-y;
u = u + Gammap*e-Gammap*b*eold;
eold = e;
```

Sistemi a segnali campionati

Il “punto di vista digitale” nell’analisi di un sistema di controllo digitale induce a considerare (si veda la Fig. 16) il sistema esternamente a tempo discreto costituito dalla serie di uno ZOH, di un sistema a tempo continuo e di un campionatore: questo sistema prende il nome di **sistema a segnali campionati**.

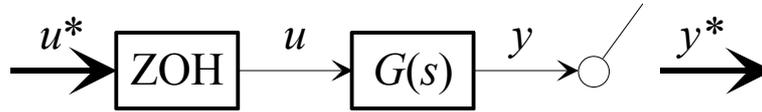


Fig. 21 : Sistema a segnali campionati

Ci proponiamo, dati la funzione di trasferimento $G(s)$ ed il tempo di campionamento T , di ricavare la funzione di trasferimento del sistema a tempo discreto da u^* a y^* . Supporremo la funzione di trasferimento $G(s)$ razionale, assumendo quindi in particolare che non presenti ritardi di tempo. Sia (A, B, C, D) una realizzazione minima di $G(s)$, ossia una quaterna di matrici, con le dimensioni di A pari al grado del denominatore di G , tali che:

$$G(s) = C(sI - A)^{-1}B + D.$$

Il sistema di funzione di trasferimento $G(s)$ può quindi essere rappresentato dalle seguenti equazioni nel dominio del tempo:

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t)$$

$$y(t) = C\mathbf{x}(t) + D\mathbf{u}(t)$$

Ricordiamo che l’espressione dello stato al generico istante t , noti lo stato ad un istante antecedente t_0 e l’ingresso a partire da t_0 , è data dalla formula di Lagrange:

$$\mathbf{x}(t) = e^{A(t-t_0)}\mathbf{x}_0 + \int_{t_0}^t e^{A(t-\tau)}B\mathbf{u}(\tau)d\tau.$$

Sceglieremo ora come istante iniziale t_0 un generico istante di campionamento e come istante t il successivo istante di campionamento, ossia $t_0 = kT$, $t = kT + T$:

$$\mathbf{x}(kT + T) = e^{AT}\mathbf{x}(kT) + \int_{kT}^{kT+T} e^{A(kT+T-\tau)}B\mathbf{u}(\tau)d\tau.$$

Poniamo ora $\mathbf{x}^*(k) = \mathbf{x}(kT)$ ed osserviamo che, in virtù del principio di funzionamento dello ZOH, la variabile u resta costante per tutto l’intervallo di campionamento, e quindi per tutto l’intervallo compreso tra gli estremi d’integrazione:

$$u(\tau) = u(kT) = u^*(k) \quad \tau \in [kT, kT + T)$$

Pertanto:

$$\mathbf{x}^*(k+1) = e^{AT}\mathbf{x}^*(k) + \left[\int_{kT}^{kT+T} e^{A(kT+T-\tau)}B d\tau \right] u^*(k).$$

Eseguito il cambiamento di variabili:

$$\sigma = kT + T - \tau$$

otteniamo:

$$\mathbf{x}^*(k+1) = e^{AT} \mathbf{x}^*(k) + \left[\int_0^T e^{A\sigma} \mathbf{B} d\sigma \right] u^*(k)$$

In definitiva, il sistema a tempo discreto si lascia descrivere dalle seguenti equazioni:

$$\begin{aligned} \mathbf{x}^*(k+1) &= \mathbf{A}^* \mathbf{x}^*(k) + \mathbf{B}^* u^*(k) \\ y^*(k) &= \mathbf{C} \mathbf{x}^*(k) + \mathbf{D} u^*(k) \end{aligned}$$

con:

$$\mathbf{A}^* = e^{AT}, \quad \mathbf{B}^* = \int_0^T e^{A\sigma} \mathbf{B} d\sigma,$$

per cui la sua funzione di trasferimento è:

$$G^*(z) = \mathbf{C} (z\mathbf{I} - \mathbf{A}^*)^{-1} \mathbf{B}^* + \mathbf{D}.$$

Dalla relazione precedentemente ricavata tra la matrice \mathbf{A} del sistema a tempo continuo e la matrice \mathbf{A}^* del sistema a tempo discreto, discende la relazione tra un autovalore s_i della matrice \mathbf{A} e un autovalore z_i della matrice \mathbf{A}^* :

$$z_i = e^{s_i T}, \quad i = 1, \dots, n.$$

Pertanto i **poli** delle due funzioni di trasferimento sono legati dalla trasformazione:

$$z = e^{sT},$$

che prende il nome di **trasformazione di campionamento**.

