

Paolo Rocco

**Dispensa ad uso degli studenti del
Politecnico di Milano**

Fondamenti di Automatica

**E' vietato qualsiasi uso di questo materiale
non autorizzato dall'autore**

Avvertenza

Questa dispensa raccoglie, sotto forma di appunti sintetici, il materiale didattico per i corsi introduttivi di Fondamenti di Automatica.

La dispensa vuole essere un ausilio agli studenti per rivedere i propri appunti o per un rapido ripasso della materia, ma non sostituisce, né vuole sostituire, un testo organico¹ sull'Automatica, al quale si raccomanda di fare riferimento per un apprendimento più consapevole della materia.

Ultimo aggiornamento: Febbraio 2020

¹ Si segnala in particolare il testo:

“Fondamenti di Controlli Automatici” di P.Bolzern, R.Scattolini, N.Schiavoni, Mc-Graw Hill Italia, 4^a Ed.

Sommario

	pag.
Capitolo 1: Introduzione	5
Capitolo 2: Sistemi dinamici nel dominio del tempo	15
Capitolo 3: Sistemi dinamici nel dominio delle trasformate	57
Capitolo 4: Raggiungibilità e osservabilità	73
Capitolo 5: Risposte canoniche dei sistemi del primo e secondo ordine	83
Capitolo 6: Risposta in frequenza	97
Capitolo 7: Schemi a blocchi	111
Capitolo 8: Requisiti di un sistema di controllo	117
Capitolo 9: Stabilità dei sistemi di controllo	123
Capitolo 10: Prestazioni dinamiche dei sistemi di controllo	137
Capitolo 11: Prestazioni statiche dei sistemi di controllo	151
Capitolo 12: Progetto del controllore	159
Capitolo 13: Controllori PID	183
Capitolo 14: Schemi avanzati di controllo	191
Capitolo 15: Luogo delle radici	207
Capitolo 16: Assegnamento degli autovalori	215
Capitolo 17: Sistemi dinamici a tempo discreto	237
Capitolo 18: Controllo digitale	259

Capitolo 1: Introduzione

Con il termine automatica si fa riferimento ad una disciplina che studia tutti gli aspetti metodologici e concettuali che stanno alla base dell'**automazione**, ossia del trasferimento alle macchine di operazioni di governo e controllo di dispositivi, processi e sistemi di svariata natura. Si parla di automazione ogniqualvolta un'operazione viene eseguita da una macchina senza, o con ridotto, intervento dell'uomo.

I comparti applicativi in cui si presenta l'automazione sono i più svariati e toccano da vicino la vita quotidiana: si pensi agli elettrodomestici (frigoriferi, lavatrici, condizionatori), ai sistemi di frenatura e sterzo servoassistiti, alle sospensioni attive o al controllo della velocità di crociera nelle automobili, al pilota automatico negli aerei, ai processi manifatturieri automatizzati (fabbrica automatica), al controllo di motori elettrici, al controllo degli impianti per la generazione di energia, e così via.

Una tale vastità di applicazioni in cui l'automazione riveste un ruolo rilevante può far nascere il legittimo dubbio che l'automatica si riduca ad una rassegna o tutt'al più ad una classificazione delle applicazioni più significative.

In effetti inizialmente (al principio del ventesimo secolo) non vi era alcuna consapevolezza del carattere comune delle applicazioni di controllo. Le applicazioni, che pur esistevano (controllo di livello in serbatoi, controllo di velocità delle macchine a vapore, controllo del moto delle pale di mulini a vento), evolvevano in modo pionieristico e del tutto indipendente tra loro.

E' stato solo con il formarsi, e quindi con il consolidarsi, di una teoria matematica che l'automatica ha cominciato a prendere le forme di una disciplina scientifica. Tale teoria matematica va sotto il nome di **teoria dei sistemi**. Il suo indubbio pregio risiede nel fornire gli strumenti per lo studio delle caratteristiche del sistema, oggetto di automazione, in modo sostanzialmente indipendente dal contesto applicativo. Grazie alla teoria dei sistemi, tutti i sistemi di automazione elencati sommariamente in precedenza possono essere studiati con la stessa metodologia matematica.

Lo studio dei fondamenti della teoria dei sistemi, che occuperà la prima parte di questo corso, consentirà da un lato di dotarsi di strumenti molto efficaci per l'analisi di sistemi (non solo tecnologici, ma anche economici, ecologici o biologici) in cui è importante formalizzare l'evoluzione nel tempo delle variabili, dall'altro preparerà la strada allo studio dei sistemi di **controllo automatico**, che occuperà la seconda parte del corso. L'obiettivo primario dello studio sarà la valutazione oggettiva delle prestazioni dei sistemi di controllo, per mezzo di parametri che formalizzano concetti intuitivi, quali la stabilità, la velocità di risposta, la precisione del sistema di controllo. Saranno forniti anche elementi per la progettazione del dispositivo che esegue il controllo automatico e per la sua realizzazione in tecnologia digitale.

Il problema del controllo

Un problema di controllo nasce nel momento in cui si vuole imporre ad un “oggetto” (la cui natura va di volta in volta precisata) un comportamento desiderato, per mezzo di opportune azioni esercitate sull’oggetto stesso. Operiamo la seguente distinzione:

Controllo automatico: l’azione di controllo viene esercitata da dispositivi che operano in modo autonomo senza, o con ridotto, intervento umano;

Controllo manuale: l’azione di controllo viene esercitata dall’operatore umano.

Consideriamo un esempio tratto dalla vita quotidiana, la **guida di un automobile**.

Nella guida di un automobile intervengono una serie di problemi di controllo, molti dei quali tuttora risolti in modalità manuale.

Vediamone alcuni:

Controllo di percorso

- si vuole raggiungere una località di destinazione in un tempo desiderato;
- si può agire sulla traiettoria desiderata dell’automobile e sulla sua velocità desiderata.

Controllo di traiettoria

- si vuole fare in modo che l’automobile segua la traiettoria desiderata, non uscendo dalla carreggiata;
- si misura, a occhio, la posizione dell’automobile, rispetto ai limiti della carreggiata;
- si può agire sul volante.

Controllo di velocità

- si desidera che l’automobile viaggi a una velocità desiderata, senza collidere con altri veicoli;
- si misura la velocità dell’automobile tramite un tachimetro e, a occhio, la distanza dai veicoli che precedono;
- si può agire sui pedali del freno e dell’acceleratore.

Controllo di velocità del motore

- si desidera che il motore operi a numero di giri né troppo basso, né troppo alto;
- si misura la velocità del motore tramite il contagiri o semplicemente a orecchio;
- si può agire sulla leva del cambio.

Per la maggior parte, i problemi di controllo elencati sono tuttora risolti con controllo manuale, sebbene vi siano su alcune automobili sistemi automatizzati (cambio automatico, controllo della velocità di crociera, sistemi di assistenza alla guida) che risolvono almeno parzialmente i problemi di controllo.

Oltre a quelli appena citati, in un'automobile vi sono poi diversi altri dispositivi che funzionano in quanto dotati di un controllo automatico: il condizionatore, i tergicristalli, i sistemi di frenata assistita (ABS), il controllo di stabilità in curva, ecc.

Si osservi inoltre che sull'automobile agiscono una serie di altre variabili, quali le condizioni meteorologiche, le condizioni della strada, la presenza o meno di traffico, le condizioni dell'automobile stessa, che sono *indipendenti* dalla volontà di chi affronta il problema di controllo e largamente *incerte* (il loro valore non è noto a priori): tali variabili prendono il nome di **disturbi**.

Elementi di un problema di controllo

Quali sono gli elementi di un problema di controllo?

A) Il sistema sotto controllo

E' il sistema oggetto dell'azione di controllo. Su di esso agiscono delle *variabili manipolabili*, o *di controllo* (u), e dei *disturbi* (d) (variabili indipendenti e incerte), mentre le sue uscite (y) costituiscono le *variabili controllate* (di cui interessa cioè controllare l'andamento nel tempo).

B) L'andamento desiderato delle variabili controllate

Sono le variabili (y°) che esprimono l'andamento che le variabili controllate dovrebbero assumere per garantire un corretto funzionamento del sistema controllato. Verranno anche chiamate *referimenti* o *setpoint*.

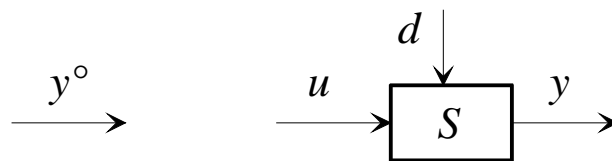


Fig. 1 : Elementi di un problema di controllo

Problema di controllo: determinare, ad ogni istante, il valore delle variabili di controllo u in modo tale che le variabili controllate y assumano un andamento quanto più possibile simile all'andamento desiderato y° , qualunque siano, tra quelli ritenuti ragionevoli, gli andamenti dei riferimenti y° e dei disturbi d .

Controllore: oggetto che determina ed esercita l'azione di controllo.

Legge di controllo: criterio secondo il quale agisce il controllore.

Un esempio: il frigorifero

Consideriamo il problema di controllo della temperatura in un frigorifero (Fig. 2). Come è noto, il principio di funzionamento del frigorifero si basa sulla presenza di una serpentina in cui circola un fluido che, sottoposto a cicli di compressione-espansione, sottrae calore all'ambiente interno del frigorifero. Il fluido circola quando il compressore è attivato da un motore elettrico, che può essere alimentato o meno a seconda della posizione di un interruttore.

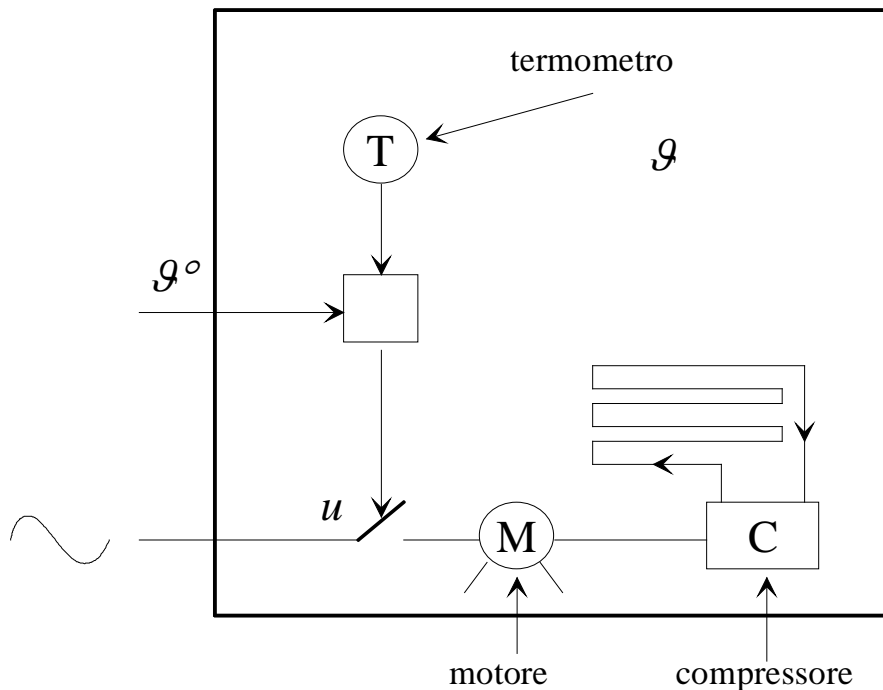


Fig. 2 : Un frigorifero

Obiettivo del controllo:	Mantenere approssimativamente costante la temperatura all'interno del frigo.
Riferimento:	g° valore desiderato per la temperatura all'interno del frigo (lo si imposta con una manopola).
Variabile di controllo:	u posizione (binaria: 0 oppure 1) dell'interruttore di alimentazione del motore del compressore.
Disturbi:	g_{d1} temperatura dell'ambiente esterno; g_{d2} temperatura degli oggetti inseriti.
Variabile controllata:	g temperatura all'interno del frigorifero (può essere misurata o no).

STRATEGIA DI CONTROLLO 1

Si calcola la quantità di calore che deve essere estratta per mantenere una certa temperatura

desiderata ϑ° . Servendosi di un timer, si accende e spegne il motore ad intervalli regolari.

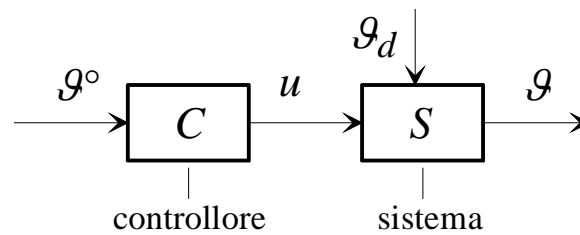


Fig. 3 : Strategia di controllo 1

Tipico andamento temporale della variabile di controllo u :

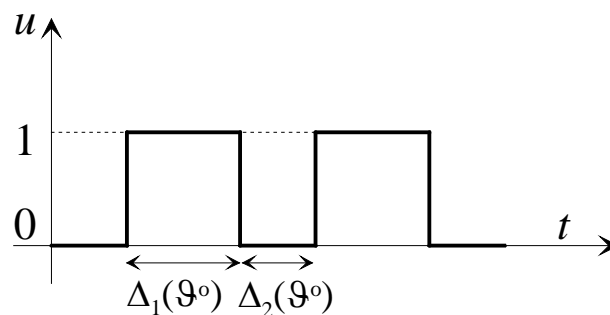


Fig. 4 : Posizione dell'interruttore

Osservazioni

- La legge di controllo si basa esclusivamente sul modello matematico (bilancio termico), per mezzo del quale si determinano gli intervalli di tempo di accensione del compressore (Δ_1) e di spegnimento (Δ_2) in base alla temperatura desiderata ϑ°
- Non è richiesto l'uso di un termometro
- Gli eventuali disturbi (porta del frigo lasciata a lungo aperta, oggetti inseriti particolarmente caldi, ecc.) compromettono l'efficacia della regolazione della temperatura.

STRATEGIA DI CONTROLLO 2

Si utilizza la misura ϑ^m della temperatura ϑ , fornita da un termometro.

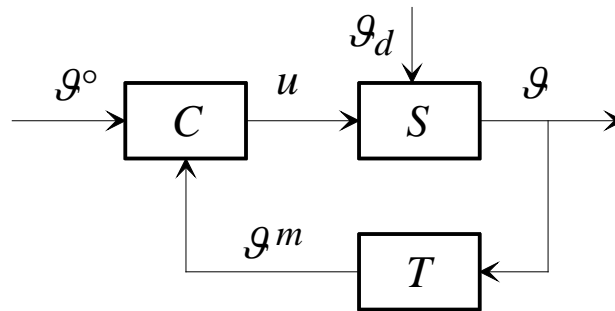


Fig. 5 : Strategia di controllo 2

Si alimenta il motore quando la differenza $\vartheta^m - \vartheta^\circ$ supera una certa soglia ε e lo si spegne quando tale differenza scende al di sotto di un'altra soglia $-\varepsilon$ (controllo a *relè*).

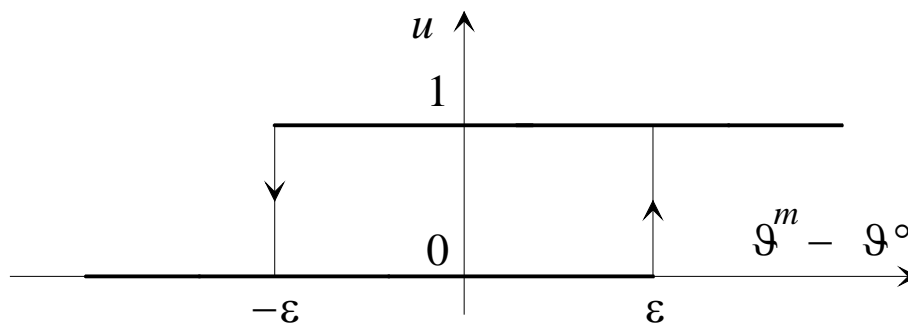


Fig. 6 : Controllore a relè

Osservazioni

- La legge di controllo non si basa sul modello. La taratura della soglia $\pm\varepsilon$ viene eseguita sulla base di prove empiriche
- E' richiesto l'uso di un termometro
- In presenza di eventuali disturbi la temperatura viene comunque regolata efficacemente.

Controllo in anello aperto e in anello chiuso

Generalizzando l'esempio del frigorifero, possiamo individuare due strategie di controllo:

Controllo in anello aperto (*feedforward control*)

Non viene eseguita alcuna misura sulle variabili del sistema, oppure le eventuali variabili misurate, ed utilizzate nella legge di controllo, non dipendono dai valori assunti dalla variabile di controllo u (strategia 1 nell'esempio precedente).

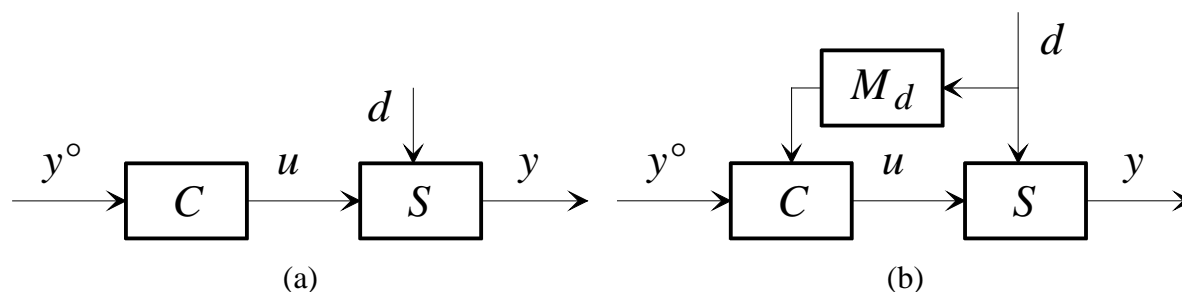


Fig. 7 : Schemi di controllo in anello aperto

Lo schema di Fig. 7b prende il nome di compensazione del disturbo: se il disturbo è misurabile, si esercita un'azione di controllo che dipende dalla misura del disturbo stesso

Controllo in anello chiuso (*feedback control*)

L'azione di controllo viene esercitata sulla base di misure di grandezze il cui valore dipende anche dal valore assunto dalla variabile u (strategia 2 nell'esempio precedente). In questo modo si viene a chiudere un anello nel rapporto di causa ed effetto tra le variabili (la variabile y dipende da u che, a sua volta, dipende da y ...): per questo si parla di "anello chiuso".

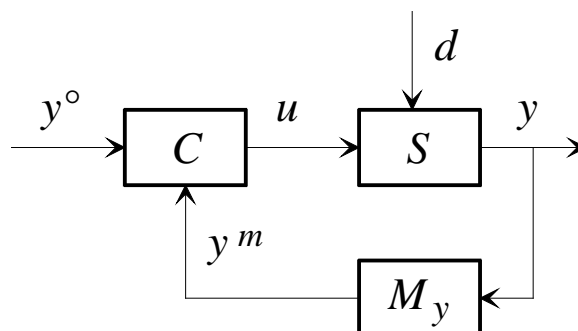


Fig. 8 : Schemi di controllo in anello chiuso

	Anello aperto	Anello chiuso
Misura di y	No	Sì
Modello matematico accurato	Sì	No
Sensibilità ai disturbi	Elevata	Bassa

Strumentazione

La strumentazione è costituita dai dispositivi (trasduttore e attuatore) che interfacciano il processo sotto controllo con il controllore.

Trasduttori: misurano una grandezza fisica del sistema sotto controllo (tipicamente la variabile controllata) e ne inviano la misura al controllore in una forma compatibile con la sua tecnologia.

Attuatori: traducono l'azione di controllo determinata dal controllore in un'azione efficace sul sistema, operando sulle sue variabili manipolabili (tipicamente con stadi intermedi di amplificazione e conversione di potenza).

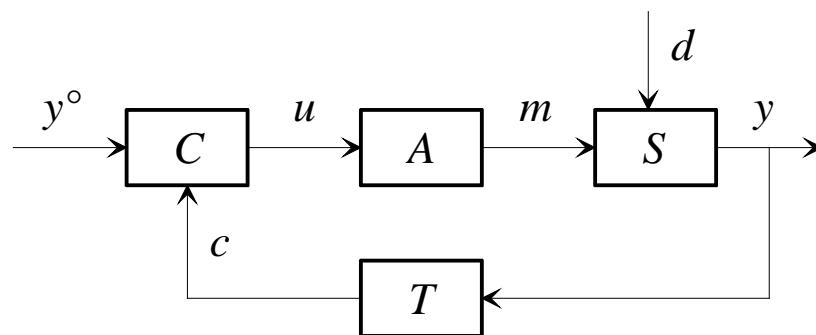


Fig. 9 : Schema di controllo completo di strumentazione

Si osservi che nello schema si è operata la distinzione tra la variabile di controllo u e la variabile manipolabile m e tra la variabile controllata y e la sua misura c .

Capitolo 2: Sistemi dinamici nel dominio del tempo

La necessità di adottare modelli dinamici nello studio dei sistemi di controllo può essere compresa facendo riferimento a un esempio di un nastro trasportatore di sabbia:

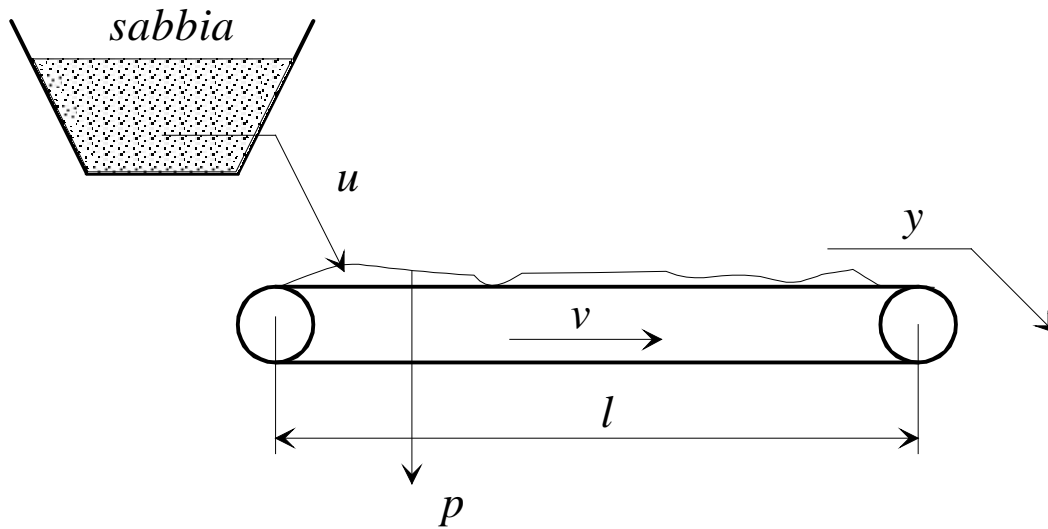


Fig. 1 : Un nastro trasportatore di sabbia

u : portata di sabbia all'inizio del nastro

y : portata di sabbia alla fine del nastro

p : perdite di sabbia lungo il nastro

v : velocità (costante) del nastro

l : lunghezza del nastro

Problema di controllo

Fare in modo che la portata y in uscita al nastro sia quanto più possibile simile ad un valore costante prefissato y° , nonostante le perdite p , agendo sulla portata u di sabbia all'ingresso del nastro.

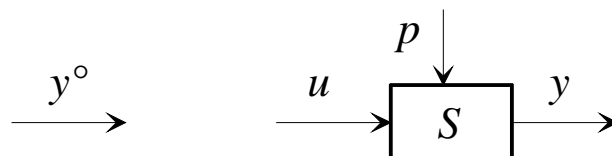


Fig. 2 : Il problema di controllo

Modello matematico

Il modello matematico traduce in un'equazione il fatto che, ad ogni istante di tempo t , la

portata in uscita uguaglia, a meno delle perdite, la portata manifestatasi in ingresso, τ istanti prima, dove τ è il tempo di percorrenza del nastro:

$$y(t) = u(t - \tau) - p(t), \quad \tau := l/v$$

Si suppone inoltre che le perdite siano calcolabili come la somma di un valore medio costante noto \bar{p} e di uno scostamento imprevedibile Δp :

$$p(t) = \bar{p} + \Delta p(t).$$

Strategia di controllo in anello aperto

La più ovvia strategia di controllo in anello aperto consiste nell'imporre un valore di portata in ingresso costante, uguale alla somma del valore desiderato in uscita e del valore medio delle perdite:

$$u(t) = y^\circ + \bar{p}.$$

Risulta però:

$$y(t) = y^\circ + \bar{p} - (\bar{p} + \Delta p(t)) = y^\circ - \Delta p(t),$$

ossia:

$$y^\circ - y(t) = \Delta p(t).$$

Pertanto il sistema di controllo è completamente “indifeso” rispetto al disturbo Δp (tutto il disturbo si traduce in errore).

Strategia di controllo in anello chiuso

Se la portata in uscita è misurabile, si somma alla precedente azione di controllo in anello aperto un termine correttivo, proporzionale all'errore tra valore desiderato ed effettivo di y :

$$u(t) = y^\circ + \bar{p} + \mu(y^\circ - y(t)),$$

dove μ è un parametro di progetto.

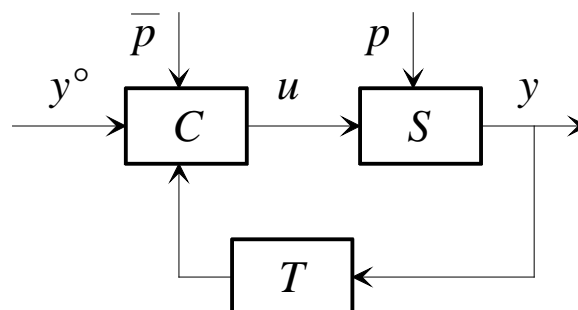


Fig. 3 : Strategia di controllo in anello chiuso

Risulta allora:

$$y(t) = y^\circ + \bar{p} + \mu(y^\circ - y(t - \tau)) - (\bar{p} + \Delta p(t)) = (1 + \mu)y^\circ - \mu y(t - \tau) - \Delta p(t).$$

Studiamo anzitutto il comportamento a regime (*analisi statica*), supponendo costanti le

perdite ($\Delta p(t) = \Delta \bar{p}$). Tutte le variabili risulteranno allora costanti, ed in particolare si avrà:

$$y(t) = y(t - \tau) = \bar{y}.$$

Facendo i conti si ottiene:

$$y^\circ - \bar{y} = \frac{\Delta \bar{p}}{1 + \mu}.$$

Sembra quindi che pur di scegliere il parametro μ positivo sufficientemente grande, si possa ridurre arbitrariamente l'errore.

Il problema è risolto? Non proprio...

Studiamo un *transitorio*, ossia il passaggio da una condizione di regime ad un'altra (**analisi dinamica**). In particolare, ipotizziamo che l'andamento nel tempo delle perdite sia rappresentato dal grafico di Fig. 4.