La trasformazione mette in relazione regioni del piano complesso s (nel quale si indicano i poli del sistema a tempo continuo) con regioni del piano z (relativo ai poli del sistema a tempo discreto), come mostra la figura:

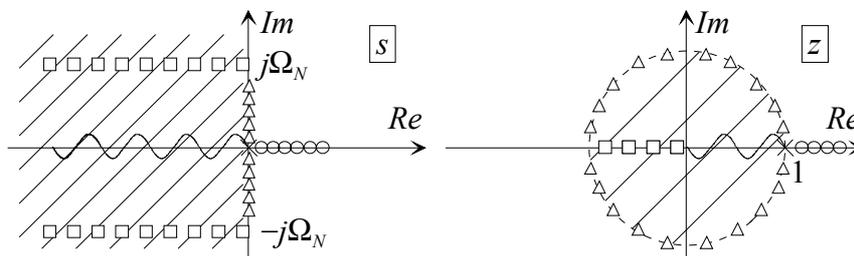


Fig. 22 : Trasformazione di campionamento

Si osservi in particolare che:

- i punti a parte reale negativa in s si trasformano in punti a modulo minore di 1 in z ;
- l'origine in s si trasforma in $z=1$;

- la trasformazione non è biunivoca: presi due punti in s le cui parti immaginarie differiscono per multipli della pulsazione di campionamento, essi si trasformano nello stesso punto in z :

$$s_2 = s_1 + jh\Omega_c \Rightarrow e^{s_2 T} = e^{s_1 T} e^{jh2\pi} = e^{s_1 T}, \quad h \text{ intero}$$

La prima osservazione implica che se il sistema a tempo continuo è asintoticamente stabile, lo è anche il sistema a segnali campionati.

Il legame tra gli **zeri** di G e G^* è più articolato:

- Se $G(s)$ ha m zeri, ci sono m zeri in $G^*(z)$ che tendono, al tendere di T a zero, alle immagini secondo la trasformazione di campionamento degli zeri di $G(s)$;
- Se $n > m + 1$, compaiono in $G^*(z)$ $n - m - 1$ zeri aggiuntivi (detti **zeri del campionamento**) la cui posizione tende, al tendere a zero di T , in alcune posizioni fisse (indipendenti da G).

La funzione di trasferimento del sistema a segnali campionati si può anche ottenere in modo più semplice, osservando che uno scalino a tempo discreto sulla variabile $u^*(k)$ genera, per come opera lo ZOH, uno scalino a tempo continuo su $u(t)$. Ne consegue la seguente **procedura**:

1. Detta $Y(s) = G(s)/s$ la trasformata di Laplace della risposta di $G(s)$ allo scalino, si antitrasforma Y per ottenere la risposta allo scalino nel dominio del tempo $y(t)$;
2. Si valuta la risposta allo scalino negli istanti di campionamento, ottenendo $y^*(k) = y(kT)$;
3. Si calcola la trasformata Zeta $Y^*(z)$ di $y^*(k)$;
4. Si ricava $G^*(z)$ come rapporto tra $Y^*(z)$ e la trasformata dello scalino a tempo discreto:

$$G^*(z) = \frac{Y^*(z)}{\frac{z}{z-1}} = \frac{z-1}{z} Y^*(z).$$

La funzione di trasferimento che si ottiene con questa procedura è ovviamente identica a quella ottenuta precedentemente. Si ponga attenzione al fatto che quanto ottenuto con le due procedure espone un risultato esatto e non va confuso con le formule introdotte per la discretizzazione della legge di controllo (formule di Eulero e di Tustin).