Fig. 4 : Andamento temporale delle perdite di sabbia

Studiamo il comportamento del sistema al trascorrere del tempo t :

$$\underline{t < \tau}$$

Supponiamo che il sistema parta dalla condizione di regime, che mantiene fino all'intervento della perturbazione all'istante τ :

$$y(t) = y^\circ$$

$$u(t) = y^\circ + \bar{p}$$

$$\underline{t = \tau}$$

La variazione istantanea di p determina una variazione istantanea di y (in base al modello) e quindi una variazione istantanea di u (in base alla legge di controllo):

$$y(\tau) = u(0) - p(\tau) = y^\circ + \bar{p}$$

$$u(\tau) = y^\circ + \bar{p} + \mu(y^\circ - y(\tau)) = y^\circ + \bar{p}(1 - \mu)$$

$$\tau < t < 2\tau$$

Sia y sia u rimangono costanti, perché la variazione di u all'istante τ si ripercuote su y solo dopo un intervallo di durata τ :

$$y(t) = y(\tau) = y^o + \bar{p}$$

$$u(t) = u(\tau) = y^o + \bar{p}(1 - \mu)$$

$$t = 2\tau$$

La variazione di u all'istante τ determina la variazione di y (in base al modello) e quindi una variazione di u (in base alla legge di controllo):

$$y(2\tau) = u(\tau) - p(2\tau) = y^o + \bar{p}(1 - \mu)$$

$$u(2\tau) = y^o + \bar{p} + \mu(y^o - y(2\tau)) = y^o + \bar{p}(1 - \mu + \mu^2)$$

Procedendo in questo modo, si conclude che y è **costante a tratti** e cambia valore agli istanti multipli di τ , secondo la legge:

$$y(k\tau) = y^o + \bar{p}(1 - \mu + \mu^2 - \mu^3 + \dots + (-1)^{k-1} \mu^{k-1}) = y^o + \bar{p} \sum_{i=0}^{k-1} (-\mu)^i$$

Il parametro μ influenza quindi pesantemente l'andamento temporale della portata in uscita y , come mostrano i seguenti grafici:

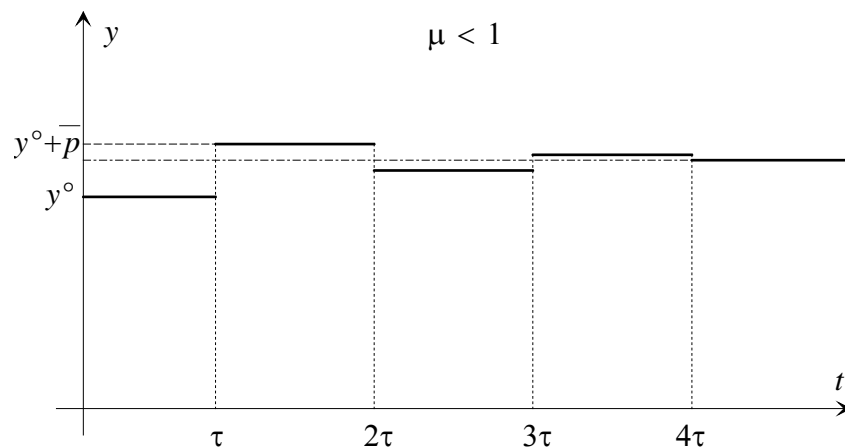


Fig. 5 : Andamento temporale della portata in uscita: $\mu < 1$

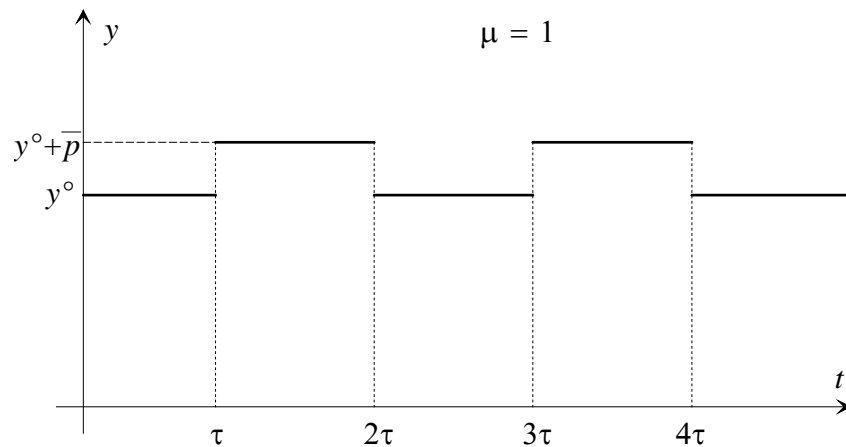


Fig. 6 : Andamento temporale della portata in uscita: $\mu=1$

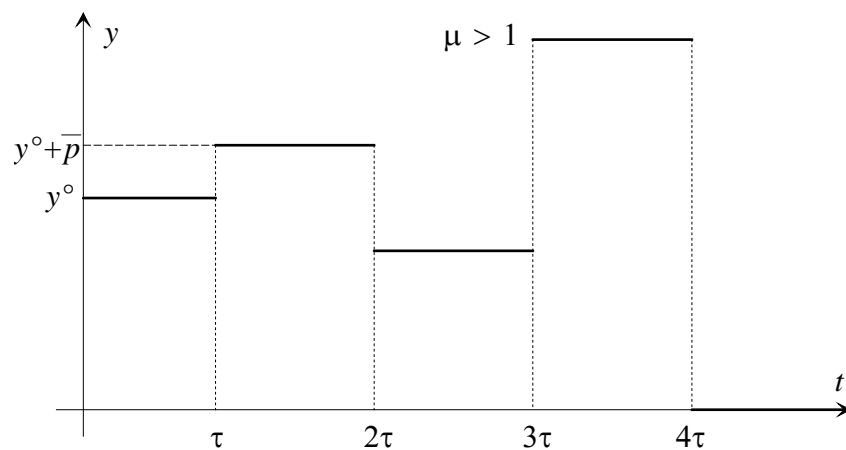


Fig. 7 : Andamento temporale della portata in uscita: $\mu > 1$

Tipo di transitorio	
$\mu < 1$	Oscillazioni convergenti (*)
$\mu = 1$	Oscillazioni permanenti
$\mu > 1$	Oscillazioni divergenti

(*) Dalle precedenti formule si trova che le oscillazioni convergono al valore $y^o + \bar{p}/(1+\mu)$, coerente con l'analisi statica, tenendo conto che nel nuovo punto di equilibrio $\Delta\bar{p} = -\bar{p}$.

Conclusioni

- L'analisi statica non è sufficiente per lo studio del comportamento dei sistemi di controllo. A volte (vedi i casi $\mu=1$ e $\mu>1$) può dare risultati inconsistenti, predicendo un errore a regime tanto più piccolo quanto più grande è il guadagno, quando in realtà la situazione di regime non può essere raggiunta (oscillazioni permanenti o divergenti).
- E' solo con un'analisi dinamica del sistema di controllo che è possibile comprenderne

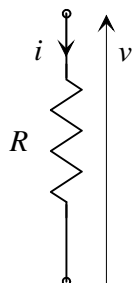
appieno il funzionamento.

- Un modello matematico che descrive l'evoluzione nel tempo delle variabili del sistema prende il nome di *modello dinamico*.
- Nel caso del nastro trasportatore, la dinamica del sistema era espressa da un ritardo di tempo. Nel prosieguo del corso, lo strumento matematico che useremo per formulare i modelli matematici sarà quello delle *equazioni differenziali*.

Modelli dinamici di sistemi elementari

Introduciamo alcuni modelli dinamici di sistemi elementari:

Resistore



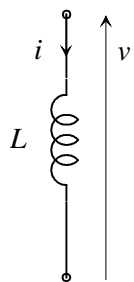
R : resistenza

i : corrente

v : tensione

$$v(t) = Ri(t)$$

Induttore



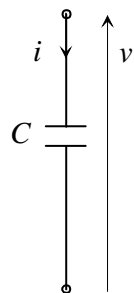
L : induttanza

i : corrente

v : tensione

$$v(t) = L \frac{di(t)}{dt}$$

Condensatore



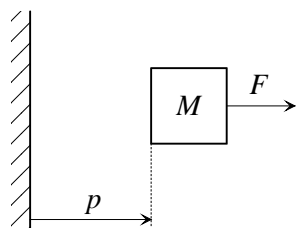
C : capacità

i : corrente

v : tensione

$$i(t) = C \frac{dv(t)}{dt}$$

Massa



M : massa

p : posizione

v : velocità

a : accelerazione

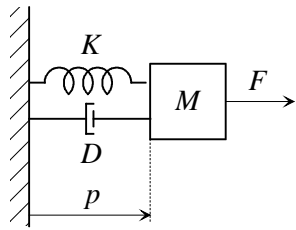
F : forza

$$v(t) = \frac{dp(t)}{dt}$$

$$a(t) = \frac{dv(t)}{dt}$$

$$F(t) = Ma(t)$$

Oscillatore meccanico



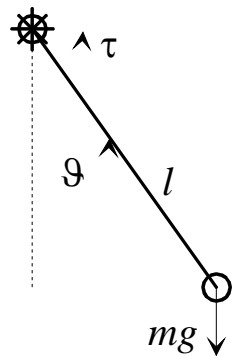
M : massa
 K : costante elastica
 D : coefficiente di attrito
 p : posizione
 v : velocità
 a : accelerazione
 F : forza

$$v(t) = \frac{dp(t)}{dt}$$

$$a(t) = \frac{dv(t)}{dt}$$

$$F(t) = Kp(t) + Dv(t) + Ma(t)$$

Pendolo



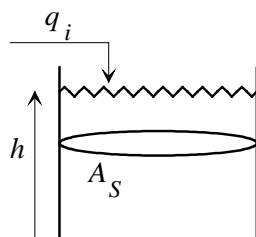
l : lunghezza dell'asta (priva di massa)
 m : massa concentrata
 g : accelerazione di gravità
 ϑ : posizione angolare
 ω : velocità angolare
 α : accelerazione angolare
 τ : coppia

$$\omega(t) = \frac{d\vartheta(t)}{dt}$$

$$\alpha(t) = \frac{d\omega(t)}{dt}$$

$$\tau(t) = ml^2\alpha(t) + mgl \sin(\vartheta(t))$$

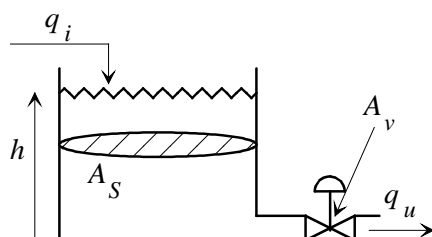
Serbatoio cilindrico



A_S : area sezione serbatoio
 h : livello liquido
 q_i : portata di liquido entrante

$$q_i(t) = A_S \frac{dh(t)}{dt}$$

Serbatoio cilindrico con valvola d'efflusso



A_S : area sezione serbatoio
 A_v : area di efflusso della valvola
 k : coefficiente caratteristico della valvola
 h : livello liquido
 q_i : portata di liquido entrante

$$q_i(t) = A_S \frac{dh(t)}{dt} + kA_v \sqrt{h(t)}$$

Sistemi dinamici

I modelli matematici precedentemente elencati possono tutti essere messi in una forma molto generale che prende il nome di sistema dinamico.

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 che influenzano il comportamento del sistema, di uscita quelle che caratterizzano il sistema e sulle quali soffermiamo il nostro interesse (tipicamente perché costituiscono l’obiettivo del controllo).

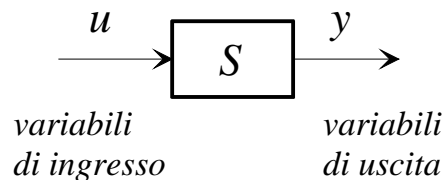


Fig. 8 : Ingressi e uscite di un sistema

La relazione che sussiste tra variabili di ingresso e di uscita è di causa-effetto e non ha nulla a che vedere con relazioni di afflusso ed efflusso di materia o energia (la portata di uscita in un serbatoio può essere variabile di ingresso per il sistema, se per esempio è comandata da una pompa).

E’ sufficiente descrivere il comportamento dinamico di un sistema mediante relazioni algebriche tra i suoi ingressi e le sue uscite? Quasi sempre no (nei nostri esempi, solo per il resistore), per due motivi: occorre conoscere i valori assunti dalle variabili di ingresso a partire dall’istante iniziale ed occorre conoscere una o più condizioni iniziali.

Consideriamo a titolo di esempio il condensatore, in cui l’ingresso è costituito dalla corrente ($u(t) = i(t)$), l’uscita dalla tensione ($y(t) = v(t)$). Avremo quindi:

$$C\dot{y}(t) = u(t) \Rightarrow y(t) = y(t_0) + \frac{1}{C} \int_{t_0}^t u(\tau) d\tau.$$

Occorre quindi conoscere il valore iniziale della tensione e l’andamento della corrente dall’istante iniziale. 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 indicheremo con n .

Per descrivere l’evoluzione dinamica del sistema è quindi sufficiente assegnare, istante per istante, n valori, ovvero dare l’andamento nel tempo di n variabili: indicheremo con x_1, x_2, \dots, x_n queste variabili e le definiremo **variabili di stato**.

Noto lo **stato** del sistema a un dato istante e l’andamento degli ingressi da quell’istante in poi, sarà quindi possibile determinare l’andamento di tutte le uscite dall’istante considerato. La formalizzazione matematica del sistema dinamico passa allora per la scrittura delle equazioni differenziali di cui le variabili di stato sono le soluzioni, noti gli ingressi esterni, e del legame tra le variabili di uscita e quelle di stato e di ingresso.

Sia m il numero delle variabili di ingresso e p il numero di variabili di uscita:

$$\left. \begin{aligned} \dot{x}_1(t) &= f_1(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\ \dot{x}_2(t) &= f_2(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\ &\vdots \\ \dot{x}_n(t) &= f_n(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \end{aligned} \right\} \text{equazioni di stato}$$

$$\left. \begin{aligned} y_1(t) &= g_1(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\ y_2(t) &= g_2(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\ &\vdots \\ y_p(t) &= g_p(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \end{aligned} \right\} \text{trasformazioni di uscita}$$

Queste sono le equazioni di un **sistema dinamico**.

Introduciamo i vettori:

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

e le funzioni vettoriali:

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

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

Possiamo riscrivere le equazioni del sistema dinamico in forma compatta vettoriale:

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

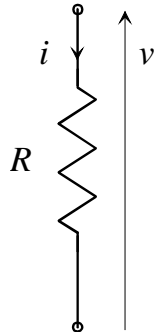
Si osservi che la dipendenza esplicita dal tempo delle due funzioni vettoriali rende il sistema genericamente *tempo variante*. In realtà nel presente corso ci limiteremo allo studio di sistemi in cui questa dipendenza esplicita non esista (sistemi *tempo invarianti* o *stazionari*). Per questi sistemi la scelta dell'asse dei tempi è del tutto convenzionale, ossia come istante iniziale sarà sempre possibile scegliere l'istante $t=0$.

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.

Torniamo ai nostri esempi:

Resistore



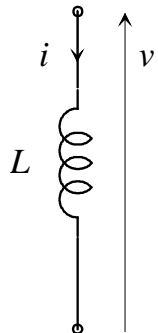
ingresso: $u = v$

uscita: $y = i$

variabili di stato: nessuna

$$y(t) = \frac{1}{R}u(t)$$

Induttore



ingresso: $u = v$

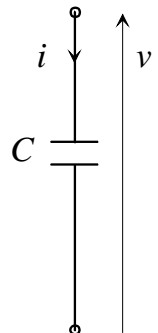
uscita: $y = i$

variabili di stato: $x_1 = i$

$$\dot{x}_1(t) = \frac{1}{L}u(t)$$

$$y(t) = x_1(t)$$

Condensatore



ingresso: $u = i$

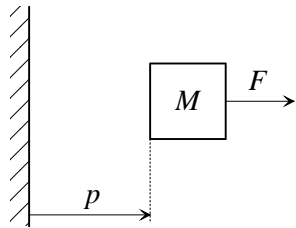
uscita: $y = v$

variabili di stato: $x_1 = v$

$$\dot{x}_1(t) = \frac{1}{C}u(t)$$

$$y(t) = x_1(t)$$

Massa



ingresso: $u = F$

uscita: $y = p$

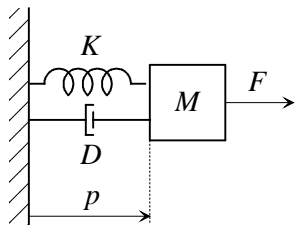
variabili di stato: $x_1 = p, x_2 = v$

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{1}{M}u(t)$$

$$y(t) = x_1(t)$$

Oscillatore meccanico



ingresso: $u = F$

uscita: $y = p$

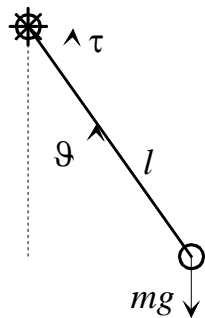
variabili di stato: $x_1 = p, x_2 = v$

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{1}{M}(-Kx_1(t) - Dx_2(t) + u(t))$$

$$y(t) = x_1(t)$$

Pendolo



ingresso: $u = \tau$

uscita: $y = \theta$

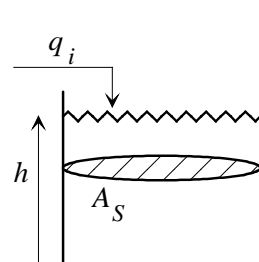
variabili di stato: $x_1 = \theta, x_2 = \omega$

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = -\frac{g}{l}\sin(x_1(t)) + \frac{1}{ml^2}u(t)$$

$$y(t) = x_1(t)$$

Serbatoio cilindrico



ingresso: $u = q_i$

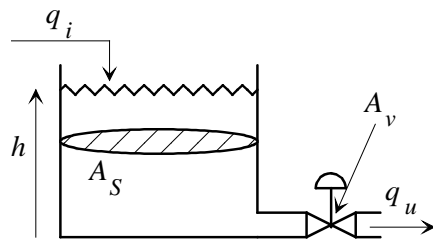
uscita: $y = h$

variabili di stato: $x_1 = h$

$$\dot{x}_1(t) = \frac{1}{A_S}u(t)$$

$$y(t) = x_1(t)$$

Serbatoio cilindrico con valvola d'efflusso



ingresso: $u = q_i$

uscita: $y = h$

variabili di stato: $x_1 = h$

$$\dot{x}_1(t) = -k \frac{A_v}{A_S} \sqrt{x_1(t)} + \frac{1}{A_S} u(t)$$

$$y(t) = x_1(t)$$

Gli esempi evidenziano che, di norma, le variabili di stato sono associate a fenomeni di accumulo (di energia elettrica, di energia potenziale, di energia cinetica, di massa...).

Sistemi dinamici lineari

Nei sistemi dinamici lineari le equazioni di stato e le trasformazioni di uscita sono lineari nelle variabili di stato e nelle variabili di ingresso. Se il sistema è anche tempo-invariante, potremo scrivere:

$$\left. \begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) + b_{11}u_1(t) + b_{12}u_2(t) + \dots + b_{1m}u_m(t) \\ \dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t) + b_{21}u_1(t) + b_{22}u_2(t) + \dots + b_{2m}u_m(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t) + b_{n1}u_1(t) + b_{n2}u_2(t) + \dots + b_{nm}u_m(t) \end{aligned} \right\} \text{equazioni di stato}$$

$$\left. \begin{aligned} y_1(t) &= c_{11}x_1(t) + c_{12}x_2(t) + \dots + c_{1n}x_n(t) + d_{11}u_1(t) + d_{12}u_2(t) + \dots + d_{1m}u_m(t) \\ y_2(t) &= c_{21}x_1(t) + c_{22}x_2(t) + \dots + c_{2n}x_n(t) + d_{21}u_1(t) + d_{22}u_2(t) + \dots + d_{2m}u_m(t) \\ &\vdots \\ y_p(t) &= c_{p1}x_1(t) + c_{p2}x_2(t) + \dots + c_{pn}x_n(t) + d_{p1}u_1(t) + d_{p2}u_2(t) + \dots + d_{pm}u_m(t) \end{aligned} \right\} \text{trasformazioni di uscita}$$

Introduciamo le matrici:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \dots & c_{pn} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1m} \\ d_{21} & d_{22} & \dots & d_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{p1} & d_{p2} & \dots & d_{pm} \end{bmatrix}$$

Il sistema dinamico lineare potrà allora essere riscritto in forma compatta vettoriale come segue:

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

Tutti i precedenti esempi sono descritti da sistemi dinamici lineari, tranne il pendolo (a causa della funzione trigonometrica) ed il serbatoio con valvola di efflusso (per via della radice quadrata).

Movimento

In un sistema dinamico il movimento (o moto) dello stato è definito come l'evoluzione nel tempo del vettore delle variabili di stato, a partire da un istante iniziale in cui sia dato il valore dello stato stesso, e noti gli andamenti degli ingressi da quell'istante in poi. Analoga definizione si dà per il movimento dell'uscita. Di fatto quindi il movimento dello stato costituisce la soluzione del sistema di equazioni differenziali che forma il sistema dinamico.

Formalmente, 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:

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

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

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

Equilibrio

Si supponga che l'ingresso (o gli ingressi) del sistema dinamico (lineare o no) siano costanti. Uno stato di equilibrio è un particolare movimento caratterizzato dal fatto che tutte le variabili di stato (e quindi anche le variabili di uscita) del sistema rimangono costanti nel tempo.

Consideriamo l'equazione di stato (vettoriale):

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

e assumiamo l'ingresso costante: $\mathbf{u}(t) = \bar{\mathbf{u}}$.

Se il sistema si trova all'equilibrio, $\mathbf{x}(t) = \bar{\mathbf{x}}$, e la derivata di \mathbf{x} è nulla. Pertanto:

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

Questa equazione, nell'incognita $\bar{\mathbf{x}}$, consente di trovare il punto di equilibrio del sistema. La corrispondente *uscita di equilibrio* sarà data da:

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

Non è detto che lo stato di equilibrio esista e, se esiste, non è detto che sia unico.

Esempio

Si consideri il sistema, non lineare, del secondo ordine:

$$\dot{x}_1 = x_1^3 + u$$

$$\dot{x}_2 = x_1 + x_2^2$$

$$y = x_1 x_2 + \sqrt{u}$$

Si vogliono individuare eventuali punti di equilibrio in corrispondenza dell'ingresso costante $\mathbf{u}(t) = \bar{\mathbf{u}} = 1$.

Annullando le derivate si ottiene:

$$\bar{x}_1^3 + 1 = 0$$

$$\bar{x}_1 + \bar{x}_2^2 = 0$$

Dalla prima equazione si ricava, come unica soluzione reale, $\bar{x}_1 = -1$, che, sostituita nella seconda, comporta le due soluzioni: $\bar{x}_2 = -1$ e $\bar{x}_2 = 1$.

Pertanto il sistema soggetto all'ingresso costante assegnato ammette due punti di equilibrio:

$$(\bar{x}_1 = -1; \bar{x}_2 = -1), (\bar{x}_1 = -1; \bar{x}_2 = 1).$$

In corrispondenza del primo punto di equilibrio l'uscita di equilibrio vale:

$$\bar{y} = \bar{x}_1 \bar{x}_2 + \sqrt{\bar{u}} = 2$$

mentre in corrispondenza del secondo:

$$\bar{y} = \bar{x}_1 \bar{x}_2 + \sqrt{\bar{u}} = 0.$$

Esempio

Si consideri un oscillatore meccanico:

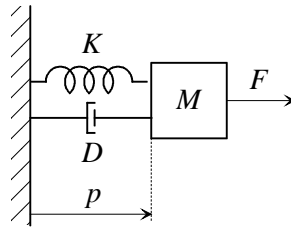


Fig. 9 : Oscillatore meccanico

Scelte le variabili:

ingresso: $u = F$

uscita: $y = p$

variabili di stato: $x_1 = p, x_2 = v$

le equazioni del sistema si scrivono:

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{1}{M}(-Kx_1(t) - Dx_2(t) + u(t))$$

$$y(t) = x_1(t)$$

Volendo determinare gli stati di equilibrio corrispondenti a una generica forza costante \bar{u} , dovremo risolvere il sistema:

$$\begin{cases} \bar{x}_2 = 0 \\ -K\bar{x}_1 - D\bar{x}_2 + \bar{u} = 0 \end{cases}$$

L'unica soluzione è:

$$\begin{cases} \bar{x}_1 = \frac{\bar{u}}{K} \\ \bar{x}_2 = 0 \end{cases}$$

che corrisponde a massa ferma, in una posizione tale per cui la forza di richiamo elastica equilibra la forza esterna applicata.

Esempio

Si consideri il pendolo nel piano verticale:

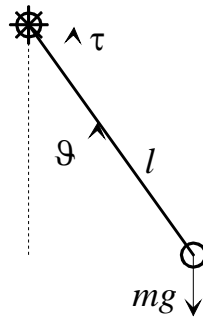


Fig. 10 : Pendolo nel piano verticale

Scelte le variabili:

ingresso: $u = \tau$

uscita: $y = \vartheta$

variabili di stato: $x_1 = \vartheta, x_2 = \omega$

le equazioni del sistema si scrivono:

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = -\frac{g}{l} \sin(x_1(t)) + \frac{1}{ml^2} u(t)$$

$$y(t) = x_1(t)$$

Volendo determinare gli stati di equilibrio corrispondenti a coppia in ingresso nulla, $\bar{u} = 0$, dovremo risolvere il sistema:

$$\begin{cases} \bar{x}_2 = 0 \\ -\frac{g}{l} \sin(\bar{x}_1) + \frac{1}{ml^2} \bar{u} = 0 \end{cases} \Rightarrow \begin{cases} \bar{x}_2 = 0 \\ \sin(\bar{x}_1) = 0 \end{cases}$$

Le soluzioni sono:

$$\begin{cases} \bar{x}_1 = 0 \\ \bar{x}_2 = 0 \end{cases} \quad \text{pendolo fermo rivolto verso il basso}$$

$$\begin{cases} \bar{x}_1 = \pi \\ \bar{x}_2 = 0 \end{cases} \quad \text{pendolo fermo rivolto verso l'alto}$$



Stabilità

La nozione di stabilità fa riferimento alla proprietà del sistema di reagire a perturbazioni ad esso applicate, ritornando alla condizione preesistente alla perturbazione, o quantomeno non allontanandosene indefinitamente. La definizione di stabilità che adotteremo 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.

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, $\forall \varepsilon > 0$, esiste in corrispondenza un valore $\delta_\varepsilon > 0$, tale che $\forall \mathbf{x}_{0p}$ che soddisfi la condizione:

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

risulti:

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

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:

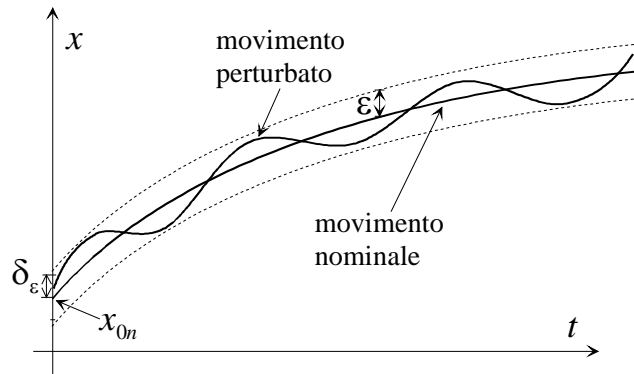


Fig. 11 : Movimento stabile

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.

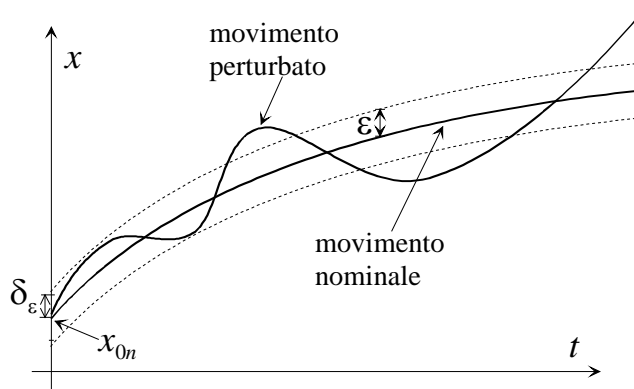


Fig. 12 : Movimento instabile

Infine un movimento $x_n(t)$ si dice **asintoticamente stabile** se, oltre ad essere stabile, soddisfa la condizione:

$$\lim_{t \rightarrow \infty} \|x_p(t) - x_n(t)\| = 0, \quad \forall x_{0p} : \|x_{0p} - x_{0n}\| < \delta_\varepsilon.$$

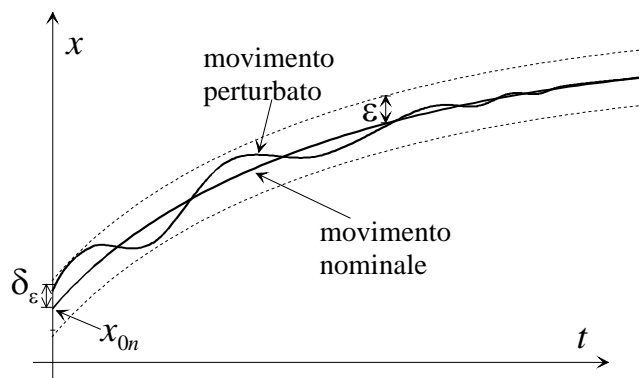


Fig.13 : Movimento asintoticamente stabile

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 di 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**.

Si consideri a titolo di esempio 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:

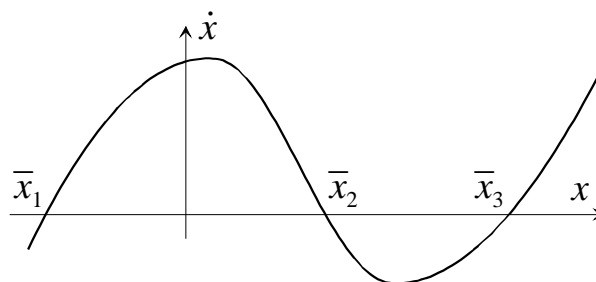


Fig. 14 : Derivata dello stato

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) .

Sistemi LTI: cambiamento di variabili di stato

Come già osservato, 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:

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

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.

Si consideri per esempio il circuito elettrico RLC di figura:

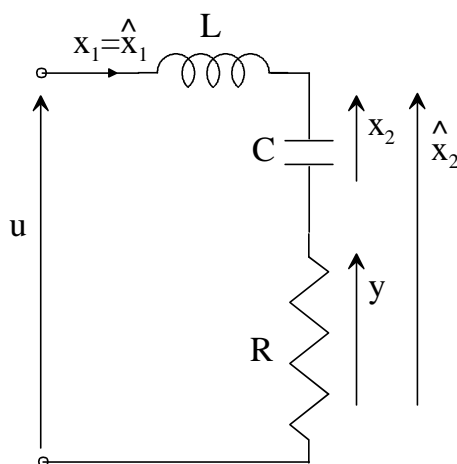


Fig. 15 : Differenti scelte di variabili di stato

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:

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

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:

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

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:

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

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

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

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

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} .

Sistemi LTI: equilibrio

Consideriamo un sistema dinamico lineare tempo invariante (LTI):

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

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

e supponiamo che il sistema sia soggetto all'ingresso costante: $\mathbf{u}(t) = \bar{\mathbf{u}}$

Gli stati di equilibrio sono soluzioni dell'equazione:

$$\mathbf{A}\bar{\mathbf{x}} + \mathbf{B}\bar{\mathbf{u}} = \mathbf{0}$$

Se la matrice \mathbf{A} è non singolare si ha uno e uno solo stato di equilibrio:

$$\bar{\mathbf{x}} = -\mathbf{A}^{-1}\mathbf{B}\bar{\mathbf{u}}$$

Pertanto:

$$\bar{\mathbf{y}} = \mu_s \bar{\mathbf{u}}$$

dove la quantità (scalare se il sistema è SISO):

$$\mu_s = -\mathbf{C}\mathbf{A}^{-1}\mathbf{B} + \mathbf{D}$$

è detta **guadagno statico** del sistema.

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 $\phi(\tau, t)$

(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:

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

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**.

Sovrapposizione degli effetti

Si consideri un sistema dinamico **lineare**.

Si eseguano sul sistema tre esperimenti:

1. Lo stato iniziale valga \mathbf{x}'_0 e si assegni l'ingresso $\mathbf{u}'(t)$, per $t \geq 0$.

Siano $\mathbf{x}'(t)$ e $\mathbf{y}'(t)$ i corrispondenti movimenti di stato e uscita.

2. Lo stato iniziale valga \mathbf{x}''_0 e si assegni l'ingresso $\mathbf{u}''(t)$, per $t \geq 0$.

Siano $\mathbf{x}''(t)$ e $\mathbf{y}''(t)$ i corrispondenti movimenti di stato e uscita.

3. Lo stato iniziale valga $\mathbf{x}'''_0 = \alpha \mathbf{x}'_0 + \beta \mathbf{x}''_0$ e si assegni l'ingresso $\mathbf{u}'''(t) = \alpha \mathbf{u}'(t) + \beta \mathbf{u}''(t)$, per $t \geq 0$, essendo α e β due arbitrari numeri reali.

Siano $\mathbf{x}'''(t)$ e $\mathbf{y}'''(t)$ i corrispondenti movimenti di stato e uscita.

Il principio di sovrapposizione degli effetti, valido solo per sistemi **lineari**, afferma che:

$$\mathbf{x}'''(t) = \alpha \mathbf{x}'(t) + \beta \mathbf{x}''(t) ,$$

$$\mathbf{y}'''(t) = \alpha \mathbf{y}'(t) + \beta \mathbf{y}''(t) .$$

E' quindi possibile studiare separatamente l'effetto sul moto delle cause (stato iniziale e differenti ingressi) che lo generano, e quindi sovrapporre (combinare linearmente) gli effetti.

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 \mathbf{x}_n preso in considerazione, l'analisi della sua stabilità conduce sempre all'analisi del comportamento delle soluzioni dell'equazione differenziale $\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t)$, al variare della condizione iniziale. Poiché il risultato di questa analisi è ovviamente lo stesso qualunque sia il movimento \mathbf{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 $\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t)$, 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 $\delta\mathbf{x}_0$) è la seguente:

$$\delta\mathbf{x}_l(t) = e^{\mathbf{A}t}\delta\mathbf{x}_0,$$

si deduce che la stabilità del sistema si analizzerà studiando le proprietà della matrice \mathbf{A} .

D'altra parte il fatto che il moto libero si scali proporzionalmente alla condizione iniziale \mathbf{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'equazione²:

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

(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

² 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:

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

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 \mathbf{A} :

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

L'espressione del moto libero è quindi:

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

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 \mathbf{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 :

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

Per formare una generica riga della tabella (la l -esima) ci si serve delle due righe immediatamente precedenti (la k -esima e la h -esima), 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 $n+1$ -esima 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:

$$1 \quad 1 \quad a$$

$$2 \quad 1 \quad 0$$

$$k_1 \quad k_2$$

$$h_1 \quad h_2$$

$$l_1$$

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:

$$1 \quad 1 \quad a$$

$$2 \quad 1 \quad 0$$

$$0.5 \quad a$$

$$1 - 4a \quad 0$$

$$a$$

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:

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

Linearizzazione

Considerando *piccoli scostamenti* delle variabili attorno a valori di equilibrio, è possibile approssimare il comportamento di un sistema dinamico non lineare con quello di un sistema dinamico lineare.

Consideriamo un generico sistema non lineare in forma vettoriale:

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

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

soggetto all'ingresso costante $\mathbf{u}(t) = \bar{\mathbf{u}}$. Supponiamo che esista il punto di equilibrio (eventualmente non unico) caratterizzato dal valore $\bar{\mathbf{x}}$ delle variabili di stato e dal valore $\bar{\mathbf{y}}$ dell'uscita di equilibrio.

Per definizione di equilibrio sarà quindi:

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

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

Si supponga ora che lo stato iniziale (all'istante $t=0$) sia costituito dal valore di equilibrio $\bar{\mathbf{x}}$ cui si somma un piccolo scostamento:

$$\mathbf{x}_0 = \bar{\mathbf{x}} + \delta\mathbf{x}_0,$$

e che, a partire dall'istante iniziale, l'ingresso si possa esprimere come la somma del valore all'equilibrio e di un piccolo scostamento:

$$\mathbf{u}(t) = \bar{\mathbf{u}} + \delta\mathbf{u}(t), \quad t \geq 0.$$

E' senz'altro lecito esprimere anche i movimenti di stato e uscita che ne conseguono come somma dei valori di equilibrio e di scostamenti:

$$\mathbf{x}(t) = \bar{\mathbf{x}} + \delta\mathbf{x}(t)$$

$$\mathbf{y}(t) = \bar{\mathbf{y}} + \delta\mathbf{y}(t).$$

Essendo le espressioni precedenti movimenti del sistema devono soddisfarne le equazioni. Si ottiene quindi:

$$\delta\dot{\mathbf{x}}(t) = \mathbf{f}(\bar{\mathbf{x}} + \delta\mathbf{x}(t), \bar{\mathbf{u}} + \delta\mathbf{u}(t))$$

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

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

Il **sistema linearizzato** si ottiene sviluppando in serie di Taylor intorno al punto di equilibrio le equazioni di stato e le trasformazioni di uscita del sistema originario ed arrestando lo sviluppo ai termini di primo grado. Nello sviluppo compariranno le derivate parziali delle funzioni vettoriali \mathbf{f} e \mathbf{g} rispetto agli argomenti vettoriali \mathbf{x} e \mathbf{u} (matrici Jacobiane), valutate nel punto di equilibrio:

$$\dot{\delta \mathbf{x}}(t) = \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}}) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}} \delta \mathbf{x}(t) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}} \delta \mathbf{u}(t)$$

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

$$\bar{\mathbf{y}} + \delta \mathbf{y}(t) = \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}}) + \left. \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}} \delta \mathbf{x}(t) + \left. \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}} \delta \mathbf{u}(t)$$

Ponendo ora:

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

$$\mathbf{C} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}}, \quad \mathbf{D} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right|_{\bar{\mathbf{x}}, \bar{\mathbf{u}}},$$

e ricordando le relazioni valide tra le variabili che caratterizzano l'equilibrio, otteniamo:

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

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

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

che è un sistema lineare.

Per il sistema linearizzato valgono quindi le proprietà dei sistemi lineari (non valide per il sistema non lineare di partenza), limitatamente a piccole variazioni intorno alla condizione di equilibrio.

Esempio

Si consideri nuovamente il sistema del secondo ordine:

$$\dot{x}_1 = x_1^3 + u$$

$$\dot{x}_2 = x_1 + x_2^2$$

$$y = x_1 x_2 + \sqrt{u}$$

Si vogliono determinare le espressioni dei sistemi linearizzati intorno ai due punti di equilibrio corrispondenti all'ingresso costante $u(t) = \bar{u} = 1$.

Abbiamo già calcolato i due punti di equilibrio:

$$(\bar{x}_1 = -1, \bar{x}_2 = -1), (\bar{x}_1 = -1, \bar{x}_2 = 1).$$

Le equazioni del sistema linearizzato sono le seguenti:

$$\begin{aligned}\delta\dot{x}_1 &= 3\bar{x}_1^2\delta x_1 + \delta u \\ \delta\dot{x}_2 &= \delta x_1 + 2\bar{x}_2\delta x_2 \\ \delta y &= \bar{x}_2\delta x_1 + \bar{x}_1\delta x_2 + \frac{1}{2}\frac{1}{\sqrt{u}}\delta u\end{aligned}.$$

In particolare, il sistema linearizzato intorno al primo punto di equilibrio risulta:

$$\begin{aligned}\delta\dot{x}_1 &= 3\delta x_1 + \delta u \\ \delta\dot{x}_2 &= \delta x_1 - 2\delta x_2 \\ \delta y &= -\delta x_1 - \delta x_2 + \frac{1}{2}\delta u\end{aligned},$$

mentre quello linearizzato intorno al secondo punto:

$$\begin{aligned}\delta\dot{x}_1 &= 3\delta x_1 + \delta u \\ \delta\dot{x}_2 &= \delta x_1 + 2\delta x_2 \\ \delta y &= \delta x_1 - \delta x_2 + \frac{1}{2}\delta u\end{aligned}.$$

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.

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 14.

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.

Capitolo 3: Sistemi dinamici nel dominio delle trasformate

Per il calcolo della risposta (uscita) di un sistema dinamico **lineare** soggetto ad ingressi assegnati, si possono seguire due strade.

Calcolo nel “dominio del tempo”

Con i metodi dell'analisi matematica, si integra il sistema di equazioni differenziali (equazioni di stato) forzato dalle funzioni del tempo assegnate (gli ingressi). Dalla trasformazione di uscita si ricava quindi l'espressione dell'uscita.

Calcolo nel “dominio delle trasformate”

Alla funzione del tempo $u(t)$ si associa, con i metodi matematici che vedremo, una funzione U che prende il nome di **trasformata** del segnale³ di ingresso. Dalle equazioni del sistema dinamico è poi possibile ricavare facilmente il legame tra la trasformata U e la trasformata Y del segnale di uscita. Ricavata quindi la trasformata Y , le si associa la funzione del tempo $y(t)$, che ne costituisce l'**antitrasformata**, e che rappresenta la risposta del sistema cercata.

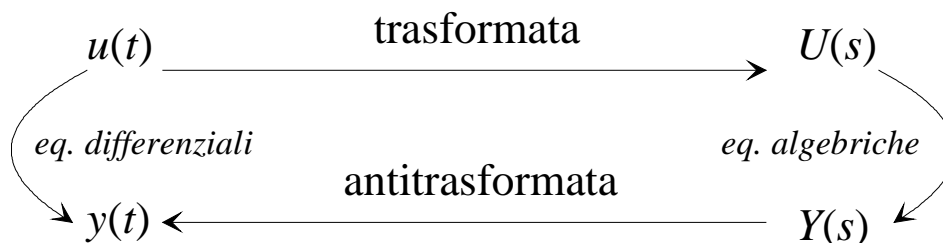


Fig. 16 : Calcolo della risposta di un sistema dinamico lineare

Qual è il vantaggio del metodo di calcolo nel dominio delle trasformate ?

Il vantaggio, notevolissimo, è che il legame tra la trasformata dell'ingresso e la trasformata dell'uscita è di natura **algebrica e non differenziale**, come accade invece tra le rispettive funzioni del tempo.

³Con il termine “segnale” intendiamo una variabile, scalare, funzione del tempo.

Trasformata di Laplace

Si consideri una funzione reale $f(t)$ della variabile reale t , definita per $t \geq 0$.

L'integrale, funzione della **variabile complessa** s :

$$F(s) = \int_0^{\infty} f(t) e^{-st} dt$$

è calcolabile per valori di s la cui parte reale sia maggiore di un certo valore scalare a , ovvero nel cosiddetto semipiano di convergenza ($s: \operatorname{Re}[s] > a$). Definiamo **trasformata di Laplace** di $f(t)$, indicandola anche con $\mathcal{L}[f(t)]$, l'unica funzione di s , definita e analitica quasi ovunque nel piano complesso, che nel semipiano di convergenza coincide con l'integrale $F(s)$

Esempio

Si consideri la funzione **scalino**:

$$f(t) = \text{sca}(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

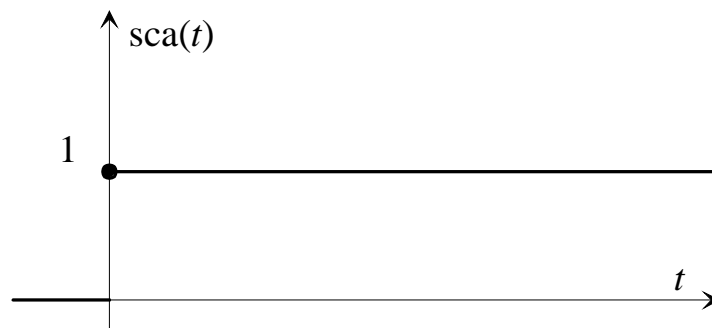


Fig. 17 : La funzione scalino

$$\mathcal{L}[\text{sca}(t)] = \int_0^{\infty} e^{-st} dt = \left[\frac{e^{-st}}{-s} \right]_0^{\infty} = \frac{1}{s}$$

Si noti che l'ultima eguaglianza è vera quando s è un numero complesso a parte reale positiva (cioè nel semipiano destro del piano complesso, che costituisce il semipiano di convergenza).

Esempio

Si consideri la funzione **impulso**:

$$f(t) = \text{imp}(t) = 0, \quad \forall t \neq 0$$

$$\int_{-\infty}^{+\infty} f(t) dt = 1$$

Tale funzione può essere vista come il limite, per $\varepsilon \rightarrow 0$, della seguente funzione:

$$f_{\varepsilon}(t) = \begin{cases} 1/\varepsilon & 0 \leq t \leq \varepsilon \\ 0 & t > \varepsilon \end{cases}$$

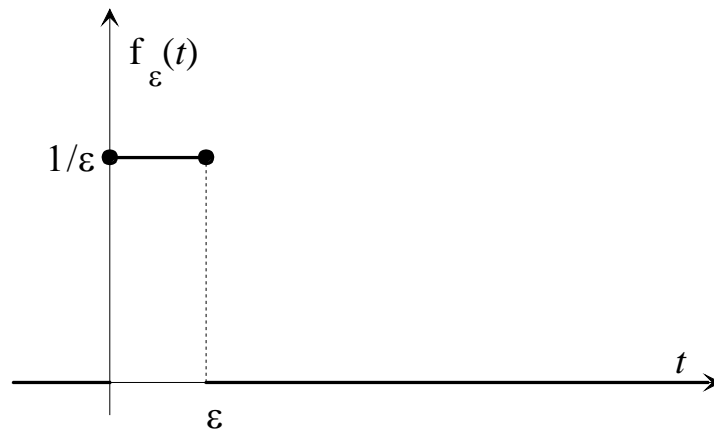


Fig. 18 : La funzione di cui l'impulso costituisce il limite

$$\begin{aligned} \mathcal{L}[\text{imp}(t)] &= \int_0^{\infty} \text{imp}(t) e^{-st} dt = \lim_{\varepsilon \rightarrow 0} \int_0^{\infty} f_{\varepsilon}(t) e^{-st} dt = \lim_{\varepsilon \rightarrow 0} \int_0^{\varepsilon} \frac{1}{\varepsilon} e^{-st} dt = \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left[\frac{e^{-st}}{-s} \right]_0^{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{1 - e^{-s\varepsilon}}{s\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{se^{-s\varepsilon}}{s} = 1 \end{aligned}$$

Proprietà notevoli della trasformata

- Linearità

$$\mathcal{L}[\alpha_1 f_1(t) + \alpha_2 f_2(t)] = \alpha_1 \mathcal{L}[f_1(t)] + \alpha_2 \mathcal{L}[f_2(t)].$$

- Traslazione nel dominio della variabile complessa

$$\text{Se } \mathcal{L}[f(t)] = F(s),$$

$$\text{allora } \mathcal{L}[e^{at} f(t)] = F(s - a).$$

- Traslazione nel dominio del tempo

$$\text{Se } \mathcal{L}[f(t)] = F(s),$$

$$\text{allora } \mathcal{L}[f(t - \tau)] = e^{-s\tau} F(s), \text{ per } \tau \geq 0.$$

- Derivazione nel dominio del tempo

$$\text{Se } \mathcal{L}[f(t)] = F(s),$$

$$\text{allora } \mathcal{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0^+).$$

- Derivazione nel dominio della variabile complessa

Se $\mathcal{L}[f(t)] = F(s)$,

allora $\mathcal{L}[tf(t)] = -\frac{dF(s)}{ds}$.

Trasformate notevoli

$f(t), t \geq 0$	$F(s)$
$\text{imp}(t)$	1
$\text{sca}(t)$	$1/s$
$\text{ram}(t)$	$1/s^2$
$\text{par}(t)$	$1/s^3$
e^{at}	$1/(s-a)$
$\sin(\omega t)$	$\omega/(s^2+\omega^2)$
$\cos(\omega t)$	$s/(s^2+\omega^2)$

dove:

$$\text{ram}(t) = \begin{cases} t & t \geq 0 \\ 0 & t < 0 \end{cases} \quad \text{par}(t) = \begin{cases} t^2/2 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Poli e zeri

I poli di una trasformata $F(s)$ sono i valori di s per cui $|F(s)| = \infty$.

Gli zeri di una trasformata $F(s)$ sono i valori di s per cui $F(s) = 0$.

Se $F(s)$ è *razionale*, ossia esprimibile come rapporto di due polinomi in s ,

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

i poli sono le radici del denominatore $D(s)$, gli zeri le radici del numeratore $N(s)$.

Antitrasformata di Laplace

Data una funzione $F(s)$ di variabile complessa, si vuole determinare la funzione $f(t)$ di cui $F(s)$ costituisce la trasformata.

I seguenti due teoremi forniscono informazioni parziali su $f(t)$.

Teorema del valore iniziale

Se $\mathcal{L}[f(t)] = F(s)$,

allora $f(0^+) = \lim_{s \rightarrow \infty} [sF(s)]$.

Se, ad esempio,

$$F(s) = \frac{s^2 + s + 1}{2s^3 + 3s^2 + 4},$$

allora $f(0^+) = \lim_{s \rightarrow \infty} \frac{s^3 + s^2 + s}{2s^3 + 3s^2 + 4} = \frac{1}{2}$.

Teorema del valore finale

Se $\mathcal{L}[f(t)] = F(s)$, e $F(s)$ è razionale e ha poli tutti a parte reale negativa oppure nell'origine del piano complesso, allora

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} [sF(s)].$$

Se, ad esempio,

$$F(s) = \frac{s^2 + s + 1}{2s^3 + 3s^2 + s},$$

$F(s)$ ha poli in 0, $-1/2$ e -1 , per cui il teorema è applicabile, e risulta:

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} \frac{s^2 + s + 1}{2s^2 + 3s + 1} = 1.$$

Metodo di Heaviside per funzioni razionali

Il metodo di Heaviside consente di ricavare l'espressione analitica dell'antitrasformata quando la trasformata è una funzione razionale, ossia un rapporto di polinomi in s :

$$F(s) = \frac{N(s)}{D(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \dots + b_n}{s^n + a_1 s^{n-1} + \dots + a_n}.$$

Il metodo viene qui presentato solo per alcuni casi particolari.

1) Poli reali semplici

Il denominatore è fattorizzabile come:

$$D(s) = (s + p_1)(s + p_2) \cdots (s + p_n), \quad p_i \in \mathfrak{R}, \quad p_i \neq p_j.$$

Ne consegue:

$$F(s) = \frac{\alpha_1}{s + p_1} + \frac{\alpha_2}{s + p_2} + \cdots + \frac{\alpha_n}{s + p_n}$$

Si ricavano i coefficienti $\alpha_1, \dots, \alpha_n$ mediante confronto tra questa espressione e l'espressione originaria di $F(s)$. Infine si antitrasformano i singoli termini:

$$f(t) = \alpha_1 e^{-p_1 t} + \alpha_2 e^{-p_2 t} + \cdots + \alpha_n e^{-p_n t}, \quad t \geq 0.$$

2) Poli reali semplici e un polo reale multiplo

Il denominatore è fattorizzabile come:

$$D(s) = (s + p_1)^k (s + p_2) \cdots (s + p_m), \quad p_i \in \mathfrak{R}, \quad p_i \neq p_j, \quad m = n - (k - 1).$$

Ne consegue:

$$F(s) = \frac{\alpha_{1k}}{(s + p_1)^k} + \frac{\alpha_{1(k-1)}}{(s + p_1)^{k-1}} + \cdots + \frac{\alpha_{11}}{s + p_1} + \frac{\alpha_2}{s + p_2} + \cdots + \frac{\alpha_m}{s + p_m}$$

Si ricavano i coefficienti $\alpha_{1k}, \dots, \alpha_{11}, \alpha_2, \dots, \alpha_m$ mediante confronto tra questa espressione e l'espressione originaria di $F(s)$. Infine si antitrasformano i singoli termini:

$$f(t) = \alpha_{1k} \frac{t^{k-1}}{(k-1)!} e^{-p_1 t} + \alpha_{1(k-1)} \frac{t^{k-2}}{(k-2)!} e^{-p_1 t} + \cdots + \alpha_{11} e^{-p_1 t} + \alpha_2 e^{-p_2 t} + \cdots + \alpha_m e^{-p_m t}, \quad t \geq 0.$$

3) Poli reali semplici e due poli complessi e coniugati semplici

Il denominatore è fattorizzabile come:

$$D(s) = (s + p_1)(s + \bar{p}_1)(s + p_2) \cdots (s + p_m), \quad p_i \in \mathfrak{R} \ (i \geq 2), \quad p_i \neq p_j, \quad m = n - 1.$$

Ne consegue (posto $p_1 = \sigma + j\omega$):

$$\begin{aligned} F(s) &= \frac{\alpha_1}{s + p_1} + \frac{\bar{\alpha}_1}{s + \bar{p}_1} + \frac{\alpha_2}{s + p_2} + \cdots + \frac{\alpha_m}{s + p_m} = \frac{\beta s + \gamma}{s^2 + 2\sigma s + \sigma^2 + \omega^2} + \frac{\alpha_2}{s + p_2} + \cdots + \frac{\alpha_m}{s + p_m} \\ &= \beta \frac{s + \sigma}{(s + \sigma)^2 + \omega^2} + \frac{-\beta\sigma + \gamma}{\omega} \frac{\omega}{(s + \sigma)^2 + \omega^2} + \frac{\alpha_2}{s + p_2} + \cdots + \frac{\alpha_m}{s + p_m} \end{aligned}$$

con β e γ parametri reali opportuni. Si ricavano i coefficienti $\beta, \gamma, \alpha_2, \dots, \alpha_m$ mediante confronto tra questa espressione e l'espressione originaria di $F(s)$. Infine si antitrasformano i singoli termini:

$$f(t) = \beta e^{-\sigma t} \cos(\omega t) + \frac{-\beta\sigma + \gamma}{\omega} e^{-\sigma t} \sin(\omega t) + \alpha_2 e^{-p_2 t} + \cdots + \alpha_m e^{-p_m t}, \quad t \geq 0.$$

Funzione di trasferimento

Consideriamo un sistema dinamico lineare in forma vettoriale:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)\end{aligned}, \quad \mathbf{x} \in \mathfrak{R}^n, \mathbf{u} \in \mathfrak{R}^m, \mathbf{y} \in \mathfrak{R}^p$$

Introduciamo i vettori $\mathbf{X}(s)$, $\mathbf{U}(s)$, $\mathbf{Y}(s)$ che contengono le trasformate di Laplace delle componenti dei vettori $\mathbf{x}(t)$, $\mathbf{u}(t)$, $\mathbf{y}(t)$, rispettivamente.

Osservando che risulta:

$$\begin{aligned}\mathcal{L}[\dot{\mathbf{x}}(t)] &= \begin{bmatrix} \mathcal{L}[\dot{x}_1(t)] \\ \mathcal{L}[\dot{x}_2(t)] \\ \vdots \\ \mathcal{L}[\dot{x}_n(t)] \end{bmatrix} = \begin{bmatrix} sX_1(s) - x_1(0) \\ sX_2(s) - x_2(0) \\ \vdots \\ sX_n(s) - x_n(0) \end{bmatrix} = s\mathbf{X}(s) - \mathbf{x}(0) \\ \mathcal{L}[\mathbf{A}\mathbf{x}(t)] &= \begin{bmatrix} \mathcal{L}[a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t)] \\ \mathcal{L}[a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t)] \\ \vdots \\ \mathcal{L}[a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t)] \end{bmatrix} = \begin{bmatrix} a_{11}X_1(s) + a_{12}X_2(s) + \dots + a_{1n}X_n(s) \\ a_{21}X_1(s) + a_{22}X_2(s) + \dots + a_{2n}X_n(s) \\ \vdots \\ a_{n1}X_1(s) + a_{n2}X_2(s) + \dots + a_{nn}X_n(s) \end{bmatrix} = \mathbf{A}\mathbf{X}(s)\end{aligned}$$

e analogamente per le altre trasformate di prodotti matrice-vettore, si ottiene, sfruttando la linearità della trasformata:

$$\begin{aligned}s\mathbf{X}(s) - \mathbf{x}(0) &= \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) &= \mathbf{C}\mathbf{X}(s) + \mathbf{D}\mathbf{U}(s)\end{aligned}$$

Si è quindi ottenuto un sistema **algebrico** nelle trasformate delle variabili. Per tutti i valori di s diversi dagli autovalori della matrice \mathbf{A} , risulta:

$$\mathbf{X}(s) = (s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B}\mathbf{U}(s) + (s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{x}(0)$$

e quindi:

$$\mathbf{Y}(s) = \left[\mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \right] \mathbf{U}(s) + \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{x}(0).$$

Di particolare interesse è la situazione in cui lo stato iniziale è nullo ($\mathbf{x}(0) = 0$). Risulta:

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{U}(s),$$

dove la matrice (di dimensioni $p \times m$):

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

prende il nome di **funzione di trasferimento** del sistema.

Nel caso SISO ($m=p=1$), la funzione di trasferimento diventa uno scalare e si può scrivere, sempre a stato iniziale nullo:

$$G(s) = \frac{Y(s)}{U(s)}.$$

La funzione di trasferimento può essere calcolata con la formula precedente (ricavando quindi le matrici A , B , C , D e invertendo una matrice $n \times n$) oppure trasformando le singole equazioni membro a membro (a stato iniziale nullo), e ricavando il legame tra $Y(s)$ e $U(s)$ mediante eliminazione delle $X_i(s)$.