Esempio

Si consideri il sistema di funzione di trasferimento:

$$G(s) = \frac{P}{s + p}.$$

Una realizzazione minima (in forma canonica di controllo) del sistema è la seguente:

$$\begin{aligned} \dot{x}(t) &= -px(t) + u(t) \\ y(t) &= px(t) \end{aligned}$$

Risulta pertanto:

$$\mathbf{A} = -p, \mathbf{B} = 1, \mathbf{C} = p, \mathbf{D} = 0.$$

Il sistema a segnali campionati è quindi caratterizzato dalle matrici:

$$\mathbf{A}^* = e^{AT} = e^{-pT}, \quad \mathbf{B}^* = \int_0^T e^{A\sigma} \mathbf{B} d\sigma = \int_0^T e^{-p\sigma} d\sigma = \left[\frac{e^{-p\sigma}}{-p} \right]_0^T = \frac{1 - e^{-pT}}{p},$$

e dalla funzione di trasferimento:

$$G^*(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A}^*)^{-1} \mathbf{B}^* + \mathbf{D} = \frac{\mathbf{CB}^*}{z - \mathbf{A}^*} = \frac{1 - e^{-pT}}{z - e^{-pT}}.$$

D'altra parte la risposta allo scalino di $G(s)$ ha trasformata:

$$Y(s) = \frac{G(s)}{s} = \frac{p}{s(s+p)} = \frac{1}{s} - \frac{1}{s+p},$$

e si antitrasforma immediatamente in:

$$y(t) = 1 - e^{-pt}, \quad t \geq 0.$$

Valutando questa risposta negli istanti di campionamento si ha:

$$y^*(k) = y(kT) = 1 - e^{-pkT}, \quad k \geq 0.$$

La trasformata Zeta di questa funzione è:

$$Y^*(z) = \frac{z}{z-1} - \frac{z}{z - e^{-pT}}.$$

Infine la funzione di trasferimento si ottiene come:

$$G^*(z) = \frac{z-1}{z} Y^*(z) = 1 - \frac{z-1}{z - e^{-pT}} = \frac{1 - e^{-pT}}{z - e^{-pT}},$$

e coincide ovviamente con quella ricavata con l'altro metodo.

Si osservi che questa funzione di trasferimento ha guadagno unitario (come quella a tempo continuo) e che il suo polo è legato al polo di $G(s)$ dalla trasformazione di campionamento.

Progetto a tempo discreto del regolatore

Per il progetto del regolatore si può ora fare riferimento ad un sistema interamente a tempo discreto:

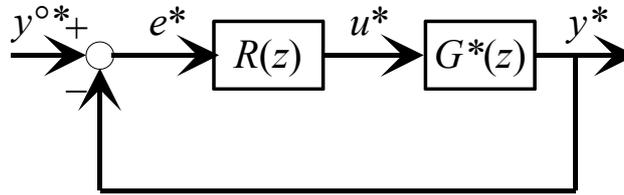


Fig. 23 :Sistema di controllo a tempo discreto

dove $G^*(z)$ è la funzione di trasferimento del sistema a segnali campionati prima discussa.

A questo punto il regolatore si può progettare con la **teoria dei sistemi a tempo discreto**. Molti dei concetti e metodi visti per i sistemi di controllo a tempo continuo restano inalterati: il criterio di Nyquist ha fondamentalmente lo stesso enunciato, come pure è uguale il tracciamento del luogo delle radici.

Si osservi che per la scelta del **periodo di campionamento**, non essendoci un preesistente progetto a tempo continuo a cui rifarsi, occorre seguire altri criteri. Ad esempio si può selezionare il tempo di campionamento in modo che nel tempo di assestamento dei transienti del sistema in anello chiuso siano compresi un numero di passi di campionamento compreso tra 5 e 50.

A tempo discreto vengono inoltre utilizzati, più che a tempo continuo, metodi che assegnano il modello del sistema in anello chiuso (ossia la funzione di trasferimento dal riferimento alla variabile controllata) e da esso ricavano direttamente la funzione di trasferimento del controllore. Uno di questi metodi, noto anche come **metodo di Ragazzini**, prevede una procedura sistematica per l'assegnamento del modello in modo che siano soddisfatte una serie di requisiti sul sistema di controllo (causalità, stabilità, prestazioni) e verrà di seguito esposto nei suoi tratti essenziali.

Sia $F^*(z)$ la funzione di sensitività complementare del sistema di controllo di Fig. 23, ossia:

$$F^*(z) = \frac{Y^*(z)}{Y^{o*}(z)} = \frac{R(z)G^*(z)}{1 + R(z)G^*(z)}.$$

Assegnata $F^*(z)$, ossia assegnato il modello del sistema in anello chiuso, si può immediatamente ricavare la funzione di trasferimento del controllore:

$$R(z) = \frac{F^*(z)}{G^*(z)[1 - F^*(z)]}.$$

Tuttavia, al fine di pervenire ad un risultato di utilità pratica, ovvero ad un regolatore realizzabile, che non comporti generazione di parti non raggiungibili o non osservabili instabili e che conferisca al sistema di controllo determinate prestazioni, occorre scegliere $F^*(z)$ rispettando alcuni vincoli. Per agevolare la trattazione, esprimiamo tutte le funzioni di trasferimento come rapporti di polinomi, primi tra loro:

$$G^*(z) = \frac{N(z)}{D(z)}, \quad R(z) = \frac{Q(z)}{P(z)}, \quad F^*(z) = \frac{B(z)}{A(z)}.$$

Siano inoltre $n_N, n_D, n_Q, n_P, n_B, n_A$ i gradi dei sei polinomi. Assumendo $F^*(z)$ strettamente propria (al fine di evitare un anello algebrico nel sistema di controllo) sarà:

$$n_A > n_B.$$

In termini di questi polinomi, la **formula risolutiva** che dà la funzione di trasferimento del controllore si può esprimere come segue:

$$R(z) = \frac{Q(z)}{P(z)} = \frac{D(z)B(z)}{N(z)[A(z) - B(z)]}.$$

Si osservi che i polinomi Q e P sono quelli che si ottengono a seguito di eventuali cancellazioni formali tra i polinomi a numeratore e a denominatore che compaiono nella formula risolutiva.