Riprendiamo gli esempi di sistemi dinamici elementari trattati in precedenza, limitandoci naturalmente a quelli lineari:

Resistore

$$y(t) = \frac{1}{R} u(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{R}$$

Induttore

$$\dot{x}_1(t) = \frac{1}{L} u(t)$$

$$y(t) = x_1(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{Ls}$$

Condensatore

$$\dot{x}_1(t) = \frac{1}{C} u(t)$$

$$y(t) = x_1(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{Cs}$$

Massa

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{1}{M} u(t)$$

$$y(t) = x_1(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{Ms^2}$$

Oscillatore meccanico

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{1}{M} (-Kx_1(t) - Dx_2(t) + u(t))$$

$$y(t) = x_1(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{Ms^2 + Ds + K}$$

Serbatoio cilindrico

$$\dot{x}_1(t) = \frac{1}{A_S} u(t)$$

$$y(t) = x_1(t) \Rightarrow G(s) = \frac{Y(s)}{U(s)} = \frac{1}{A_S s}$$

Invarianza della funzione di trasferimento

Dato un sistema dinamico LTI descritto dalle equazioni:

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

abbiamo visto 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:

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

dove:

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

Calcoliamo la funzione di trasferimento per questo sistema:

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

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

Struttura della funzione di trasferimento

Si consideri un sistema SISO, per cui la funzione di trasferimento è uno scalare:

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

Osserviamo che:

$$(s\mathbf{I}_n - \mathbf{A})^{-1} = \frac{1}{\det(s\mathbf{I}_n - \mathbf{A})} \begin{bmatrix} k_{11}(s) & k_{12}(s) & \cdots & k_{1n}(s) \\ k_{21}(s) & k_{22}(s) & \cdots & k_{2n}(s) \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1}(s) & k_{n2}(s) & \cdots & k_{nn}(s) \end{bmatrix},$$

dove i polinomi $k_{ij}(s)$ sono i complementi algebrici della matrice $(s\mathbf{I}_n - \mathbf{A})$ e hanno, per costruzione, grado non superiore a $n-1$ (mentre il determinante a denominatore ha ovviamente grado n).

Nel formare lo scalare $C(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B}$ si combinano linearmente i polinomi k_{ij} , ottenendo un polinomio che non può avere grado maggiore dei singoli polinomi. A questa espressione va poi sommato D , se il sistema non è strettamente proprio.

Concludiamo quindi che la funzione di trasferimento è razionale (rapporto di polinomi):

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

che il denominatore $D(s)$ ha grado n , mentre per il numeratore:

$$N(s) = \begin{cases} \text{polinomio di grado } \leq (n-1), & \text{se il sistema è strett. proprio } (D=0) \\ \text{polinomio di grado } = n, & \text{se il sistema non è strett. proprio } (D \neq 0) \end{cases}$$

Si osservi quindi che il grado del numeratore non può mai eccedere quello del denominatore.

Si ricorda inoltre che un polinomio di grado n a coefficienti reali ammette nel piano complesso n radici, reali o a coppie complesse e coniugate (teorema fondamentale dell'algebra).

Gli **zeri** della funzione di trasferimento sono le radici del numeratore $N(s)$ (e quindi sono in numero minore o uguale a n).

I **poli** della funzione di trasferimento sono le radici del denominatore $D(s)$ (e quindi sono in numero uguale a n). I poli, in quanto radici del determinante della matrice $(s\mathbf{I}_n - \mathbf{A})$, coincidono con gli **autovalori** della matrice \mathbf{A} .

Queste conclusioni non contemplano esplicitamente il caso in cui numeratore e denominatore abbiano una o più radici comuni. Nel formare l'espressione della funzione di trasferimento tali radici si semplificano, per cui il denominatore avrà grado minore di n (e il numeratore grado minore o uguale a quello del denominatore). In questo caso i poli della funzione di trasferimento formano un sottoinsieme degli autovalori della matrice \mathbf{A} .

Nel piano complesso, i poli vengono di norma rappresentati con una crocetta, gli zeri con un pallino. La funzione di trasferimento:

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

presenta due zeri, in $s = -j$ e $s = j$, e tre poli, in $s = 0$, $s = -1$ e $s = -2$, rappresentati come in figura:

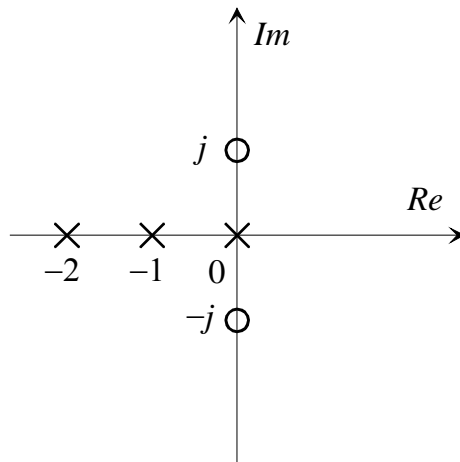


Fig. 19 : Disposizione di poli e zeri.

Parametri caratteristici della funzione di trasferimento

Si è visto che la funzione di trasferimento di un sistema dinamico è una funzione razionale della variabile complessa s , ossia è il rapporto di due polinomi:

$$G(s) = \frac{N(s)}{D(s)} = \frac{\beta_0 s^n + \beta_1 s^{n-1} + \dots + \beta_n}{s^n + \gamma_1 s^{n-1} + \dots + \gamma_n}.$$

Alternativamente si può utilizzare la seguente espressione equivalente:

$$G(s) = \rho \frac{\prod_i (s + z_i)}{\prod_i (s + p_i)}.$$

dove le produttorie corrono su tutti gli zeri e su tutti i poli, rispettivamente, mentre:

ρ : **costante di trasferimento**

$-z_i$: **zeri**

$-p_i$: **poli**

Si osservi che i parametri z_i e p_i possono anche essere complessi. Per ottenere una rappresentazione con solo numeri reali è sufficiente accoppiare i termini complessi e coniugati (a numeratore e a denominatore), nei polinomi di secondo grado a radici complesse. Questi polinomi, a loro volta sono espressi per mezzo di due parametri particolarmente significativi, indicati con ζ e ω_n :

$$s^2 + 2\zeta\omega_n s + \omega_n^2,$$

dove ω_n è un numero positivo.

Per comprendere il significato dei due parametri, osserviamo che le radici del polinomio sono:

$$s_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2},$$

e risultano effettivamente complesse e coniugate se $|\zeta| < 1$.

Il significato dei parametri ζ e ω_n è allora illustrato dalla seguente figura:

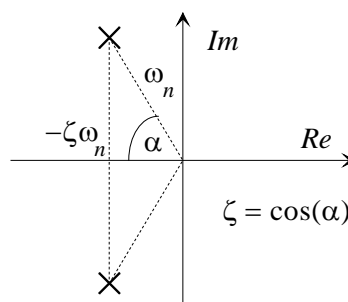


Fig. 20 : Significato dei parametri ζ e ω_n

ω_n , **pulsazione naturale**: è il modulo delle due radici, ossia la loro distanza dall'origine.

ζ , **smorzamento**: è il coseno dell'angolo α formato dalla congiungente l'origine con le radici, rispetto al semiasse reale negativo

Poiché la parte reale dei poli vale $-\zeta\omega_n$ e ω_n è un numero positivo, si ha:

$\zeta > 0$: due radici nel semipiano sinistro

$\zeta = 0$: due radici sull'asse immaginario

$\zeta < 0$: due radici nel semipiano destro

Possiamo a questo punto esprimere la funzione di trasferimento per mezzo di soli parametri reali nella seguente forma:

$$G(s) = p \frac{\prod_i (s + z_i) \prod_i (s^2 + 2\zeta_{zi}\omega_{nzi}s + \omega_{nzi}^2)}{\prod_i (s + p_i) \prod_i (s^2 + 2\zeta_{pi}\omega_{npi}s + \omega_{npi}^2)},$$

con $\omega_{nzi}, \omega_{npi} > 0$, $|\zeta_{zi}|, |\zeta_{pi}| \leq 1$.

Un'ulteriore espressione della funzione di trasferimento è la seguente:

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1 + s\tau_i)}{\prod_i (1 + sT_i)}$$

dove le produttorie corrono su tutti gli zeri e su tutti i poli diversi da zero, rispettivamente, mentre:

μ : **guadagno**

g : **tipo**

τ_i : **costanti di tempo degli zeri**

T_i : **costanti di tempo dei poli**

Si osservi che il rapporto delle due produttorie valutato in $s = 0$ è pari a 1. Per ottenere questo risultato si sono raggruppati gli eventuali poli o zeri in $s = 0$ nel termine a denominatore s^g . Pertanto g è un numero intero, uguale, se positivo, al numero di poli in $s=0$, se negativo, al numero di zeri in $s=0$ (se è nullo non vi sono né poli né zeri in $s=0$).

Se $g=0$, risulta inoltre:

$$\mu = \lim_{s \rightarrow 0} G(s) = G(0) = -\mathbf{CA}^{-1}\mathbf{B} + \mathbf{D},$$

espressione che prende il nome, come già discusso, di **guadagno statico**, in quanto corrisponde al rapporto tra ingresso e uscita all'equilibrio.

Più in generale:

$$\mu = \lim_{s \rightarrow 0} [s^g G(s)].$$

Anche questa forma della funzione di trasferimento può essere espressa in termini solo di parametri reali:

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1 + s\tau_i) \prod_i \left(1 + 2 \frac{\zeta_{zi}}{\omega_{nzi}} s + \frac{s^2}{\omega_{nzi}^2} \right)}{\prod_i (1 + sT_i) \prod_i \left(1 + 2 \frac{\zeta_{pi}}{\omega_{npi}} s + \frac{s^2}{\omega_{npi}^2} \right)},$$

con $\omega_{nzi}, \omega_{npi} > 0$, $|\zeta_{zi}|, |\zeta_{pi}| \leq 1$.

Calcolo delle risposte temporali

Dato un sistema dinamico lineare e un ingresso trasformabile secondo Laplace, è possibile ricavare l'espressione analitica dell'uscita del sistema dinamico forzata da tale ingresso. Occorre:

1. Ricavare, se non è già data, la funzione di trasferimento $G(s)$ del sistema
2. Ricavare la trasformata $U(s)$ dell'ingresso
3. Calcolare la trasformata dell'uscita $Y(s) = G(s)U(s)$
4. Antitrasformare

Sia ad esempio:

$$G(s) = \frac{2s+1}{s^2+5s+4}, \quad u(t) = \text{sca}(t).$$

Sappiamo allora che:

$$U(s) = \frac{1}{s}, \quad Y(s) = G(s)U(s) = \frac{2s+1}{s(s^2+5s+4)} = \frac{2s+1}{s(s+1)(s+4)}$$

Applichiamo il metodo di Heaviside per l'antitrasformazione di $Y(s)$:

$$Y(s) = \frac{\alpha_1}{s} + \frac{\alpha_2}{s+1} + \frac{\alpha_3}{s+4} = \frac{\alpha_1(s+1)(s+4) + \alpha_2s(s+4) + \alpha_3s(s+1)}{s(s+1)(s+4)} = \frac{2s+1}{s(s+1)(s+4)}.$$

Imponendo l'uguaglianza dei due numeratori, in particolare nei punti $s = 0$, $s = -1$, $s = -4$, si ottiene:

$$\left. \begin{array}{l} 4\alpha_1 = 1 \\ -3\alpha_2 = -1 \\ 12\alpha_3 = -7 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \alpha_1 = 1/4 \\ \alpha_2 = 1/3 \\ \alpha_3 = -7/12 \end{array} \right.$$

Pertanto:

$$y(t) = \frac{1}{4} \text{sca}(t) + \frac{1}{3} e^{-t} - \frac{7}{12} e^{-4t} = \frac{1}{4} + \frac{1}{3} e^{-t} - \frac{7}{12} e^{-4t}, \quad t \geq 0.$$

Capitolo 4: Raggiungibilità e osservabilità

Facendo riferimento ad un sistema SISO, per il quale la funzione di trasferimento è scalare, si è visto 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:

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

e calcoliamone la funzione di trasferimento:

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

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:

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

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:

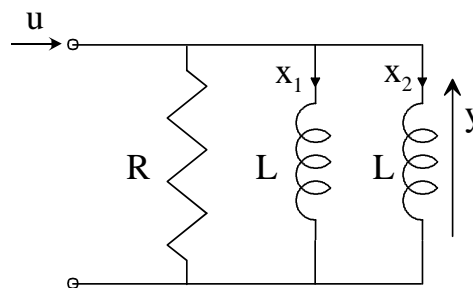


Fig. 21 : Rete elettrica

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

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

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:

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

e quindi la funzione di trasferimento:

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

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:

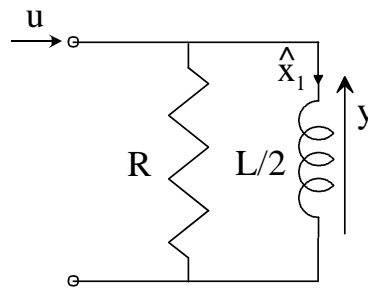


Fig. 22 : Rete elettrica equivalente agli effetti esterni

Raggiungibilità e osservabilità

La presenza di “dinamiche nascoste” in un sistema, che non si manifestano nel legame ingresso-uscita, può esser 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 \mathfrak{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 \mathfrak{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 parti⁴, con le mutue influenze e le influenze con ingresso e uscita:

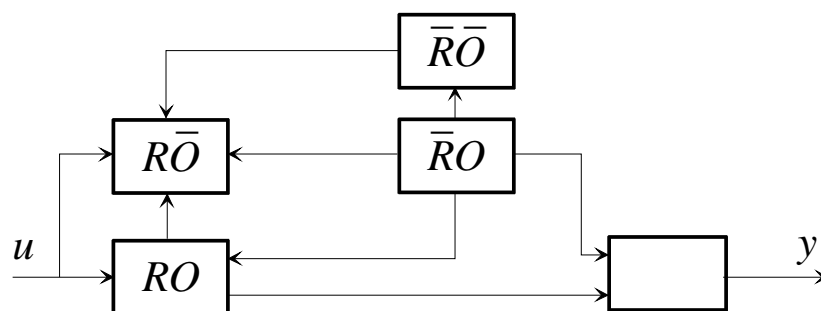


Fig. 23 : Scomposizione canonica di Kalman

⁴ Le parti non raggiungibili sono indicate con \bar{R} , quelle non osservabili con \bar{O} .

Si osservi che nel percorso tra ingresso e uscita del sistema interviene solo la parte raggiungibile ed osservabile del sistema, ovvero la **funzione di trasferimento** del sistema coinvolge solo la parte **raggiungibile ed osservabile**. Le altre parti costituiscono dinamiche nascoste, che non partecipano al legame ingresso-uscita.

Esiste un semplice **test di raggiungibilità**, che consente di stabilire se un sistema è completamente raggiungibile. Si costruisca la matrice di raggiungibilità (a n righe e $n \times m$ colonne):

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

Il sistema è completamente raggiungibile se e solo se la matrice di raggiungibilità \mathbf{K}_r ha rango n , ossia è possibile estrarre n colonne che formino una matrice quadrata di ordine n non singolare. Si osservi che se il sistema è a singolo ingresso ($m=1$), il test si riduce a verificare che la matrice \mathbf{K}_r sia **non singolare**.

Analogamente, esiste un **test di osservabilità**, che consente di stabilire se un sistema è completamente osservabile. Si costruisca la matrice di osservabilità (a n righe e $n \times p$ colonne):

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

Il sistema è completamente osservabile se e solo se la matrice di osservabilità \mathbf{K}_o ha rango n , ossia è possibile estrarre n colonne che formino una matrice quadrata di ordine n non singolare. Si osservi che se il sistema è a singola uscita ($p=1$), il test si riduce a verificare che la matrice \mathbf{K}_o sia **non singolare**.

ESEMPI

Si riprendano gli esempi del paragrafo precedente. Nel primo esempio risulta:

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

$$\mathbf{K}_o = [\mathbf{C}^T \quad \mathbf{A}^T\mathbf{C}^T] = \begin{bmatrix} 2 & -1 \\ 3 & -3 \end{bmatrix} \Rightarrow \det(\mathbf{K}_o) \neq 0$$

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

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:

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

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

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:

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

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

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

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

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

Poiché T è non singolare, il rango di \hat{K}_r coincide con il rango di 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:

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

Poiché T (e quindi T^{-T}) è non singolare, il rango di \hat{K}_o coincide con il rango di 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 l'**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:

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

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

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:

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

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:

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

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

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

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 \mathbf{T} può dunque essere il seguente:

Data (\mathbf{A}, \mathbf{B}) coppia raggiungibile:

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

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 $(\mathbf{A}, \mathbf{B}, \mathbf{C})$, la sua funzione di trasferimento, scalare, coinciderà ovviamente con la sua trasposta:

$$G(s) = \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B} = G^T(s) = [\mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B}]^T = \mathbf{B}^T (s\mathbf{I}_n - \mathbf{A})^{-T} \mathbf{C}^T = \mathbf{B}^T (s\mathbf{I}_n - \mathbf{A}^T)^{-1} \mathbf{C}^T.$$

Posto allora:

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

risulta:

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

per cui il sistema di matrici $(\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}})$, 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:

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

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

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

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

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

$$\mathbf{C} = [0 \quad 1], \quad \mathbf{D} = 1.$$

Capitolo 5: Risposte canoniche dei sistemi del primo e del secondo ordine

Per risposte canoniche si intendono le risposte dei sistemi dinamici ai segnali cosiddetti canonici (impulso, scalino, rampa), ovvero quei segnali utilizzabili come test per evidenziare le proprietà dinamiche del sistema. Ci concentreremo sui sistemi del primo e secondo ordine in quanto rappresentativi dei modelli di prima approssimazione di larga parte dei sistemi fisici.

Sul tracciato di una generica risposta allo scalino potremo definire alcuni **parametri caratteristici**:

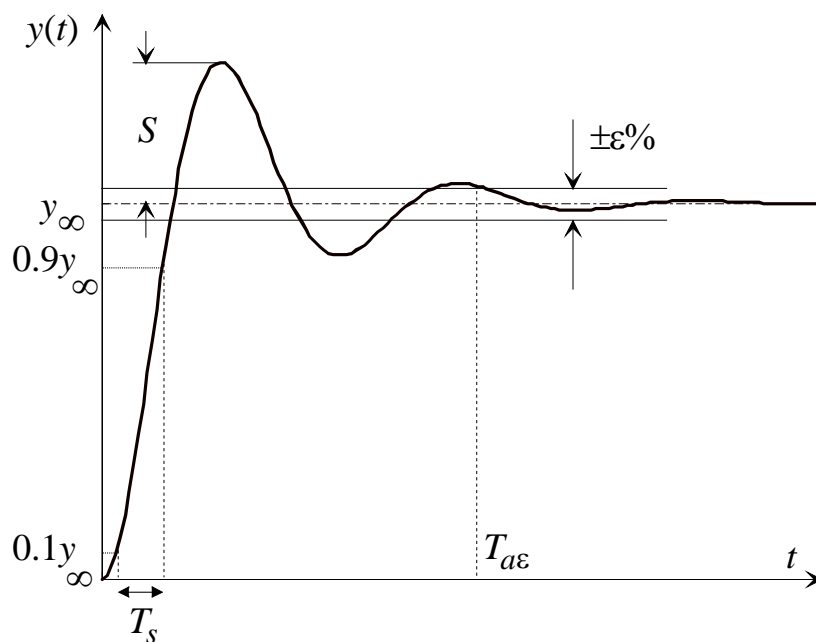


Fig. 24 : Parametri caratteristici della risposta allo scalino

- **Tempo di salita T_s** : è il tempo impiegato dalla risposta a passare dal 10% al 90% del valore di regime.
- **Tempo di assestamento al $(100-\epsilon)\%$ $T_{a\epsilon}$** : è il tempo impiegato dalla risposta ad entrare definitivamente in una fascia compresa tra $\pm\epsilon\%$ del valore di regime.
- **Sovraelongazione percentuale massima S_E** : è l'escursione massima della risposta rispetto al valore di regime, rapportata in percentuale al valore di regime stesso:

$$S_E = \frac{S}{y_\infty} 100 = \frac{\max\{y(t)\} - y_\infty}{y_\infty} 100.$$

Sistemi del primo ordine

L'espressione più generale della funzione di trasferimento per un sistema del primo ordine (ossia con un solo polo) è la seguente:

$$G(s) = \frac{\mu}{s^g} \frac{1+s\tau}{1+sT}.$$

Sistemi strettamente propri

Sono i sistemi in cui il grado del denominatore è maggiore del grado del numeratore. Sistemi del primo ordine strettamente propri non possono quindi presentare zeri:

$$G(s) = \frac{\mu}{s^g} \frac{1}{1+sT} = \begin{cases} \frac{\mu}{1+sT}, & g=0 \\ \frac{\mu}{s}, & g=1 \end{cases}$$

$$\boxed{g=0}$$

$$G(s) = \frac{\mu}{1+sT}$$

Studiamo la *risposta allo scalino* ($u(t) = \text{sca}(t) \Rightarrow U(s) = 1/s$):

$$Y(s) = G(s)U(s) = \frac{\mu}{1+sT} \frac{1}{s} = \frac{\mu}{s} - \frac{\mu T}{1+sT} = \mu \left(\frac{1}{s} - \frac{1}{s+1/T} \right).$$

Antitrasformando:

$$y(t) = \mu(1 - e^{-t/T}), \quad t \geq 0.$$

A seconda del segno di T l'andamento di y risulta molto diverso⁵:

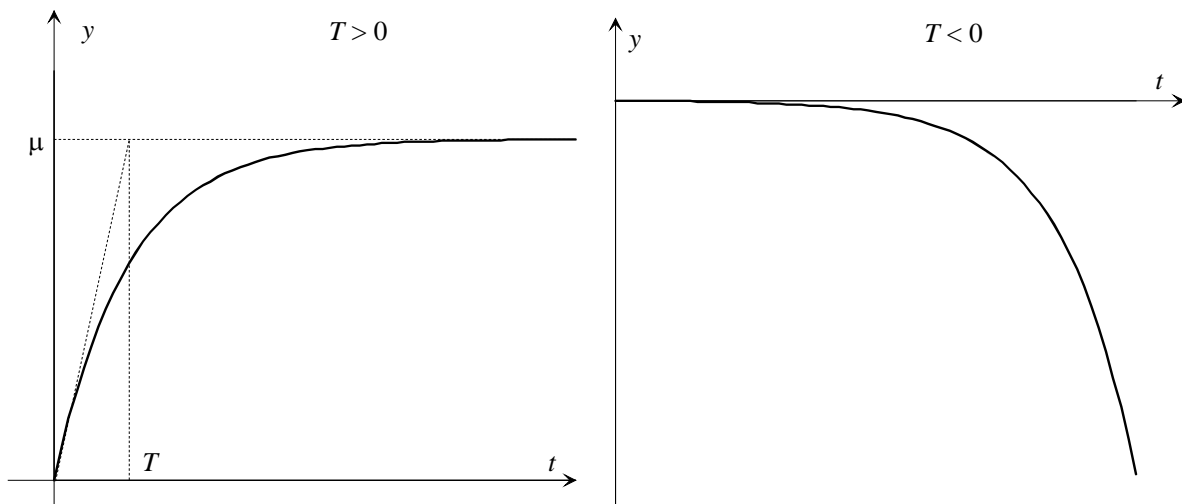


Fig. 25 : Risposta allo scalino per $T > 0$ e $T < 0$

⁵Qui e nel seguito si assumerà, senza alcuna perdita di generalità, il parametro μ positivo.

Quando $T > 0$, la risposta y si assesta su un valore finito, mentre quando $T < 0$ la risposta diverge all'infinito. Si osservi che il sistema di funzione di trasferimento $G(s)$ presenta un polo per $s = -1/T$. Pertanto il sistema risulta *asintoticamente stabile* per $T > 0$, *instabile* per $T < 0$.

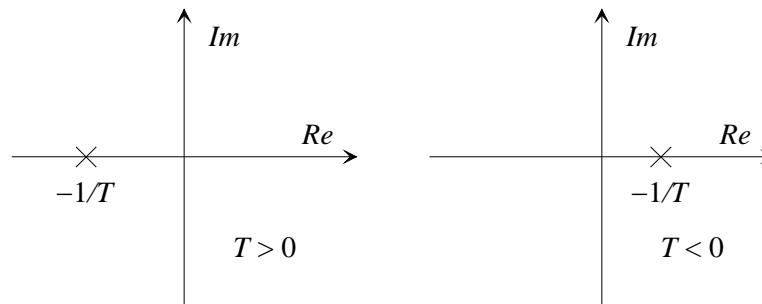


Fig. 26 : Posizione del polo per $T > 0$ e $T < 0$

Considerando solo il caso asintoticamente stabile ($T > 0$), si può calcolare il valore limite (per $t \rightarrow \infty$) della risposta allo scalino con il teorema del valore finale:

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} \left[s \frac{\mu}{1+sT} \frac{1}{s} \right] = \mu.$$

Pertanto la risposta allo scalino tende al *guadagno* μ del sistema: in altre parole il rapporto tra il valore limite dell'uscita ed il valore limite dell'ingresso (che in questo caso vale 1, essendo l'ingresso uno scalino), è pari al guadagno del sistema. Ciò costituisce una circostanza generale.

La forma del transitorio dipende invece solo dalla *costante di tempo* T . All'istante iniziale ($t=0$) la derivata di y vale μ/T : pertanto inizialmente la curva è tangente alla retta che passa per l'origine e che intercetta la retta orizzontale di ordinata μ (ossia la retta a cui tende la risposta), in corrispondenza dell'istante $t=T$ (fig. 1). Ne consegue che il transitorio è tanto più veloce quanto più piccolo è il valore della costante di tempo T . Si può verificare che la risposta y raggiunge praticamente (al 98÷99%) il valore di regime dopo un tempo pari a 4÷5 volte la costante di tempo T . Si osservi che da queste considerazioni emerge anche con molta evidenza un metodo grafico per tracciare l'andamento approssimato della risposta allo scalino.

Studiamo anche la *risposta all'impulso* ($u(t) = \text{imp}(t) \Rightarrow U(s) = 1$):

$$Y(s) = G(s)U(s) = \frac{\mu}{1+sT} = \frac{\mu}{T} \frac{1}{s+1/T}.$$

Antitrasformando:

$$y(t) = \frac{\mu}{T} e^{-t/T}, \quad t \geq 0.$$

Si noti che la risposta all'impulso risulta uguale alla derivata rispetto al tempo della risposta allo scalino (circostanza generale).

Nel caso di sistema asintoticamente stabile ($T > 0$) si ottiene l'andamento di figura:

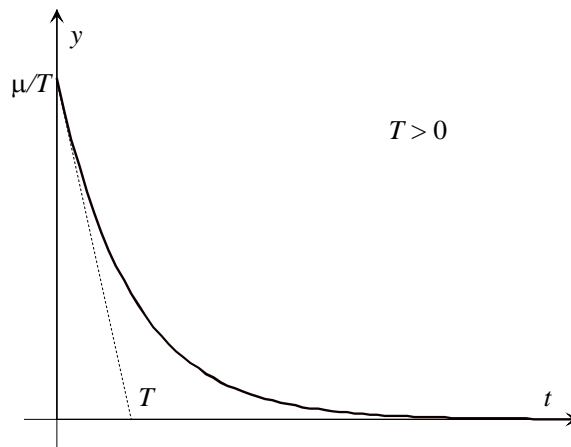


Fig. 27 : Risposta all'impulso per $T > 0$

$$\boxed{g = 1}$$

$$G(s) = \frac{\mu}{s}.$$

Il sistema ha un polo in $s=0$: è pertanto semplicemente stabile.