Vediamo ora da quali considerazioni scaturiscono i vincoli sulla scelta di $F^*(z)$.

1) Causalità del controllore

Il controllore deve essere un sistema causale (o fisicamente realizzabile) ovvero deve costituire un sistema dinamico proprio: il suo **grado relativo** (differenza tra il numero di poli ed il numero degli zeri) non deve essere negativo. Questo comporta che:

$$n_P \geq n_Q.$$

D'altra parte, dalla formula risolutiva si ottiene:

$$n_P - n_Q = n_N + n_{A-B} - (n_D + n_B) = n_N + n_A - (n_D + n_B),$$

dove si è utilizzato il fatto che il grado di $A-B$ coincide con il grado di A .

Pertanto deve risultare:

$$n_N + n_A - (n_D + n_B) \geq 0,$$

ossia:

$$\boxed{n_A - n_B \geq n_D - n_N}.$$

Pertanto il grado relativo di F^* deve essere almeno pari al grado relativo di G^* .

2) Stabilità e cancellazioni

Affinché il sistema in anello chiuso sia asintoticamente stabile, le radici di $A(z)$ devono ovviamente essere tutte a modulo minore di 1. Bisogna però anche evitare che eventuali zeri o poli a modulo maggiore o uguale a 1 in G^* siano cancellati dal regolatore, in modo da non formare parti non raggiungibili o non osservabili non asintoticamente stabili. Dalla formula risolutiva si traggono facilmente le conclusioni.

- Ogni zero z_i di $G^*(z)$ ($N(z_i)=0$), a modulo maggiore o uguale a 1 deve essere anche radice di $B(z)$:

$$\boxed{N(z_i) = 0, \quad |z_i| \geq 1 \Rightarrow B(z_i) = 0}$$

- Ogni polo p_i di $G^*(z)$ ($D(p_i)=0$), a modulo maggiore o uguale a 1 deve essere anche radice di $A(z)-B(z)$:

$$\boxed{D(p_i) = 0, \quad |p_i| \geq 1 \Rightarrow A(p_i) = B(p_i)}$$

3) Precisione statica

Per avere errore nullo a regime con riferimento a scalino, $F^*(z)$ deve avere guadagno unitario ($F^*(1)=1$):

$$\boxed{A(1) = B(1)}$$

4) Precisione dinamica

Per quanto concerne la scelta di poli e zeri di $F^*(z)$, una delle strategie più comuni, peculiare dei sistemi a tempo discreto, è il **controllo in tempo finito**. Con questa strategia, la variabile controllata y^* raggiunge, a seguito di una variazione a scalino del riferimento, il valore di regime in un numero finito di passi. Si osservi che questa circostanza è impossibile a tempo continuo. Per imporre questo comportamento al sistema in anello chiuso occorre che tutti i **poli di F^* siano nell'origine del piano complesso**, ossia che F^* assuma l'espressione:

$$\boxed{F^*(z) = \frac{B(z)}{z^{n_A}}}$$

In questo caso, infatti, risulta:

$$\frac{Y^*(z)}{Y^{o*}(z)} = \frac{b_0 z^{n_A} + b_1 z^{n_A-1} + \dots + b_{n_A}}{z^{n_A}} = b_0 + b_1 z^{-1} + \dots + b_{n_A} z^{-n_A},$$

e, nel dominio del tempo:

$$y^*(k) = b_0 y^{o*}(k) + b_1 y^{o*}(k-1) + \dots + b_{n_A} y^{o*}(k-n_A).$$

Si deduce quindi che effettivamente, se $y^{o*}(k) = \text{sca}(k)$, dopo n_A passi y^* non cambia più valore.