Studiamo la *risposta allo scalino* ($u(t) = \text{sca}(t) \Rightarrow U(s) = 1/s$):

$$Y(s) = G(s)U(s) = \frac{\mu}{s} \frac{1}{s} = \frac{\mu}{s^2}.$$

Antitrasformando:

$$y(t) = \mu \text{ ram}(t).$$

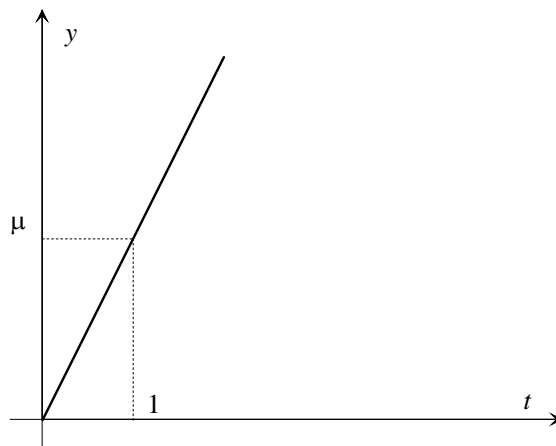


Fig. 28 : Risposta allo scalino

Per quanto riguarda la risposta all'impulso:

$$Y(s) = G(s)U(s) = \frac{\mu}{s}.$$

e quindi:

$$y(t) = \mu sca(t) .$$

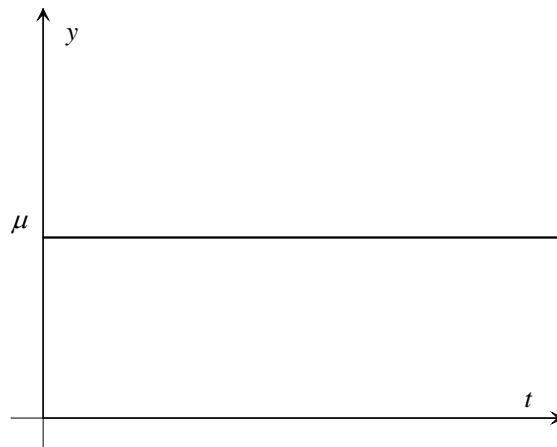


Fig. 29 : Risposta all'impulso

In entrambi i casi l'uscita equivale all'integrale dell'ingresso, moltiplicato per il fattore μ .

Sistemi propri non strettamente

Sono i sistemi in cui il grado del denominatore è uguale al grado del numeratore. Sistemi del primo ordine propri non strettamente presentano quindi uno zero:

$$G(s) = \frac{\mu}{s^g} \frac{1+s\tau}{1+sT} = \begin{cases} \mu \frac{1+s\tau}{1+sT}, & g=0 \\ \mu \frac{1+s\tau}{s}, & g=1 \\ \mu \frac{s}{1+sT}, & g=-1 \end{cases}$$

$$\boxed{g=0}$$

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

Studiando la *risposta allo scalino* si perviene alla seguente espressione:

$$y(t) = \mu \left[1 + \left(\frac{\tau}{T} - 1 \right) e^{-t/T} \right], \quad t \geq 0 .$$

Al variare del valore relativo di τ e T (e quindi della posizione relativa del polo e dello zero) la risposta allo scalino cambia sensibilmente (si considera solo il caso asintoticamente stabile, $T > 0$):

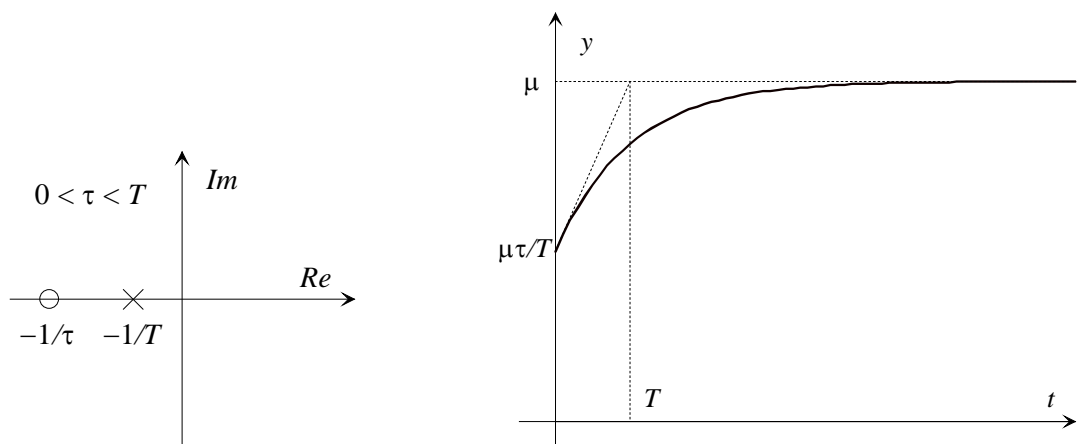


Fig. 30 : Risposta allo scalino per $0 < \tau < T$

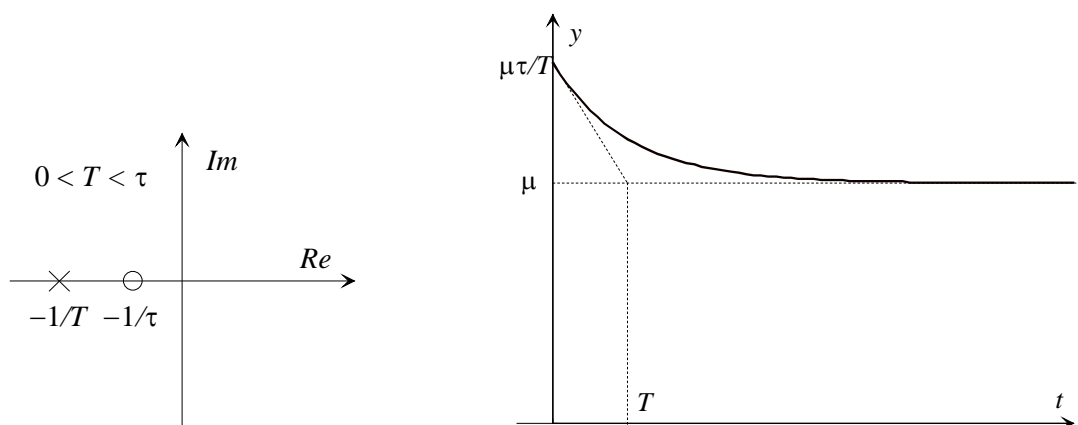


Fig. 31 : Risposta allo scalino per $0 < T < \tau$

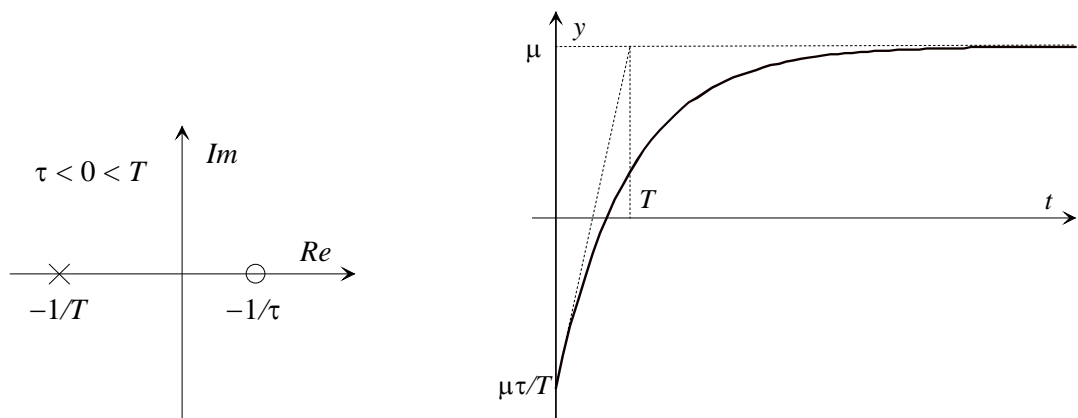


Fig. 32 : Risposta allo scalino per $\tau < 0 < T$

Uno zero nel semipiano sinistro “anticipa” la risposta rispetto al caso di sistema privo di zero, nel senso che la risposta stessa si porta inizialmente ad un valore diverso da zero, dello stesso segno del valore di regime.

Uno zero nel semipiano destro “ritarda” la risposta rispetto al caso di sistema privo di zero, nel senso che la risposta stessa si porta inizialmente ad un valore diverso da zero, di segno opposto al valore di regime (*risposta inversa*). Questo tipo di comportamento è tipico dei

sistemi con zeri nel semipiano destro, che per ragioni che saranno chiare più avanti nel corso, prendono anche il nome di sistemi *a fase non minima*.

$$g = 1$$

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

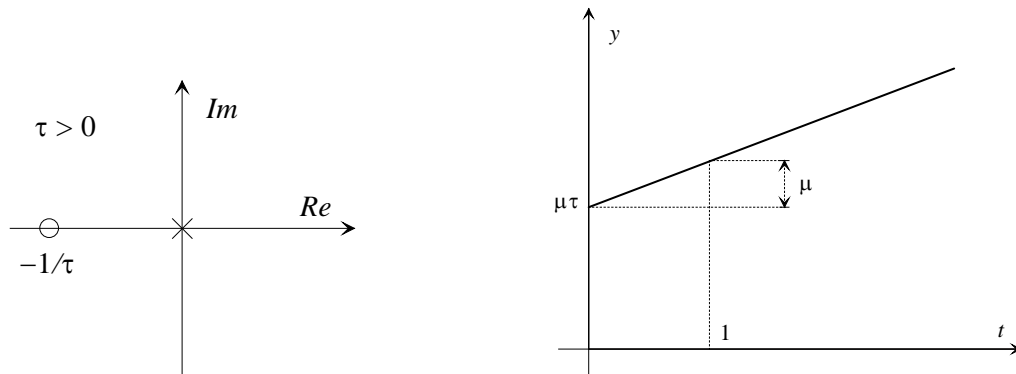


Fig. 33 : Risposta allo scalino per $\tau > 0$

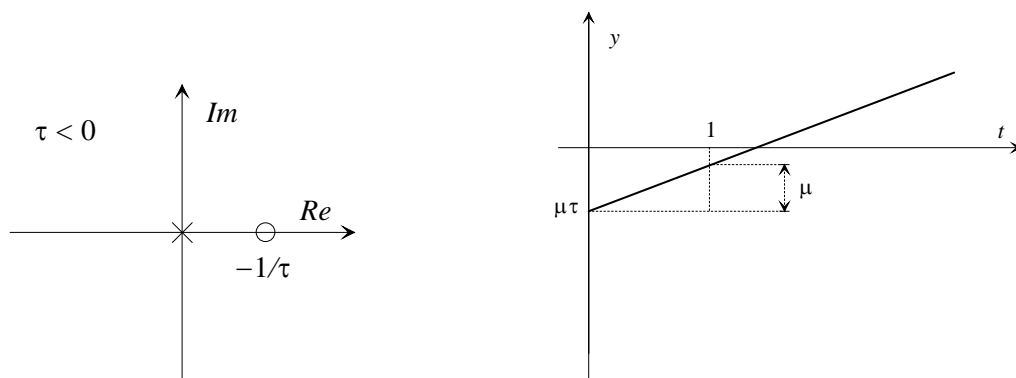


Fig. 34 : Risposta allo scalino per $\tau < 0$

$$g = -1$$

$$G(s) = \mu \frac{s}{1 + sT}$$

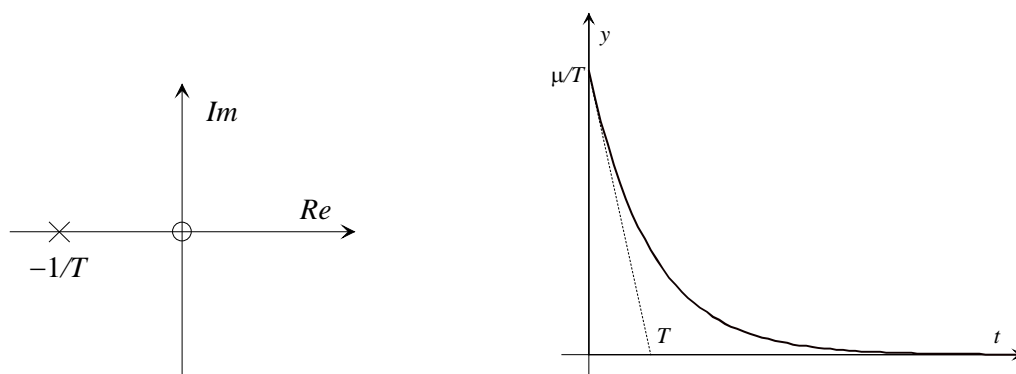


Fig. 35 : Risposta allo scalino (zero in $s=0$)

Sistemi del secondo ordine

Per i sistemi del secondo ordine (che presentano cioè due poli) ci limiteremo ad esaminare alcuni casi particolari, rinunciando alla casistica completa.

Sistemi con poli reali e nessuno zero

$$G(s) = \frac{\mu}{(1 + sT_1)(1 + sT_2)}.$$

Assumiamo T_1 e T_2 positivi, ossia il sistema asintoticamente stabile (i suoi poli, in $s = -1/T_1$, $s = -1/T_2$, sono nel semipiano sinistro).

La trasformata di Laplace della risposta allo scalino è data da:

$$Y(s) = \frac{\mu}{s(1 + sT_1)(1 + sT_2)}.$$

In base al teorema del valore iniziale, otteniamo:

$$y(0) = \lim_{s \rightarrow \infty} [sY(s)] = \lim_{s \rightarrow \infty} \frac{\mu}{(1 + sT_1)(1 + sT_2)} = 0$$

$$\dot{y}(0) = \lim_{s \rightarrow \infty} [s\mathcal{L}(\dot{y})] = \lim_{s \rightarrow \infty} [s(sY(s) - y(0))] = \lim_{s \rightarrow \infty} [s^2Y(s)] = \lim_{s \rightarrow \infty} \frac{\mu s}{(1 + sT_1)(1 + sT_2)} = 0$$

$$\ddot{y}(0) = \lim_{s \rightarrow \infty} [s\mathcal{L}(\ddot{y})] = \lim_{s \rightarrow \infty} [s(s^2Y(s) - \dot{y}(0))] = \lim_{s \rightarrow \infty} [s^3Y(s)] = \lim_{s \rightarrow \infty} \frac{\mu s^2}{(1 + sT_1)(1 + sT_2)} = \frac{\mu}{T_1 T_2} > 0$$

mentre in base al teorema del valore finale:

$$y(\infty) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} \frac{\mu}{(1 + sT_1)(1 + sT_2)} = \mu.$$

La risposta parte quindi da zero, con tangente orizzontale e concavità rivolta verso l'alto. Tende poi al valore μ .

L'antitrasformata si può ottenere con il metodo di Heaviside:

$$y(t) = \mu \left[1 - \frac{T_1}{T_1 - T_2} e^{-t/T_1} + \frac{T_2}{T_1 - T_2} e^{-t/T_2} \right], \quad t \geq 0.$$

L'andamento tipico è a forma di "S":

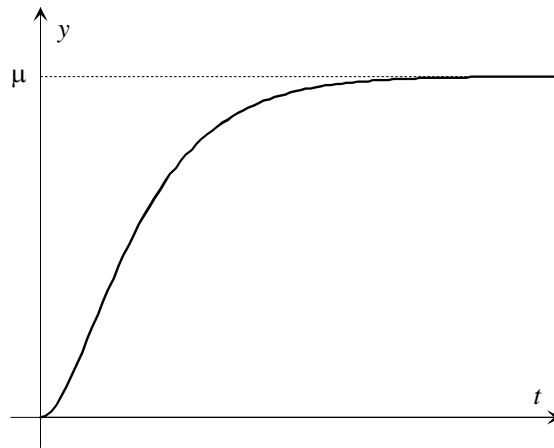


Fig. 36 : Risposta allo scalino

La durata del transitorio può essere facilmente legata alle due costanti di tempo T_1 e T_2 solo se i due valori sono molto diversi tra loro: in tal caso, infatti, conta solo il valore della costante di tempo più grande.

Se, ad esempio $T_1 \gg T_2$, allora:

$$y(t) \approx \mu \left[1 - e^{-t/T_1} \right], \quad t \geq 0.$$

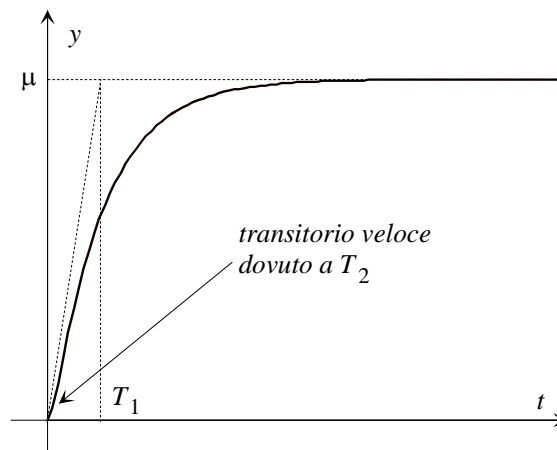


Fig. 37 : Risposta allo scalino con $T_1 \gg T_2$

Sistemi con poli reali e uno zero

$$G(s) = \mu \frac{1 + s\tau}{(1 + sT_1)(1 + sT_2)}.$$

Assumiamo T_1 e T_2 positivi, ossia il sistema asintoticamente stabile.

La trasformata di Laplace della risposta allo scalino è data da:

$$Y(s) = \mu \frac{1 + s\tau}{s(1 + sT_1)(1 + sT_2)}.$$

In base al teorema del valore iniziale, otteniamo:

$$y(0) = \lim_{s \rightarrow \infty} [sY(s)] = \lim_{s \rightarrow \infty} \left[\mu \frac{1+s\tau}{(1+sT_1)(1+sT_2)} \right] = 0$$

$$\dot{y}(0) = \lim_{s \rightarrow \infty} [s\mathcal{L}(\dot{y})] = \lim_{s \rightarrow \infty} [s(sY(s) - y(0))] = \lim_{s \rightarrow \infty} [s^2 Y(s)] = \lim_{s \rightarrow \infty} \left[\mu s \frac{1+s\tau}{(1+sT_1)(1+sT_2)} \right] = \frac{\mu\tau}{T_1 T_2}$$

mentre in base al teorema del valore finale:

$$y(\infty) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} \left[\mu \frac{1+s\tau}{(1+sT_1)(1+sT_2)} \right] = \mu.$$

La risposta parte quindi da zero, con tangente rivolta verso l'alto per $\tau > 0$, verso il basso per $\tau < 0$. La risposta tende poi al valore μ .

Qualitativamente, gli andamenti della risposta allo scalino saranno:

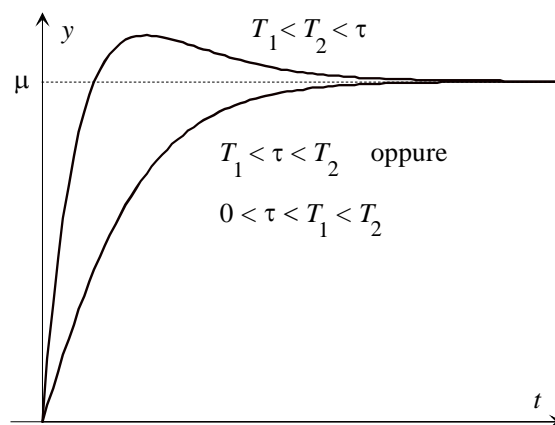


Fig. 38 : Risposta allo scalino con $\tau > 0$

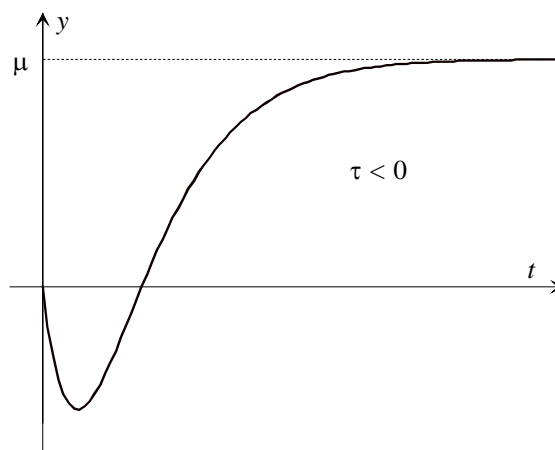


Fig. 39 : Risposta allo scalino con $\tau < 0$

Si osservi il tratto di risposta inversa nel caso di zero nel semipiano destro ($\tau < 0$).

Sistemi con poli complessi e coniugati e nessuno zero

In questo caso è comodo riscrivere l'espressione della funzione di trasferimento nella forma equivalente:

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

dove ζ , **smorzamento**, e ω_n , **pulsazione naturale**, sono stati definiti nella precedente lezione.

Si ricorda che il significato dei parametri ζ e ω_n è illustrato dalla seguente figura:

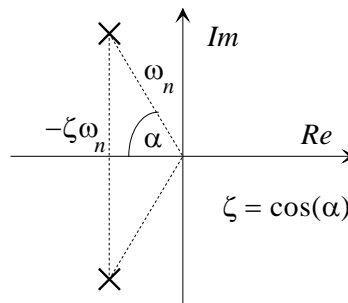


Fig. 40 : Significato dei parametri ζ e ω_n

Si osservi inoltre che:

$\zeta > 0$: due poli nel semipiano sinistro \Rightarrow sistema *asintoticamente stabile*

$\zeta = 0$: due poli sull'asse immaginario \Rightarrow sistema *semplicemente stabile*

$\zeta < 0$: due poli nel semipiano destro \Rightarrow sistema *instabile*

Studiamo la *risposta allo scalino*:

$$Y(s) = \frac{\mu}{s} \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}.$$

In base al teorema del valore iniziale, otteniamo:

$$y(0) = \lim_{s \rightarrow \infty} [sY(s)] = \lim_{s \rightarrow \infty} \frac{\mu\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = 0$$

$$\dot{y}(0) = \lim_{s \rightarrow \infty} [s\mathcal{L}(\dot{y})] = \lim_{s \rightarrow \infty} [s(sY(s) - y(0))] = \lim_{s \rightarrow \infty} [s^2Y(s)] = \lim_{s \rightarrow \infty} \frac{\mu\omega_n^2 s}{s^2 + 2\zeta\omega_n s + \omega_n^2} = 0$$

$$\ddot{y}(0) = \lim_{s \rightarrow \infty} [s\mathcal{L}(\ddot{y})] = \lim_{s \rightarrow \infty} [s(s^2Y(s) - \dot{y}(0))] = \lim_{s \rightarrow \infty} [s^3Y(s)] = \lim_{s \rightarrow \infty} \frac{\mu\omega_n^2 s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \mu\omega_n^2 > 0$$

mentre, per $\zeta > 0$, in base al teorema del valore finale:

$$y(\infty) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} \frac{\mu\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \mu.$$

L'espressione analitica della risposta allo scalino, ottenibile per antitrasformazione, è la seguente:

$$y(t) = \mu \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta \omega_n t} \sin(\omega_n \sqrt{1-\zeta^2} t + \alpha) \right],$$

dove $\zeta = \cos(\alpha)$.

Per $\zeta=0$, si ha:

$$y(t) = \mu [1 - \cos(\omega_n t)]$$

ossia una cosinusoide di pulsazione ω_n :

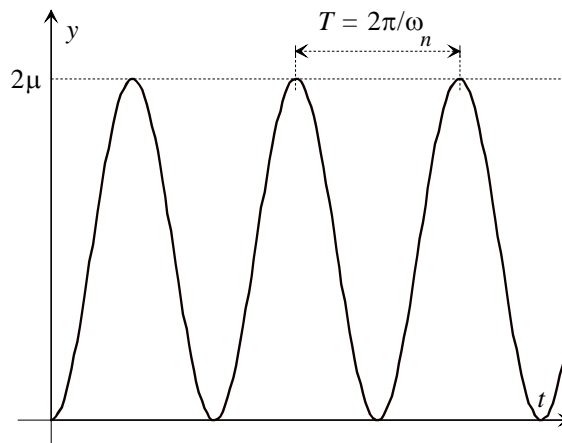


Fig. 41 : Risposta allo scalino per $\zeta=0$

Per $\zeta \neq 0$, la risposta ha l'andamento di una sinusoide involupata da due esponenziali (convergenti nel caso asintoticamente stabile, $\zeta > 0$, divergenti nel caso instabile, $\zeta < 0$).

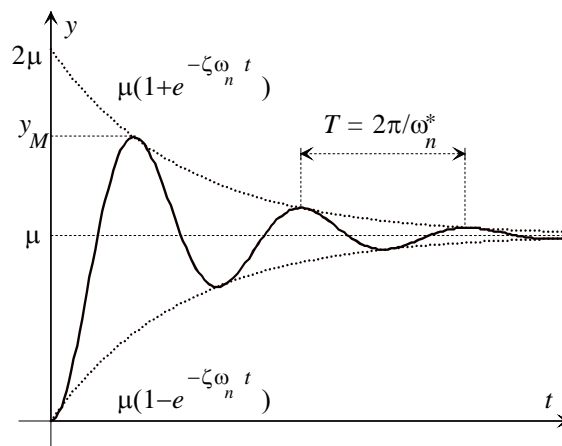


Fig. 42 : Risposta allo scalino per $\zeta > 0$ ($\omega_n^* = \omega_n \sqrt{1-\zeta^2}$)

Si può dimostrare che, nel caso asintoticamente stabile, la *sovraelongazione percentuale massima*, ossia il rapporto percentuale tra l'escursione del primo picco della risposta rispetto al valore di regime ed il valore di regime stesso, dipende esclusivamente dal fattore di smorzamento ζ :

$$S_E = 100 \frac{y_M - \mu}{\mu} = 100 e^{-\frac{\zeta \pi}{\sqrt{1-\zeta^2}}}$$

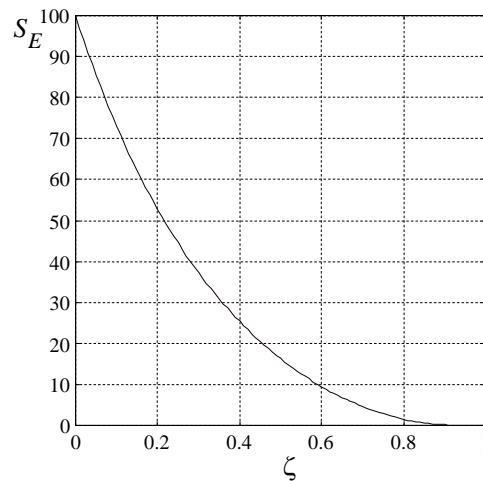


Fig. 43 : Sovraelongazione percentuale massima rispetto al fattore di smorzamento

Per fare in modo che la sovraelongazione percentuale massima sia inferiore ad un valore assegnato, occorrerà quindi che i poli del sistema appartengano ad un determinato settore del semipiano sinistro del piano complesso (come quello tratteggiato in figura):

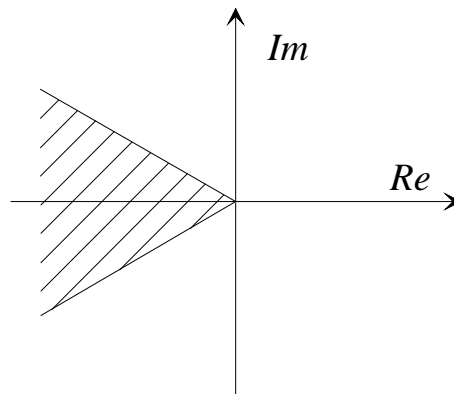


Fig. 44 : Settore del piano complesso per limitare la sovraelongazione

Il **tempo di assestamento** può invece essere determinato con buona approssimazione (per eccesso) facendo riferimento anziché alla risposta ad uno dei suoi involucri. Volendo quindi calcolare ad esempio il tempo di assestamento al 99% (T_{a1}), si imporrà:

$$\mu(1 - e^{-\zeta \omega_n T_{a1}}) = 0.99\mu \Rightarrow e^{-\zeta \omega_n T_{a1}} = 0.01 \Rightarrow \zeta \omega_n T_{a1} = \ln 100$$

e quindi:

$$T_{a1} = \frac{\ln 100}{\zeta \omega_n} \approx \frac{4.6}{\zeta \omega_n}.$$

Il tempo di assestamento risulta quindi inversamente proporzionale al modulo della **parte reale dei poli**. Per limitare il tempo di assestamento occorrerà quindi che i poli del sistema siano caratterizzati da un prodotto $\zeta \omega_n$ sufficientemente grande, ossia che appartengano ad un

semipiano incluso nel semipiano sinistro del piano complesso sufficientemente lontano dall'asse immaginario (come quello tratteggiato in figura):

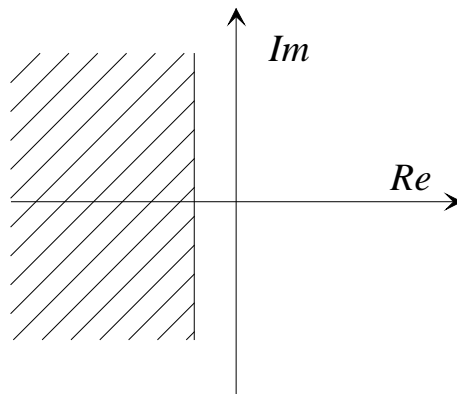


Fig. 45 : Semipiano del piano complesso per limitare il tempo di assestamento

Volendo contenere sia la sovralonga sia il tempo di assestamento, i poli della funzione di trasferimento dovranno trovarsi in una regione del piano complesso intersezione delle due regioni tratteggiate nelle precedenti figure.