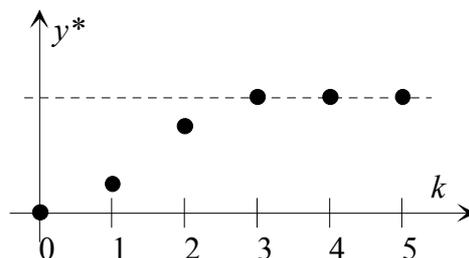


Fig. 24 : Controllo in tempo finito

Se si vuole in particolare che l'uscita raggiunga il valore di regime nel **tempo minimo**, sarà sufficiente porre:

$$\boxed{F^*(z) = \frac{1}{z^{\delta_G}}}$$

con $\delta_G = n_D - n_N$ grado relativo di G^* , per rispettare la condizione sulla causalità del controllore. La risposta di F^* allo scalino è quindi quella di un ritardo puro:

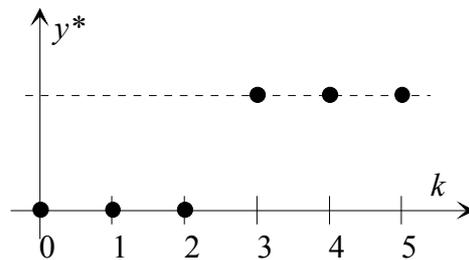


Fig. 25 : Controllo in tempo minimo

Si osservi che questa soluzione è possibile solo se G^* è un sistema strettamente proprio, asintoticamente stabile e a fase minima. Per ogni polo o zero di G^* a modulo maggiore o uguale a 1, occorrerà infatti aggiungere uno zero alla funzione di trasferimento ed un polo nell'origine, in modo da poter applicare la condizione di non cancellazione riportata precedentemente.

Pur se attraente dal punto di vista concettuale, la soluzione del controllo in tempo minimo così come presentata nasconde alcune difficoltà: la variabile di controllo è sollecitata in modo molto rilevante e può insorgere il fenomeno delle “**oscillazioni nascoste**”, ovvero di oscillazioni della variabile controllata y a tempo continuo che vengono mascherate dal processo di campionamento:

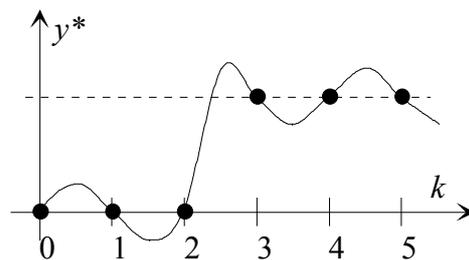


Fig. 26 : Oscillazioni nascoste

Per ovviare a queste difficoltà bisogna modificare il procedimento di sintesi, secondo tecniche alternative qui non affrontate.

Esempio

Sia:

$$G^*(z) = \frac{1}{z-2}.$$

Si vuole progettare $R(z)$ in modo che il sistema in anello chiuso sia asintoticamente stabile, la risposta di y^* ad uno scalino in y^{o*} si esaurisca in tempo minimo, senza errore a regime.

Poiché G^* ha grado relativo 1, anche F^* deve avere grado relativo 1. Si osservi che G^* presenta un polo a modulo maggiore di 1. Per poter imporre sia la condizione sulla non cancellazione di questo polo che la condizione sulla precisione statica (guadagno unitario di F^*), diamo a F^* la seguente espressione:

$$F^*(z) = \rho \frac{z+a}{z^2}.$$

Imponiamo quindi le due condizioni:

$$\begin{cases} F^*(1) = 1 & (\text{precisione statica}) \\ A(2) = B(2) & (\text{non cancellazione}) \end{cases}$$

ossia:

$$\begin{cases} \rho(1+a) = 1 \\ \rho(2+a) = 4 \end{cases} \Rightarrow \frac{2+a}{1+a} = 4 \Rightarrow a = -\frac{2}{3} \Rightarrow \rho = \frac{1}{1+a} = 3.$$

Pertanto:

$$F^*(z) = 3 \frac{z - \frac{2}{3}}{z^2} = \frac{3z-2}{z^2}.$$

Ne consegue la funzione di trasferimento del regolatore:

$$R(z) = \frac{\frac{3z-2}{z^2}}{\frac{1}{z-2} \left[1 - \frac{3z-2}{z^2} \right]} = \frac{(3z-2)(z-2)}{z^2 - 3z + 2} = \frac{(3z-2)(z-2)}{(z-1)(z-2)} = \frac{3z-2}{z-1}.$$

Si osservi la presenza nel regolatore di un polo in $z=1$ (azione integrale).