Capitolo 6: Risposta in frequenza

Consideriamo un generico sistema dinamico lineare, di funzione di trasferimento $G(s)$:

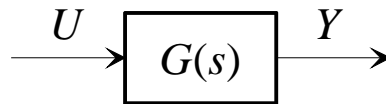


Fig. 46 : Un sistema dinamico lineare

e imponiamo il seguente andamento sinusoidale all'ingresso u

$u(t) = A \sin(\bar{\omega}t + \varphi)$

A : ampiezza
 $\bar{\omega}$: pulsazione
 φ : fase (iniziale)

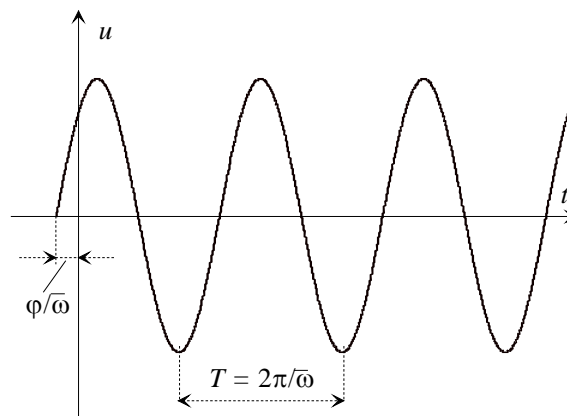


Fig. 47 : Ingresso sinusoidale

Teorema della risposta sinusoidale

Se il sistema è **asintoticamente stabile**, esaurito un transitorio iniziale, anche l'uscita è sinusoidale, con la stessa pulsazione della sinusoide in ingresso, e risulta in particolare:

$$y(t) = B \sin(\bar{\omega}t + \psi)$$

con

$$\begin{cases} B = A |G(j\bar{\omega})| \\ \psi = \varphi + \angle G(j\bar{\omega}) \end{cases}$$

dove j è l'unità immaginaria.

Risposta in frequenza

Si definisce risposta in frequenza la seguente funzione complessa della variabile reale ω :

$$G(j\omega), \quad \omega > 0.$$

A partire dall'espressione della funzione di trasferimento, l'espressione della risposta in frequenza si ottiene semplicemente sostituendo a s il prodotto $j\omega$, e limitando il campo di variabilità di ω ai valori positivi.

Per l'applicazione del teorema della risposta in frequenza, per i sistemi asintoticamente stabili, occorre poi valutare il numero complesso $G(j\omega)$ (e quindi il suo modulo e la sua fase) in corrispondenza ad un particolare valore di ω (ossia in corrispondenza alla pulsazione della sinusoide in ingresso). Coerentemente con il significato assunto nel teorema, la variabile ω prende il nome di **pulsazione**.

Si osservi che la definizione di risposta in frequenza si dà per tutti i sistemi lineari, indipendentemente dalla stabilità.

Esempio

Sia:

$$G(s) = \frac{10}{1+s}, \quad u(t) = 2 \sin(5t + 0.3)$$

Il sistema è asintoticamente stabile, per cui il teorema è applicabile. L'espressione della risposta in frequenza è la seguente:

$$G(j\omega) = \frac{10}{1+j\omega}.$$

Siamo interessati a valutare la risposta in frequenza in corrispondenza della pulsazione $\omega=5$, ed in particolare il modulo e la fase del numero complesso risultante:

$$|G(j5)| = \left| \frac{10}{1+j5} \right| = \frac{10}{|1+j5|} = \frac{10}{\sqrt{1+25}} = \frac{10}{\sqrt{26}} = 1.961$$

$$\angle G(j5) = \angle \frac{10}{1+j5} = \angle 10 - \angle(1+j5) = 0 - \arctan(5) = -1.373$$

In base al teorema della risposta in frequenza, risulterà quindi, a transitorio esaurito:

$$y(t) = 2|G(j5)| \sin(5t + 0.3 + \angle G(j5)) = 3.922 \sin(5t - 1.073).$$

Rappresentazione grafica della risposta in frequenza

Come tutte le funzioni, la risposta in frequenza è suscettibile di rappresentazione grafica. Occorre tuttavia considerare che la risposta in frequenza è una funzione complessa della variabile reale ω . Sono allora utilizzabili varie forme di rappresentazione grafica, tra le quali assumono rilevanza le seguenti due:

Diagrammi polari

Per ogni valore di ω si riporta il punto nel piano complesso $G(j\omega)$. Congiungendo i punti si ottiene una linea che prende il nome di diagramma polare.

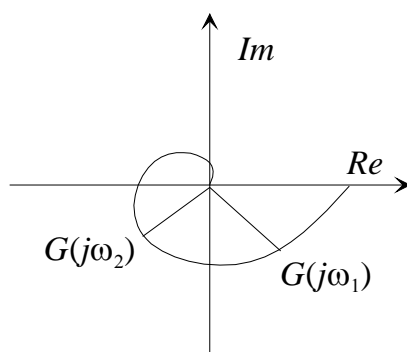


Fig. 48 : Diagramma polare

Diagrammi Cartesiani

Si tratta di una coppia di diagrammi, che rappresentano il modulo e la fase della risposta in frequenza rispetto alla pulsazione ω :

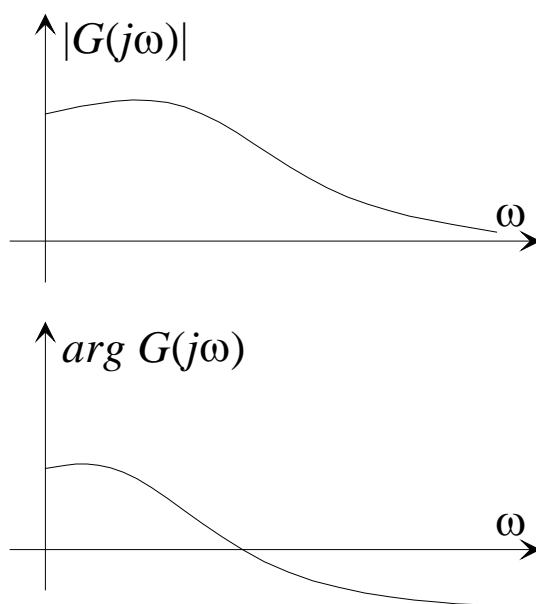


Fig. 49 : Diagrammi Cartesiani

Diagrammi di Bode

I **diagrammi di Bode** sono una coppia di diagrammi Cartesiani della risposta in frequenza, in cui le scale degli assi dell'ascissa e dell'ordinata sono scelte secondo un opportuno criterio che facilita il tracciamento dei diagrammi.

Sia nel diagramma del modulo che nel diagramma della fase l'asse delle ascisse (ossia l'asse delle pulsazioni) è in **scala logaritmica**. La distanza tra due generici punti che rappresentano le pulsazioni ω_1 e ω_2 è proporzionale alla differenza tra i logaritmi di ω_1 e ω_2 . In altre parole, date quattro pulsazioni $\omega_1, \omega_2, \omega_3$ e ω_4 tali che:

$$\frac{\omega_2}{\omega_1} = \frac{\omega_4}{\omega_3},$$

la distanza sulla scala logaritmica tra ω_1 e ω_2 è uguale alla distanza tra ω_3 e ω_4 :

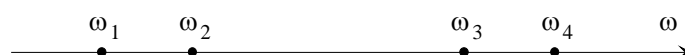


Fig. 50 : Scala logaritmica

In particolare la distanza tra due pulsazioni aventi rapporto pari a dieci prende il nome di **decade**:

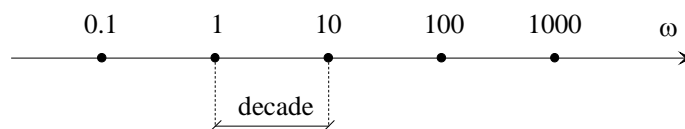


Fig. 51 : Decadi

Nel **diagramma del modulo** si rappresenta sull'asse delle ordinate il modulo in **decibel**, ossia il logaritmo in base 10 del modulo, moltiplicato per il fattore 20:

$$\left| G(j\omega) \right|_{dB} = 20 \log_{10} \left| G(j\omega) \right|.$$

I valori del modulo in decibel vengono poi rappresentati su una scala lineare. Pertanto il diagramma del modulo viene tracciato su una **carta semilogaritmica**:

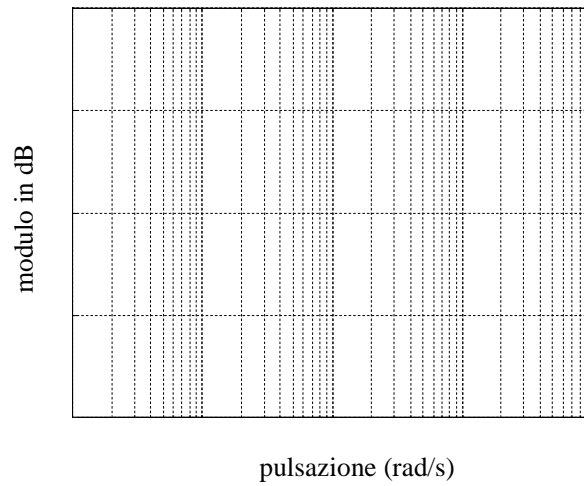


Fig. 52 : Diagramma del modulo in carta semilogaritmica

Nel **diagramma della fase** si rappresenta sull'asse delle ordinate la fase della risposta in frequenza in gradi, su scala lineare. Anche questo diagramma va quindi tracciato su carta semilogaritmica:

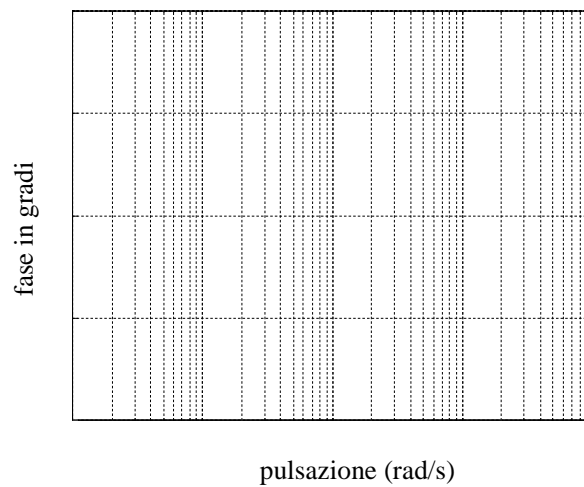


Fig. 53 : Diagramma della fase in carta semilogaritmica

Diagramma di Bode del modulo

Per una generica funzione di trasferimento:

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1 + s\tau_i)}{\prod_k (1 + s\tau_k)},$$

il modulo della risposta in frequenza assume l'espressione:

$$|G(j\omega)| = \frac{|\mu|}{|j\omega|^g} \frac{\prod_i |1 + j\omega\tau_i|}{\prod_k |1 + j\omega\tau_k|}.$$

Il modulo in decibel si scriverà quindi come:

$$\begin{aligned} |G(j\omega)|_{dB} &= 20\log_{10}|G(j\omega)| = 20\log_{10}|\mu| + 20\log_{10} \frac{1}{|j\omega|^g} \\ &\quad + \sum_i 20\log_{10}|1 + j\omega\tau_i| + \sum_k 20\log_{10} \frac{1}{|1 + j\omega\tau_k|} \end{aligned}$$

Vediamo come si tracciano i diagrammi dei singoli addendi di questa somma.

Guadagno

$$G(s) = \mu \Rightarrow |G(j\omega)|_{dB} = 20\log_{10}|\mu|$$

Si tratta di una retta orizzontale.

Esempi:

$$\begin{aligned} \mu = 100 &\Rightarrow |\mu|_{dB} = 40 \\ \mu = -10 &\Rightarrow |\mu|_{dB} = 20 \\ \mu = 0.1 &\Rightarrow |\mu|_{dB} = -20 \end{aligned}$$

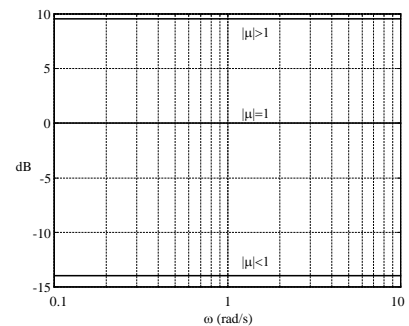


Fig 54 : Diagramma del modulo del guadagno

Zeri e poli nell'origine

$$G(s) = \frac{1}{s^g} \Rightarrow |G(j\omega)|_{dB} = 20\log_{10} \frac{1}{|j\omega|^g} = -20g \log_{10} \omega$$

Si tratta di una retta di pendenza $-20g$ dB/decade, che taglia l'asse a 0 dB per $\omega=1$.

Si dice anche che la retta ha pendenza $-g$.

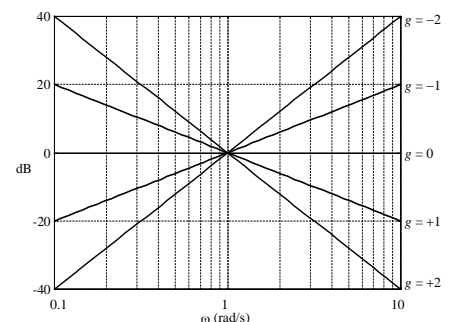


Fig 55 : Diagramma del modulo di zeri/poli in $s=0$

Zeri e poli reali

$$G(s) = (1 + s\tau)^{\pm 1} \Rightarrow$$

$$|G(j\omega)|_{dB} = \pm 20 \log_{10} |1 + j\omega\tau| = \pm 20 \log_{10} \sqrt{1 + \omega^2 \tau^2}$$

Per facilitare il tracciamento a mano si introducono i diagrammi asintotici.

$$|G(j\omega)|_{dB} \approx \begin{cases} 20 \log_{10} 1 = 0 & \omega \leq 1/|\tau| \\ \pm 20 \log_{10} |\omega\tau| & \omega > 1/|\tau| \end{cases}$$

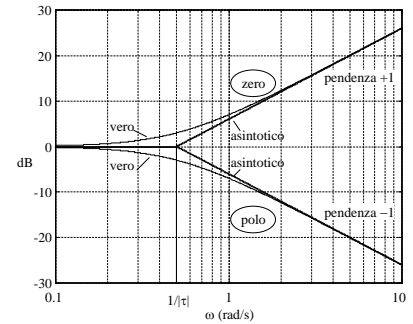


Fig 56 : Diagramma del modulo di poli/zeri reali

L'errore massimo tra diagramma vero ed asintotico si ha per $\omega = 1/|\tau|$, e vale $20 \log_{10} \sqrt{2} \approx 3dB$.

Il diagramma è del tutto indipendente dal segno di τ .

Se vi sono più zeri (poli) reali coincidenti, i diagrammi si sommano.

Zeri e poli complessi e coniugati

$$G(s) = [(1 + s\tau)(1 + s\bar{\tau})]^{\pm 1} = \left[1 + 2\frac{\zeta}{\omega_n} s + \frac{s^2}{\omega_n^2} \right]^{\pm 1} \Rightarrow$$

$$|G(j\omega)|_{dB} = \pm 20 \log_{10} \sqrt{\left(1 - \frac{\omega^2}{\omega_n^2} \right)^2 + 4\zeta^2 \frac{\omega^2}{\omega_n^2}}$$

I diagrammi asintotici si tracciano sostituendo ai due zeri (poli) due zeri (poli) reali coincidenti alla pulsazione ω_n . L'approssimazione è buona solo per valori di $|\zeta|$ elevati ($|\zeta| > 0.5$). Il diagramma non dipende dal segno di ζ .

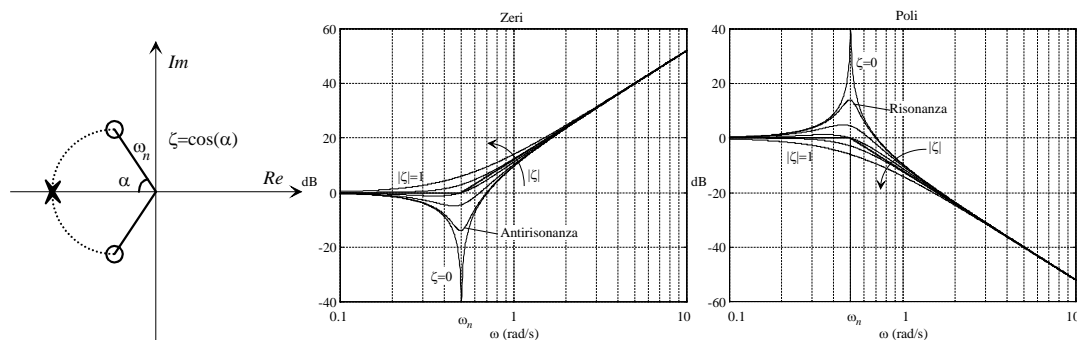


Fig.57 : Diagrammi del modulo di poli/zeri complessi e coniugati

Esempio di tracciamento

$$G(s) = \frac{10}{s} \frac{1-s}{(1+0.1s)^2}$$

$$|G(j\omega)|_{dB} = \overset{(1)}{20\log_{10}|10|} + \overset{(2)}{20\log_{10}\frac{1}{|j\omega|}} + \overset{(3)}{20\log_{10}|1-j\omega|} + \overset{(4)}{20\log_{10}\frac{1}{|1+j0.1\omega|^2}}$$

Sommando i singoli addendi si ottiene il diagramma asintotico di Bode del modulo della risposta in frequenza:

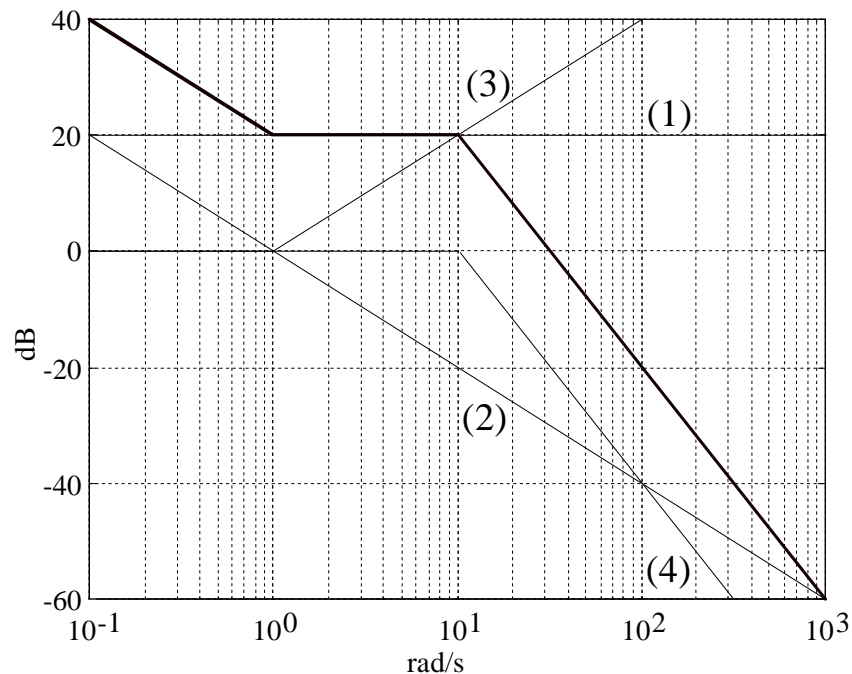


Fig. 58 : Diagramma di Bode del modulo

Per il **tracciamento veloce** del diagramma asintotico del modulo, ci si può servire delle seguenti **regole pratiche**:

1. A bassa frequenza ($\omega \rightarrow 0$) il diagramma giace sulla retta di pendenza $-g$, passante per il punto $\left[\omega = 1; |G|_{dB} = |\mu|_{dB}\right]$.
2. Ad ogni pulsazione corrispondente a p poli (zeri) reali, la pendenza diminuisce (aumenta) di p unità.
3. Ad ogni pulsazione corrispondente alla pulsazione naturale di p coppie di poli (zeri) complessi e coniugati, la pendenza diminuisce (aumenta) di $2p$ unità.
4. La pendenza finale è pari al numero degli zeri meno il numero dei poli (regola di verifica).

Diagramma di Bode della fase

Per una generica funzione di trasferimento:

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1 + s\tau_i)}{\prod_k (1 + s\tau_k)},$$

la fase della risposta in frequenza assume l'espressione:

$$\angle G(j\omega) = \angle \mu + \angle \frac{1}{(j\omega)^g} + \sum_i \angle (1 + j\omega\tau_i) + \sum_k \angle \frac{1}{1 + j\omega\tau_k}.$$

Vediamo come si tracciano i diagrammi dei singoli addendi di questa somma.

Guadagno

$$G(s) = \mu \Rightarrow \angle G(j\omega) = \angle \mu = \begin{cases} 0 & \mu \geq 0 \\ -180 & \mu < 0 \end{cases}$$

Si tratta di una retta orizzontale.

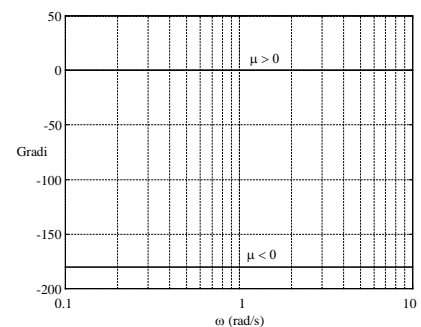


Fig. 59 : Diagramma della fase del guadagno

Zeri e poli nell'origine

$$G(s) = \frac{1}{s^g} \Rightarrow \angle G(j\omega) = \angle \frac{1}{(j\omega)^g} = -\angle (j\omega)^g = -g \angle (j\omega) = -g \cdot 90^\circ$$

Si tratta di una retta orizzontale. Ad esempio un polo nell'origine ($g=1$), ossia un integratore, dà un contributo di fase costante pari a -90° .

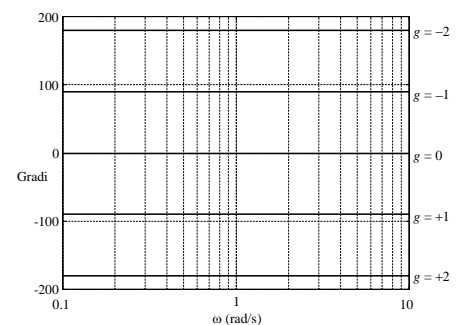


Fig. 60: Diagramm della fase di zeri/poli in $s=0$

Zeri e poli reali

$$G(s) = (1 + s\tau)^{\pm 1} \Rightarrow$$

$$\angle G(j\omega) = \angle(1 + j\omega\tau)^{\pm 1} = \pm \angle(1 + j\omega\tau) = \pm \arctan(\omega\tau)$$

Per facilitare il tracciamento a mano si introducono i diagrammi asintotici.

$$\angle G(j\omega) \approx \begin{cases} \angle 1 = 0 & \omega \leq 1/|\tau| \\ \pm \angle(j\omega\tau) & \omega > 1/|\tau| \end{cases}$$

Il diagramma dipende dal segno di τ . Infatti:

$$\angle(j\omega\tau) = \begin{cases} 90^\circ & \tau > 0 \\ -90^\circ & \tau < 0 \end{cases}$$

Se vi sono più zeri (poli) reali coincidenti, i diagrammi si sommano.

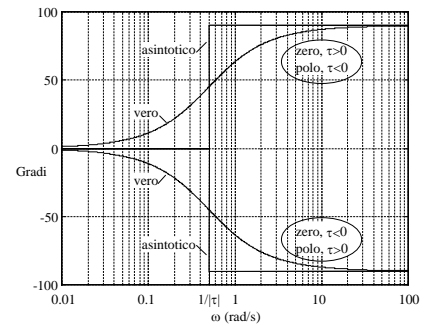


Fig. 61 : Diagramm della fase di zeri/poli reali

Zeri e poli complessi e coniugati

$$G(s) = [(1 + s\tau)(1 + s\bar{\tau})]^{\pm 1} = \left[1 + 2\frac{\zeta}{\omega_n}s + \frac{s^2}{\omega_n^2} \right]^{\pm 1} \Rightarrow \angle G(j\omega) = \pm \text{atan2}\left(1 - \frac{\omega^2}{\omega_n^2}, \frac{2\zeta}{\omega_n}\omega\right)$$

(la funzione $\text{atan2}(x, y)$ restituisce un angolo la cui tangente è y/x e il cui coseno e seno hanno segni concordi rispettivamente con x e y)

I diagrammi asintotici si tracciano sostituendo ai due zeri (poli) due zeri (poli) reali coincidenti alla pulsazione ω_n . L'approssimazione è tanto migliore quanto più piccoli sono i valori di $|\zeta|$. Il diagramma dipende dal segno di ζ .

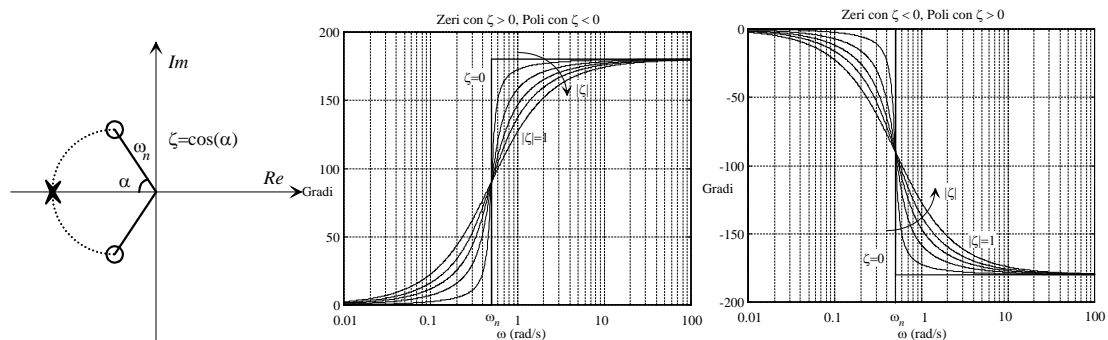


Fig. 62 : Diagrammi della fase di zeri/poli complessi e coniugati

Esempio di tracciamento

$$G(s) = \frac{10}{s} \frac{1-s}{(1+0.1s)^2}$$

$$\angle G(j\omega) = \overset{(1)}{\angle 10} + \overset{(2)}{\angle \frac{1}{j\omega}} + \overset{(3)}{\angle (1-j\omega)} + \overset{(4)}{\angle \frac{1}{(1+j0.1\omega)^2}}$$

Sommando i singoli addendi si ottiene il diagramma asintotico di Bode della fase della risposta in frequenza:

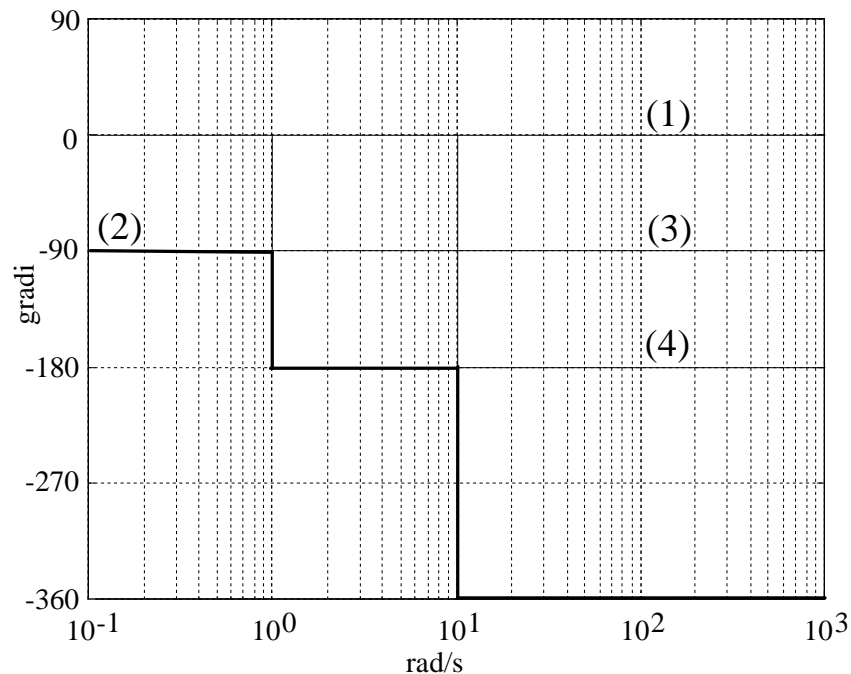


Fig. 63 : Diagramma di Bode della fase

Per il **tracciamento veloce** del diagramma asintotico della fase, ci si può servire delle seguenti **regole pratiche**:

1. A bassa frequenza ($\omega \rightarrow 0$) il diagramma giace sulla retta orizzontale di ordinata $\angle \mu - g 90^\circ$.
2. Ad ogni pulsazione corrispondente a p zeri reali nel semipiano sinistro o p poli reali nel semipiano destro, il diagramma ha un salto positivo di $p 90^\circ$.
3. Ad ogni pulsazione corrispondente a p zeri reali nel semipiano destro o p poli reali nel semipiano sinistro, il diagramma ha un salto negativo di $p 90^\circ$.
4. Ad ogni pulsazione corrispondente alla pulsazione naturale di p coppie di zeri complessi e coniugati nel semipiano sinistro o p coppie di poli complessi e coniugati nel semipiano destro, il diagramma ha un salto positivo di $p 180^\circ$.
5. Ad ogni pulsazione corrispondente alla pulsazione naturale di p coppie di zeri complessi e coniugati nel semipiano destro o p coppie di poli complessi e coniugati nel semipiano sinistro, il diagramma ha un salto negativo di $p 180^\circ$.

Azione filtrante dei sistemi dinamici

Per i sistemi dinamici lineari asintoticamente stabili si è visto che la risposta in frequenza consente di calcolare la risposta a transitorio esaurito a ingressi sinusoidali. Questo risultato si può estendere a categorie di ingressi più generali.

Ingressi periodici

Si consideri un ingresso periodico $u(t)$:

$$u(t+T) = u(t), \quad \forall t.$$

Sotto ipotesi molto generali, tra cui l'assoluta integrabilità nel periodo,

$$\int_0^T |u(t)| dt < \infty,$$

è noto che il segnale è suscettibile di sviluppo in serie di sinusoidi (**serie di Fourier**):

$$u(t) = U_0 + \sum_{n=1}^{\infty} U_n \cos(n\omega_0 t + \varphi_n), \quad \omega_0 = \frac{2\pi}{T}.$$

I coefficienti U_n e φ_n prendono il nome di coefficienti di Fourier di u ed il loro calcolo non è qui riportato.

Risulta allora che, esaurito un transitorio, anche l'uscita è periodica, con lo stesso periodo dell'ingresso, e si può scrivere:

$$y(t) = Y_0 + \sum_{n=1}^{\infty} Y_n \cos(n\omega_0 t + \psi_n),$$

con:

$$\begin{cases} Y_n = |G(jn\omega_0)| U_n \\ \psi_n = \varphi_n + \angle G(jn\omega_0) \end{cases}.$$

In altre parole il teorema della risposta in frequenza vale per tutte le sinusoidi in cui è scomponibile il segnale di ingresso.

Ingressi aperiodici

Si consideri un ingresso aperiodico $u(t)$:

$$\nexists T: u(t+T) = u(t), \quad \forall t.$$

Sotto ipotesi molto generali, tra cui l'assoluta integrabilità del segnale,

$$\int_0^{\infty} |u(t)| dt < \infty,$$

è noto che il segnale è suscettibile di sviluppo in integrale di sinusoidi (**integrale di Fourier**):

$$u(t) = \int_0^{\infty} U(\omega) \cos(\omega t + \varphi(\omega)) d\omega.$$

Le funzioni $U(\omega)$ e $\varphi(\omega)$ costituiscono la trasformata di Fourier di u ed il loro calcolo non è

qui riportato.

Risulta allora che, esaurito un transitorio, anche l'uscita è esprimibile con un integrale di Fourier:

$$y(t) = \int_0^{\infty} Y(\omega) \cos(\omega t + \psi(\omega)) d\omega ,$$

con:

$$\begin{cases} Y(\omega) = |G(j\omega)| U(\omega) \\ \psi(\omega) = \varphi(\omega) + \angle G(j\omega) \end{cases} .$$

In altre parole il teorema vale per tutte le sinusoidi in cui è scomponibile il segnale di ingresso.

La risposta in frequenza consente quindi di calcolare la risposta a qualsiasi ingresso, poiché determina come si modificano le componenti armoniche dell'ingresso. In questo senso un sistema dinamico asintoticamente stabile si può vedere sempre come un **filtro**.

Particolare rilevanza assume una tipologia di filtro che prende il nome di **filtro passabasso**. Si tratta di un sistema dinamico asintoticamente stabile, che per semplicità considereremo a guadagno unitario, caratterizzato da un modulo della risposta in frequenza del tipo di quello riportato in figura:

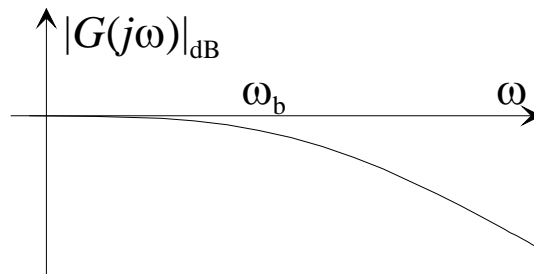


Fig. 64 : Diagramma del modulo di un filtro passabasso

Se risulta:

$$|G(j\omega)|_{dB} < 3, \forall \omega ,$$

ossia se il sistema non presenta risonanze, definiamo **banda passante** del filtro l'insieme di pulsazioni:

$$\left\{ \omega : |G(j\omega)|_{dB} > -3 \right\} = [0, \omega_b] ,$$

con ω_b estremo superiore della banda passante.

Un filtro passabasso, quindi, lascia passare le armoniche le cui pulsazioni sono interne alla sua banda passante e attenua le altre.

Capitolo 7: Schemi a blocchi

Gli schemi a blocchi costituiscono un formalismo per rappresentare graficamente le interazioni tra sistemi dinamici.

Vediamone gli elementi costitutivi:

Il blocco

Il blocco non è altro che un simbolo indicante la presenza di un sistema dinamico, avente la funzione di trasferimento riportata nel simbolo del blocco, e l'ingresso e l'uscita riportati rispettivamente sulla freccia entrante e sulla freccia uscente dal blocco:

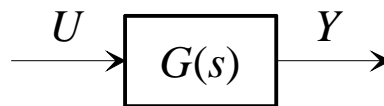


Fig. 65 : Un blocco

Il nodo sommatore

L'uscita del nodo è data dalla somma algebrica dei segnali che entrano nel nodo, ciascuno preso con il proprio segno (se non è indicato il segno, si assume per convenzione il segno positivo).

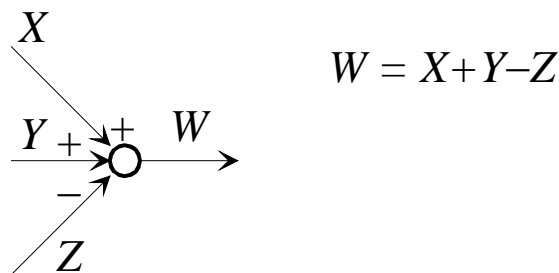


Fig. 66 : Un nodo sommatore

Il punto di diramazione

Tutti i segnali uscenti da un punto di diramazione sono uguali al segnale entrante nel punto.

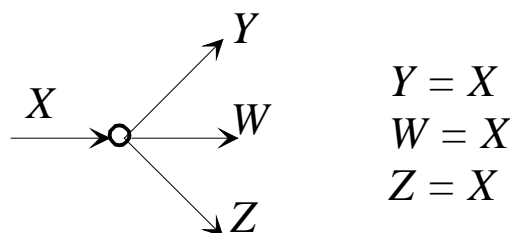


Fig. 67 : Un punto di diramazione

Schemi di interconnessione

Sistemi in cascata (o serie)

Due sistemi si dicono in cascata (o in serie) se l'uscita di uno è l'ingresso dell'altro.

Graficamente si ha la seguente situazione:

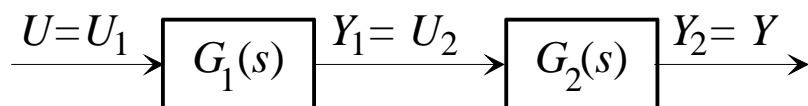


Fig. 68 : Blocchi in cascata

La funzione di trasferimento dall'ingresso del primo sistema all'uscita del secondo si ottiene come segue:

$$Y(s) = Y_2(s) = G_2(s)U_2(s) = G_2(s)Y_1(s) = G_2(s)G_1(s)U_1(s) = G_2(s)G_1(s)U(s)$$

Pertanto:

$$\frac{Y(s)}{U(s)} = G_1(s)G_2(s)$$

La funzione di trasferimento del sistema costituito dalla cascata di due sottosistemi è quindi data dal **prodotto** delle due funzioni di trasferimento parziali.

Sistemi in parallelo

Due sistemi si dicono in parallelo se hanno lo stesso ingresso, mentre le loro uscite si sommano (algebricamente) per determinare l'uscita del sistema risultante.

Graficamente si ha la seguente situazione:

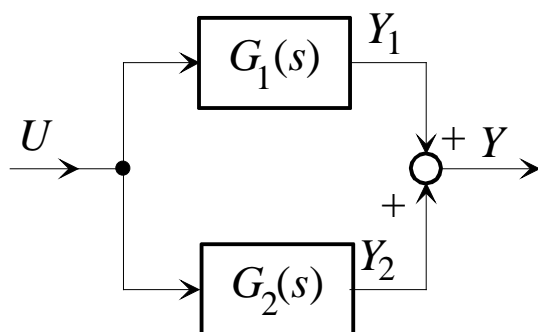


Fig. 69 : Blocchi in parallelo

La funzione di trasferimento dall'ingresso comune ai due sistemi all'uscita si ottiene come segue:

$$Y(s) = Y_1(s) + Y_2(s) = G_1(s)U(s) + G_2(s)U(s) = [G_1(s) + G_2(s)]U(s)$$

Pertanto:

$$\frac{Y(s)}{U(s)} = G_1(s) + G_2(s)$$

La funzione di trasferimento del sistema costituito dal parallelo di due sottosistemi è quindi data dalla **somma algebrica** delle due funzioni di trasferimento parziali, ciascuna presa con il segno con cui la sua uscita entra nel nodo sommatore.

Sistemi in retroazione

Due sistemi si dicono connessi in retroazione quando l'uscita del primo è l'ingresso del secondo, mentre l'uscita del secondo si somma o si sottrae ad un ingresso esterno per determinare l'ingresso del primo sistema.

Si hanno quindi due possibili schemi di connessione:

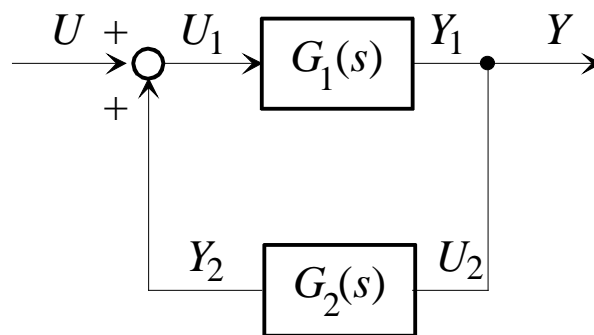


Fig. 70 : Blocchi in retroazione positiva

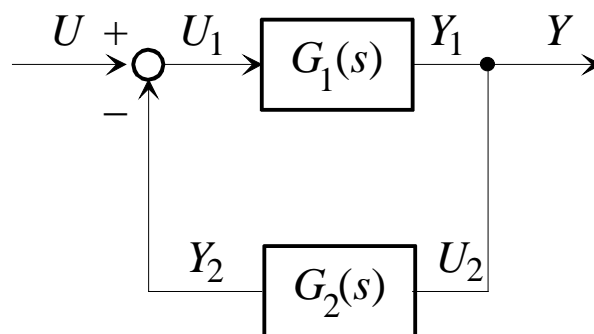


Fig. 71 : Blocchi in retroazione negativa

In entrambi i casi:

G_1 : funzione di trasferimento della **linea di andata**

G_2 : funzione di trasferimento della **linea di retroazione**

Consideriamo il caso di retroazione **positiva** e calcoliamo la funzione di trasferimento dall'ingresso U all'uscita Y :

$$\begin{aligned} Y(s) &= Y_1(s) = G_1(s)U_1(s) = G_1(s)[U(s) + Y_2(s)] = G_1(s)[U(s) + G_2(s)U_2(s)] = \\ &= G_1(s)[U(s) + G_2(s)Y(s)] = G_1(s)U(s) + G_1(s)G_2(s)Y(s) \end{aligned}$$

Pertanto:

$$\frac{Y(s)}{U(s)} = \frac{G_1(s)}{1 - G_1(s)G_2(s)}$$

Analogamente, nel caso di retroazione **negativa**:

$$\begin{aligned} Y(s) = Y_1(s) &= G_1(s)U_1(s) = G_1(s)[U(s) - Y_2(s)] = G_1(s)[U(s) - G_2(s)U_2(s)] = \\ &= G_1(s)[U(s) - G_2(s)Y(s)] = G_1(s)U(s) - G_1(s)G_2(s)Y(s) \end{aligned}$$

e quindi:

$$\frac{Y(s)}{U(s)} = \frac{G_1(s)}{1 + G_1(s)G_2(s)}$$

La funzione di trasferimento $G_1(s)G_2(s)$ prende il nome di **funzione di trasferimento d'anello**.

La regola per trovare la funzione di trasferimento del sistema complessivo (**sistema in anello chiuso**) è quindi la seguente :

$$\frac{Y(s)}{U(s)} = \frac{\text{f.d.t. linea di andata}}{1 \mp \text{f.d.t. d'anello}} \quad \begin{array}{l} - : \text{retroazione positiva} \\ + : \text{retroazione negativa} \end{array}$$

Stabilità degli schemi di interconnessione

Sistemi in cascata

Siano:

$$G_1(s) = \frac{N_1(s)}{D_1(s)}, \quad G_2(s) = \frac{N_2(s)}{D_2(s)},$$

le funzioni di trasferimento dei due sistemi in cascata, espresse come rapporti di polinomi.

La funzione di trasferimento del sistema complessivo sarà quindi:

$$G(s) = G_1(s)G_2(s) = \frac{N_1(s)}{D_1(s)} \frac{N_2(s)}{D_2(s)}.$$

Il denominatore di $G(s)$ è dato dal prodotto dei denominatori delle funzioni di trasferimento parziali: ne consegue che i poli del sistema complessivo sono la riunione dei poli dei due sottosistemi in cascata. Pertanto:

Un sistema costituito dalla cascata di due o più sottosistemi è asintoticamente stabile se e solo se lo sono tutti i sottosistemi che compongono la cascata.

Il precedente ragionamento non prevede la possibilità che vi siano radici di N_1 uguali a radici di D_2 , o radici di N_2 uguali a radici di D_1 , ossia che intervengano *cancellazioni* tra poli di una funzione di trasferimento e zeri dell'altra. Se viceversa tali cancellazioni avvengono, occorre porre attenzione al fatto che i poli cancellati siano o meno a parte reale negativa (ossia nel semipiano sinistro).

Se infatti tutti i poli cancellati sono nel semipiano sinistro, essi non hanno alcun ruolo nel determinare l'asintotica stabilità del sistema complessivo, che viene ovviamente a dipendere dai poli non cancellati. Se invece almeno uno dei poli cancellati non è nel semipiano sinistro, mentre tutti i poli non cancellati lo sono, si sarebbe indotti a ritenere che il sistema risultante sia asintoticamente stabile (il denominatore della funzione di trasferimento ottenuto a seguito delle cancellazioni presenterebbe tutte radici nel semipiano sinistro). In realtà una situazione di questo tipo corrisponderebbe alla presenza di una instabilità (o, comunque, non asintotica stabilità) *interna*: a seguito di una sollecitazione impulsiva all'ingresso, seppure la variabile di uscita del sistema si riporta, esaurito il transitorio, al valore di riposo, altre variabili interne possono crescere indefinitamente, o comunque non ritornare al valore di riposo.

Concludiamo quindi che la precedente affermazione sulla stabilità dei sistemi connessi in cascata è in realtà valida, facendo riferimento al concetto di stabilità interna, anche in presenza di cancellazioni tra poli e zeri.

Sistemi in parallelo

Siano:

$$G_1(s) = \frac{N_1(s)}{D_1(s)}, \quad G_2(s) = \frac{N_2(s)}{D_2(s)},$$

le funzioni di trasferimento dei due sistemi in parallelo. La funzione di trasferimento del sistema complessivo sarà quindi (i segni con cui avviene la somma sono irrilevanti ai fini della stabilità):

$$G(s) = G_1(s) + G_2(s) = \frac{N_1(s)}{D_1(s)} + \frac{N_2(s)}{D_2(s)} = \frac{N_1(s)D_2(s) + N_2(s)D_1(s)}{D_1(s)D_2(s)}.$$

Anche in questo caso, il denominatore di $G(s)$ è dato dal prodotto dei denominatori delle funzioni di trasferimento parziali: ne consegue che i poli del sistema complessivo sono la riunione dei poli dei due sottosistemi in cascata. Pertanto:

Un sistema costituito dal parallelo di due o più sottosistemi è asintoticamente stabile se e solo se lo sono tutti i sottosistemi che compongono il parallelo.

Un ragionamento analogo a quello sviluppato per i sistemi in cascata consente di concludere che l'affermazione è valida, con riferimento al concetto più generale di stabilità interna, anche in presenza di poli comuni tra le due funzioni di trasferimento (ovvero radici comuni di D_1 e D_2 , che comportano cancellazioni).

Sistemi in retroazione

Siano:

$$G_1(s) = \frac{N_1(s)}{D_1(s)}, \quad G_2(s) = \frac{N_2(s)}{D_2(s)},$$

le funzioni di trasferimento dei due sistemi in retroazione. La funzione di trasferimento del sistema complessivo sarà quindi:

$$G(s) = \frac{G_1(s)}{1 \mp G_1(s)G_2(s)} = \frac{\frac{N_1(s)}{D_1(s)}}{1 \mp \frac{N_1(s)}{D_1(s)} \frac{N_2(s)}{D_2(s)}} = \frac{N_1(s)D_2(s)}{D_1(s)D_2(s) \mp N_1(s)N_2(s)},$$

con l'opportuno segno a seconda che si tratti di retroazione positiva o negativa.

Pertanto i poli del sistema in anello chiuso sono le radici del denominatore:

$$D_1(s)D_2(s) \mp N_1(s)N_2(s)$$

e non hanno nessuna relazione precisa con le radici dei polinomi D_1 e D_2 , ossia con i poli dei due sottosistemi interconnessi. Pertanto:

Per un sistema costituito dalla retroazione di due sottosistemi non si può affermare nulla sulla asintotica stabilità del sistema in anello chiuso a partire dalla asintotica stabilità o meno dei due sistemi interconnessi.

Capitolo 8: Requisiti di un sistema di controllo

Essaurita la trattazione dei sistemi dinamici, si torna ora al problema di controllo, che aveva dato origine a tale studio. In figura è riportata la struttura tipica di un sistema di controllo in retroazione:

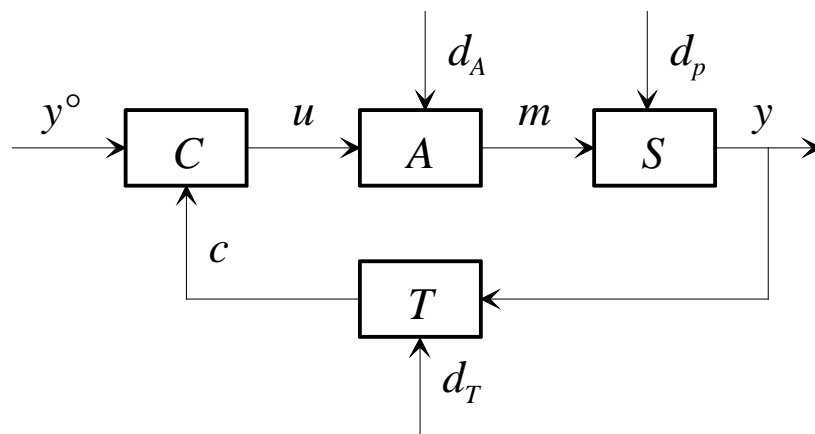


Fig. 72 : Sistema di controllo completo di strumentazione

dove:

S: sistema sotto controllo (o processo)

T: trasduttore

A: attuatore

C: controllore (o regolatore)

Sistema sotto controllo

È l'oggetto dello studio delle precedenti lezioni. Su di esso si interviene attraverso la variabile manipolabile m ed il suo comportamento è influenzato dal disturbo d_p . La sua uscita è la variabile controllata y . Se il sistema è lineare, le sue proprietà dinamiche possono essere espresse per mezzo di due funzioni di trasferimento⁶:

$$y(s) = P(s) m(s) + H(s) d_p(s).$$

⁶Indicheremo con lo stesso simbolo una variabile funzione del tempo e la sua trasformata di Laplace.

Trasduttore

E' lo strumento che misura una grandezza fisica del sistema sotto controllo (la variabile controllata y) e ne invia la misura c al controllore, in una forma compatibile con la sua tecnologia. E' generalmente caratterizzabile con due proprietà, *precisione* e *ripetibilità*.

Si può distinguere tra *precisione statica* (a transitorio esaurito il segnale che esprime la misura della grandezza è proporzionale al valore assunto dalla grandezza stessa) e *precisione dinamica* (velocità del transitorio con il quale lo strumento reagisce a variazioni nella grandezza misurata). La *ripetibilità* invece è la proprietà per cui il comportamento del trasduttore, sia statico che dinamico, non varia nel tempo.

Se il comportamento dinamico del trasduttore è approssimabile a quello di un sistema dinamico lineare, la relazione che intercorre tra le trasformate della grandezza controllata y e della misura c è esprimibile per mezzo di una funzione di trasferimento e può essere affetta da un disturbo:

$$c(s) = T(s) y(s) + d_T(s)$$

Se il trasduttore è ripetibile, $T(s)$ non varia nel tempo. In tal caso è anche possibile individuare un andamento desiderato della misura c° , elaborando con un sistema di funzione di trasferimento $T(s)$ l'andamento desiderato y° :

$$c^\circ(s) = T(s) y^\circ(s).$$

Attuatore

L'attuatore traduce l'azione di controllo elaborata dal controllore, ed espressa dalla variabile di controllo u , in un'azione efficace sulla variabile manipolabile m . Ad esso è quindi di norma associato uno stadio di *amplificazione di potenza* ed eventualmente di *conversione di potenza* (si pensi ad un motore elettrico che converte potenza elettrica in potenza meccanica).

Anche per gli attuatori ipotizzeremo un comportamento dinamico lineare affetto da disturbo, per cui:

$$m(s) = A(s) u(s) + d_A(s).$$

Controllore

Il controllore riceve in ingresso la misura c della variabile controllata ed il relativo segnale di riferimento c° . Dovendo rendere questi due segnali quanto più possibile simili, è naturale che il controllore agisca sulla loro differenza, ossia sull'errore $e_c = c^\circ - c$. Ipotizzeremo che anche il controllore abbia un comportamento dinamico lineare, per cui si avrà:

$$u(s) = R(s) e_c(s).$$

Alla luce delle precedenti considerazioni, siamo in grado di riformulare un problema di controllo sulla base di uno schema a blocchi di sistemi dinamici lineari:

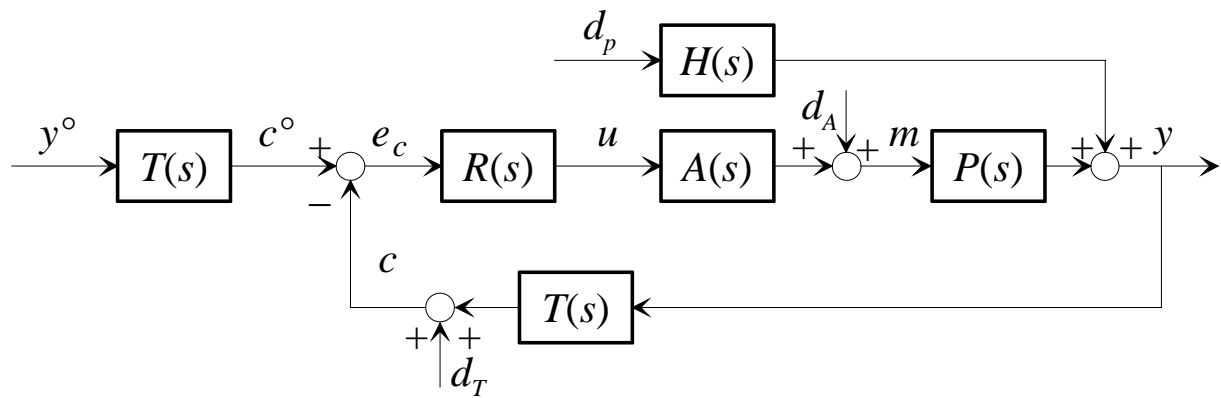


Fig. 73 : Sistema di controllo lineare

Ipotizzeremo le funzioni di trasferimento $P(s)$, $H(s)$, $T(s)$ e $A(s)$ date, insieme con l'andamento del segnale di riferimento y^o . L'incognita del problema sarà la funzione di trasferimento $R(s)$.

Formalizzazione del problema di controllo

Lo schema a blocchi di Fig. 73 può essere semplificato osservando che l'effetto dei due disturbi d_p e d_A in linea di andata equivale all'effetto di un unico disturbo d riportato direttamente sull'uscita del processo:

$$d(s) = P(s) d_A(s) + H(s) d_p(s).$$

Inoltre osserviamo che:

$$e_c(s) = c^o(s) - c(s) = T(s)y^o(s) - T(s)y(s) - d_T(s) = T(s)[y^o(s) - y(s) - n(s)],$$

dove:

$$n(s) = T(s)^{-1} d_T(s).$$

Si ottiene quindi lo schema a blocchi riportato di seguito:

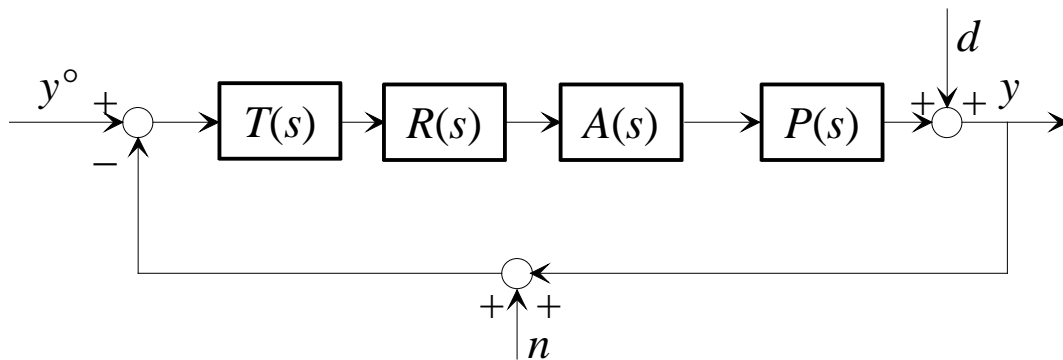


Fig. 74 : Prima elaborazione del sistema di controllo

A questo punto, sfruttando la commutatività del prodotto tra funzioni di trasferimento, si può ulteriormente semplificare lo schema a blocchi:

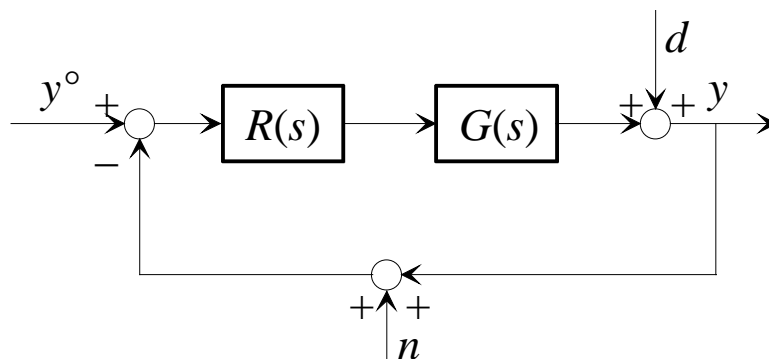


Fig. 75 : Seconda elaborazione del sistema di controllo

con:

$$G(s) = T(s) P(s) A(s).$$

Infine, una terza elaborazione porta al seguente schema:

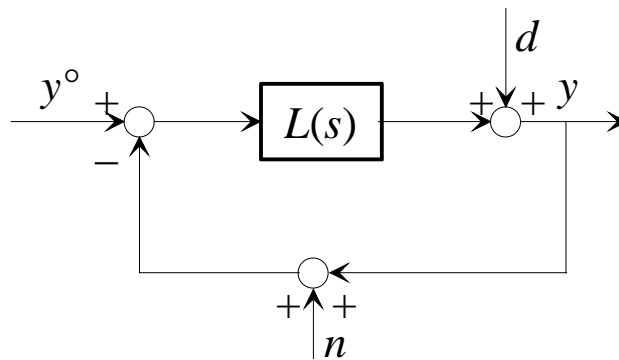


Fig. 76 : Terza elaborazione del sistema di controllo

dove:

$$L(s) = R(s) G(s).$$

Si osservi che $L(s)$ è la **funzione di trasferimento d'anello** del sistema.

L'obiettivo ideale $y \equiv y^o$ non è realizzabile, a causa dei limiti connessi alla dinamica del sistema sotto controllo, dell'attuatore e del trasduttore. Si definiscono allora una serie di **requisiti** che il progetto del controllore dovrà soddisfare:

- | | |
|--|--|
| Stabilità: | Il sistema in anello chiuso deve essere asintoticamente stabile, altrimenti qualsiasi perturbazione agente in qualsiasi punto dell'anello si amplificherebbe indefinitamente. |
| Precisione statica: | A regime, a seguito di assegnate perturbazioni (a scalino, a rampa ecc.) degli ingressi, l'errore tra riferimento e variabile controllata deve essere nullo, oppure inferiore ad una soglia prefissata. |
| Precisione dinamica: | La variabile controllata deve inseguire le variazioni del riferimento, e reagire a perturbazioni sui disturbi, con sufficiente rapidità, e senza manifestare comportamenti oscillatori. |
| Intensità dell'azione di controllo: | A causa dei limiti di funzionamento lineare degli attuatori, oltre che per non danneggiare gli attuatori stessi, occorre evitare che la variabile di controllo subisca brusche variazioni o assuma valori eccessivi. |

Capitolo 9: Stabilità dei sistema di controllo

Riprendiamo lo schema a blocchi di un sistema di controllo in retroazione:

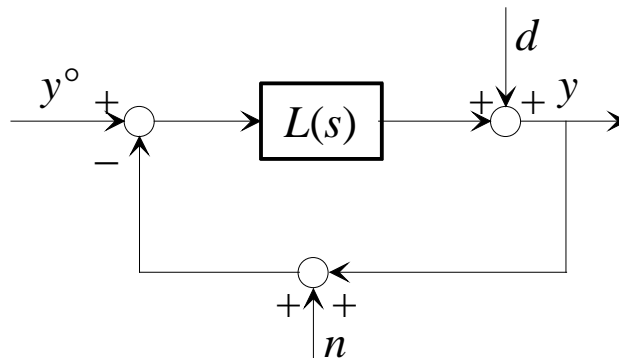


Fig. 77 : Sistema di controllo

Essendo la stabilità una proprietà del sistema, indipendente dagli ingressi, possiamo considerare il sistema di controllo privo dei disturbi:

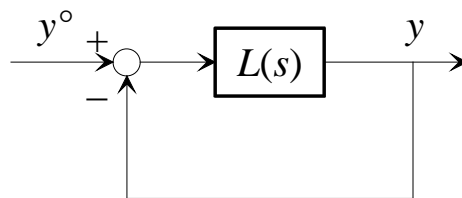


Fig. 78 : Sistema di controllo privo di disturbi

Espressa la funzione di trasferimento d'anello come rapporto di polinomi:

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

la funzione di trasferimento da y° a y assume l'espressione:

$$\frac{c(s)}{c^\circ(s)} = \frac{L(s)}{1 + L(s)} = \frac{\frac{N(s)}{D(s)}}{1 + \frac{N(s)}{D(s)}} = \frac{N(s)}{N(s) + D(s)}.$$

Definiamo il denominatore di questa funzione di trasferimento **polinomio caratteristico in anello chiuso**:

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

Le radici di tale polinomio sono quindi i **poli** del sistema in anello chiuso.

Pertanto il sistema in anello chiuso è asintoticamente stabile se e solo se tutte le radici del polinomio caratteristico hanno parte reale negativa.

Esempio

Sia

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

Poiché il polinomio caratteristico in anello chiuso:

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

non soddisfa la condizione necessaria perché le sue radici abbiano tutte parte reale negativa, il sistema in anello chiuso non è asintoticamente stabile.

Se nel formare il prodotto $L(s)$ tra $R(s)$ e $G(s)$ intervengono cancellazioni tra poli (o zeri) di $R(s)$ e zeri (o poli) di $G(s)$, i poli cancellati non compaiono più come radici del polinomio caratteristico. Ricordando tuttavia la discussione condotta precedentemente riguardo gli effetti delle cancellazioni tra poli e zeri sulla stabilità dei sistemi interconnessi, osserviamo che, se i poli cancellati non hanno parte reale negativa, il sistema nel suo complesso non può dirsi asintoticamente stabile, dal momento che nasce una instabilità (o, comunque, una non asintotica stabilità) interna.

Poiché zeri e poli della funzione di trasferimento $G(s)$ sono da ritenersi assegnati, le considerazioni precedenti conducono alle seguenti **conclusioni**:

- a) Se il sistema sotto controllo $G(s)$ ha un polo a parte reale positiva o nulla, tale polo non può essere cancellato da un corrispondente zero di $R(s)$.
- b) Se il sistema sotto controllo $G(s)$ ha uno zero a parte reale positiva o nulla, tale zero non può essere cancellato da un corrispondente polo di $R(s)$.

Il criterio di Nyquist

Il criterio di Nyquist è un criterio grafico di stabilità molto generale e di più immediata utilità del criterio del polinomio caratteristico ai fini della sintesi del controllore.

In questo corso ci si limiterà a dare l'enunciato del criterio, senza entrare in ulteriori approfondimenti.

Il criterio di Nyquist si basa sul tracciamento del cosiddetto **diagramma di Nyquist** associato alla funzione di trasferimento d'anello $L(s)$: si tratta del diagramma polare della risposta in frequenza di L , orientato nel senso delle ω crescenti, cui si aggiunge il simmetrico rispetto all'asse reale del piano complesso. Occorre poi introdurre due quantità:

P_d : numero di poli a parte reale strettamente positiva di $L(s)$

N : numero di giri compiuti dal diagramma di Nyquist intorno al punto -1 dell'asse reale, contati positivamente in senso antiorario. Se il diagramma passa per il punto -1 , N si dice non definito.

Il criterio afferma che il sistema in anello chiuso è asintoticamente stabile se e solo se N è ben definito e risulta:

$$N = P_d$$

Esempio

Sia:

$$L(s) = \frac{10}{(1+s)^2}.$$

Il diagramma polare si traccia sulla base dei diagrammi di Bode asintotici (il modulo parte da 10 e decresce monotonicamente, la fase parte da 0 e decresce monotonicamente fino a -180°). Dal diagramma polare è immediato tracciare il diagramma di Nyquist:

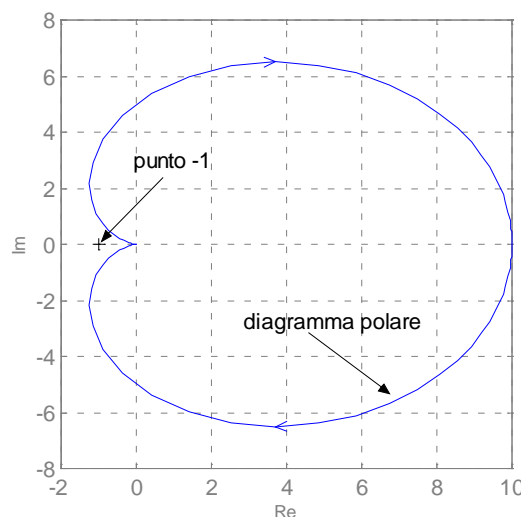


Fig. 79 : Diagramma di Nyquist di L

Si osservi che $P_d=0$, e che il diagramma di Nyquist non compie giri intorno al punto -1 , per cui $N=0$. Poiché $N=P_d$ il sistema in anello chiuso è asintoticamente stabile.

Per verifica, osserviamo che il polinomio caratteristico in anello chiuso è il seguente:

$$\chi(s) = 10 + (1+s)^2 = s^2 + 2s + 11,$$

ed ha entrambe le radici:

$$s_{1,2} = -1 \pm j\sqrt{10}$$

nel semipiano sinistro.

Se invece:

$$L(s) = \frac{10}{(1+s)^3},$$

il diagramma di Nyquist qualitativo si può tracciare di nuovo facilmente (si osservi che ora la fase della risposta in frequenza termina con il valore -270°).

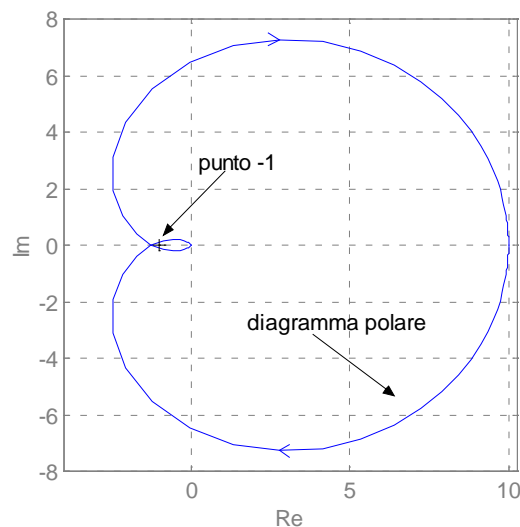


Fig. 80 : Diagramma di Nyquist di L

Anche in questo caso $P_d=0$, mentre per determinare il valore di N occorre stabilire dove si trova il punto P in cui il diagramma attraversa l'asse reale. Tale punto può essere caratterizzato come quello in cui la parte immaginaria della risposta in frequenza si annulla o come quello in cui la fase della risposta in frequenza vale -180° . Seguendo quest'ultima strada, e denominando ω_p la pulsazione cui è associato il punto P , ricaviamo ω_p dall'equazione:

$$\angle L(j\omega_p) = -180^\circ \Rightarrow -3\angle(1+j\omega_p) = -180^\circ \Rightarrow \arctan(\omega_p) = 60^\circ \Rightarrow \omega_p = \sqrt{3}.$$

Poiché:

$$\left|L(j\omega_p)\right| = \frac{10}{\left|1+j\omega_p\right|^3} = \frac{10}{\left(\sqrt{1+\omega_p^2}\right)^3} = \frac{10}{8} > 1,$$

il punto P si trova a sinistra del punto -1 , attorno al quale il diagramma compie quindi due giri in senso orario. Pertanto $N = -2 \neq P_d$ ed il sistema in anello chiuso non è asintoticamente stabile.

Per verifica, osserviamo che il polinomio caratteristico in anello chiuso è il seguente:

$$\chi(s) = 10 + (1+s)^3 = s^3 + 3s^2 + 3s + 11,$$

ed ha due radici:

$$s_{1,2} = 0.07 \pm j1.86, \quad s_3 = -3.15$$

nel semipiano destro.

Il criterio di Bode

Si consideri nuovamente il sistema in anello chiuso di Fig. 78.

Introduciamo le seguenti *ipotesi* sulla funzione di trasferimento d'anello $L(s)$:

1. $L(s)$ non ha poli a parte reale positiva.
2. Il diagramma di Bode del modulo di $L(j\omega)$ interseca l'asse a 0 dB una e una sola volta.

Diamo le seguenti *definizioni*:

Pulsazione critica ω_c : pulsazione alla quale il diagramma di $|L(j\omega)|_{dB}$ taglia l'asse a 0 dB, ossia: $|L(j\omega_c)|=1$;

Fase critica ϕ_c : fase di $L(j\omega)$ in corrispondenza della pulsazione critica, ossia $\phi_c = \angle L(j\omega_c)$;

Margine di fase ϕ_m : differenza tra 180° e la fase critica, presa in modulo, ossia: $\phi_m = 180^\circ - |\phi_c|$;

Guadagno d'anello μ_L : guadagno (generalizzato) di $L(s)$.

Il **criterio di Bode** afferma che il sistema in anello chiuso è asintoticamente stabile **se e solo se** il guadagno d'anello ed il margine di fase sono entrambi positivi:

$$\begin{cases} \mu_L > 0 \\ \phi_m > 0 \end{cases}$$

Esempio

Sia:

$$L(s) = \frac{10}{(1+s)^2}.$$

Il diagramma di Bode asintotico del modulo di L è riportato in figura:

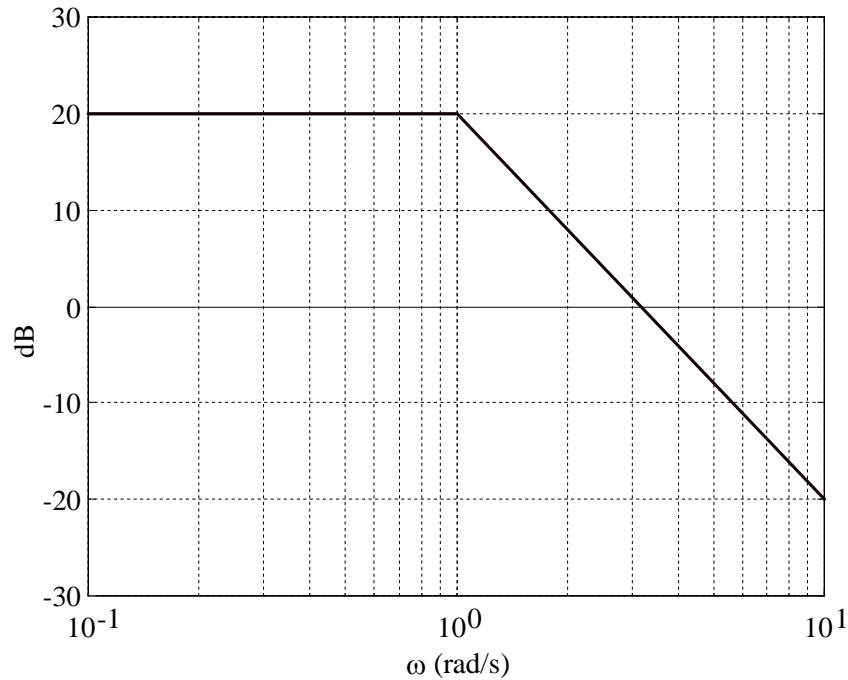


Fig. 81 : Diagramma di Bode asintotico di $|L|$

La pulsazione critica vale $\omega_c \approx 3 \text{ rad/s}$. La fase critica può essere calcolata analiticamente.

La formula che si può utilizzare nel caso generale (assumendo poli e zeri reali in $L(s)$) è la seguente:

$$\begin{aligned}\phi_c = \angle L(j\omega_c) &= \angle \mu_L - g_L 90^\circ + \sum_i \angle(1 + j\omega_c \tau_i) - \sum_k \angle(1 + j\omega_c \tau_k) = \\ &= \angle \mu_L - g_L 90^\circ + \text{sgn}(\tau_i) \sum_i \arctan\left(\frac{\omega_c}{\omega_{zi}}\right) - \sum_k \arctan\left(\frac{\omega_c}{\omega_{pk}}\right)\end{aligned}$$

dove τ_i e τ_k sono le costanti di tempo degli zeri e dei poli di $L(s)$ (queste ultime necessariamente positive per le ipotesi di applicabilità del criterio di Bode), mentre ω_{zi} e ω_{pk} sono le relative pulsazioni.

Risulta nel nostro caso

$$\phi_c = -2 \arctan\left(\frac{\omega_c}{1}\right) = -2 \arctan(3) = -2 \times 72^\circ = -144^\circ.$$

Il margine di fase è quindi:

$$\phi_m = 180^\circ - |\phi_c| = 36^\circ > 0.$$

Poiché anche il guadagno d'anello (pari a 10) è positivo, il sistema in anello chiuso è asintoticamente stabile, coerentemente con quanto determinato con il criterio di Nyquist.

Se invece:

$$L(s) = \frac{10}{(1+s)^3},$$

il diagramma asintotico del modulo di L è il seguente:

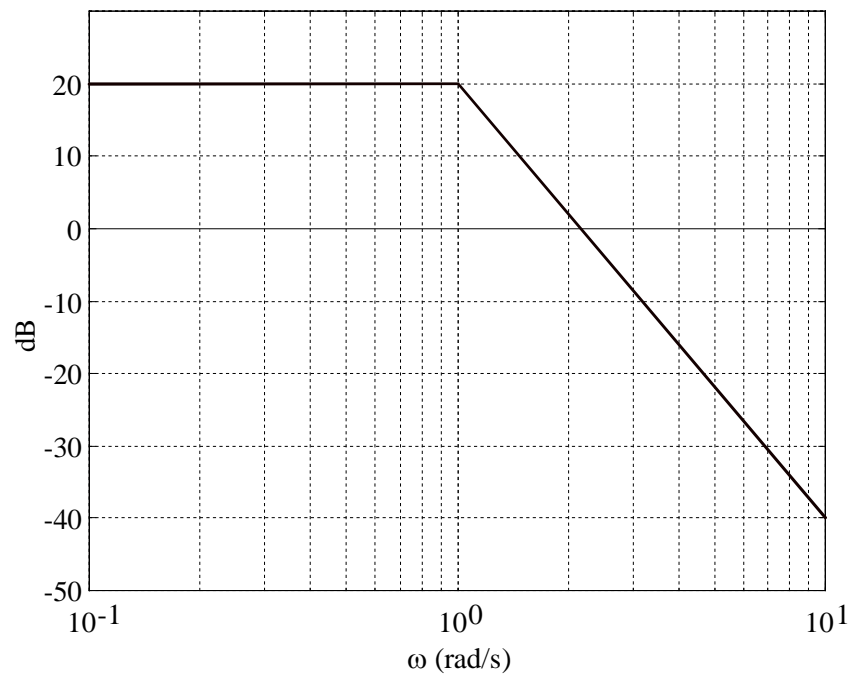


Fig. 82 : Diagramma di Bode di $|L|$

La pulsazione critica vale $\omega_c \approx 2$ rad / s. La fase critica:

$$\phi_c = -3 \arctan(2) = -3 \times 64^\circ = -192^\circ.$$

Il margine di fase è quindi:

$$\phi_m = 180^\circ - |\phi_c| = -12^\circ < 0.$$

Pertanto il sistema in anello chiuso non è asintoticamente stabile, coerentemente con quanto determinato con il criterio di Nyquist

Osservazioni sul criterio di Bode

1) Giustificazione

Il criterio di Bode è un caso particolare del criterio di Nyquist: non è infatti difficile verificare che nelle condizioni di applicabilità del criterio di Bode (che corrispondono a $P_d = 0$), le condizioni sul guadagno d'anello e sul margine di fase espresse dal criterio equivalgono a garantire l'assenza di giri del diagramma di Nyquist intorno al punto -1 ($N = 0$).

Rispetto al criterio di Nyquist, il criterio di Bode ha il vantaggio di richiedere il tracciamento dei diagrammi di Bode, di norma più agevole del tracciamento del diagramma di Nyquist.

2) Uso dei diagrammi asintotici

La pulsazione critica può essere stimata con buona precisione utilizzando il diagramma di Bode asintotico del modulo, a patto che non vi siano cambiamenti di pendenza nelle vicinanze della pulsazione critica stessa. Occorre quindi tenere presente che più vicini alla pulsazione critica sono i cambiamenti di pendenza, più inaccurata sarà la stima della pulsazione critica.

Inoltre gli eventuali poli o zeri complessi e coniugati devono essere ad alto smorzamento, altrimenti i diagrammi asintotici possono scostarsi sensibilmente da quelli reali.

3) Sistemi a fase minima

Un sistema dinamico si dice a **fase minima** se ha guadagno positivo e tutte le singolarità (poli e zeri) sono nel semipiano sinistro chiuso del piano complesso. Tale dizione deriva dal fatto che, se si tracciano i diagrammi di Bode della fase di tutte le funzioni di trasferimento a guadagno positivo e con poli nel semipiano sinistro chiuso, aventi lo stesso diagramma del modulo della funzione di trasferimento a fase minima, in quella a fase minima la fase si trova, per ciascuna pulsazione, ad un valore di ordinata superiore (e quindi, se le fasi come spesso avviene sono negative, la fase è più piccola in modulo).

Per un sistema a fase minima, ad ogni polo corrisponde un decremento unitario della pendenza del diagramma del modulo ed un decremento di 90° del valore della fase, e viceversa per gli zeri. Ne consegue che il diagramma asintotico della fase si può immediatamente ottenere da quello del modulo, semplicemente moltiplicandone in ogni tratto la pendenza per 90° .

Se allora il diagramma asintotico del modulo, in corrispondenza al **taglio** dell'asse a 0 dB ha **pendenza -1** , e non vi sono cambiamenti di pendenza nelle immediate vicinanze della pulsazione critica, allora la fase critica sarà prossima al valore asintotico (-90°). Ne consegue che il sistema in anello chiuso, essendo il margine di fase ampiamente positivo, sarà asintoticamente stabile.

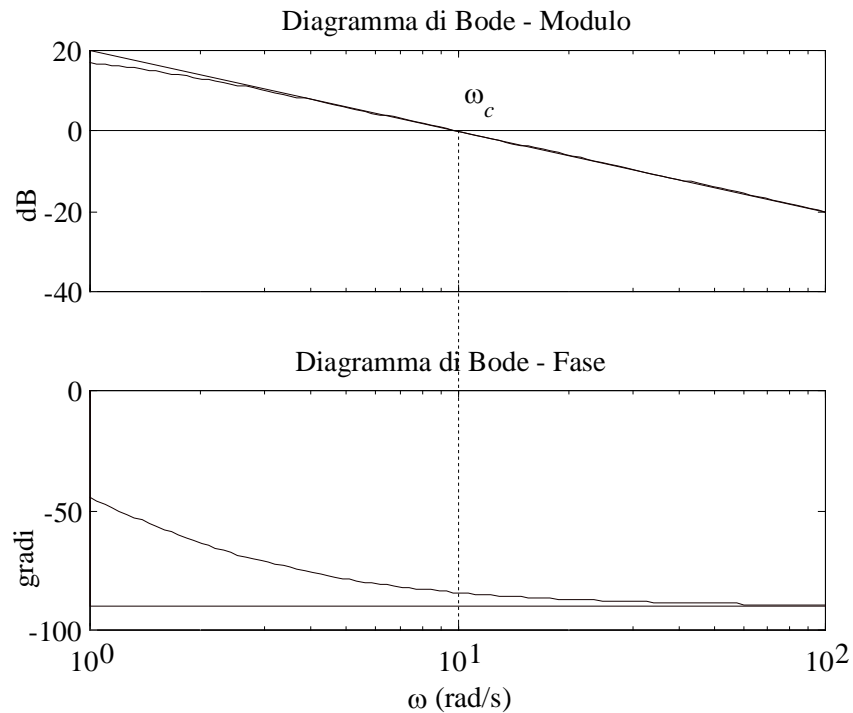


Fig. 83 : Taglio dell'asse a 0 dB con pendenza -1 per un sistema a fase minima

4) Sistemi con ritardo

Un sistema in cui ad ogni istante l'uscita ha il valore assunto dall'ingresso τ istanti prima:

$$y(t) = u(t - \tau) ,$$

prende il nome di **ritardo** (si ricordi l'esempio del nastro trasportatore nella Lezione 2).

Trasformando secondo Laplace entrambi i membri dell'equazione si ha:

$$Y(s) = e^{-s\tau} U(s) ,$$

per cui la **funzione di trasferimento** del ritardo è la seguente:

$$G(s) = e^{-s\tau} .$$

La **risposta in frequenza** associata a questa funzione di trasferimento è quindi:

$$G(j\omega) = e^{-j\omega\tau} = \cos(\omega\tau) - j \sin(\omega\tau) .$$

Poiché risulta:

$$|G(j\omega)| = |e^{-j\omega\tau}| = \sqrt{\cos^2(\omega\tau) + \sin^2(\omega\tau)} = 1 ,$$

$$\angle G(j\omega) = \angle e^{-j\omega\tau} = \arctan(-\sin(\omega\tau)/\cos(\omega\tau)) = -\omega\tau ,$$

il diagramma di Bode del modulo coincide con l'asse a 0 dB, mentre la fase decresce linearmente con ω (sulla scala logaritmica si ha una curva monotona decrescente).

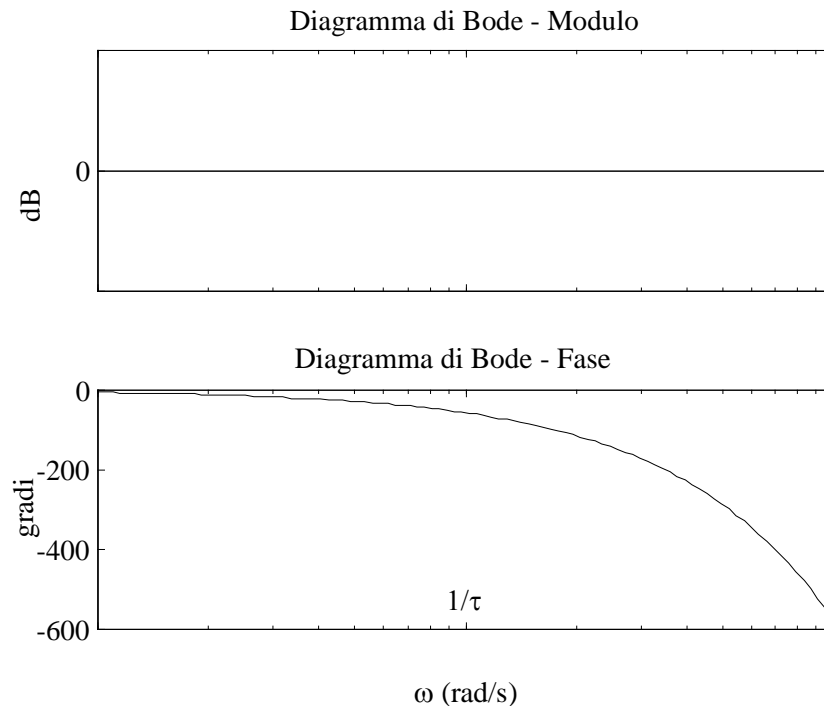


Fig. 84 : Diagrammi di Bode del ritardo

E' poi molto comune il caso in cui la funzione di trasferimento d'anello sia il prodotto di una funzione di trasferimento razionale $L_r(s)$ e della funzione di trasferimento del ritardo:

$$L(s) = L_r(s)e^{-s\tau}.$$

Si dimostra che, se il criterio di Bode è applicabile per la funzione di trasferimento $L_r(s)$ (cioè se tale funzione di trasferimento ne rispetta le ipotesi), allora rimane applicabile anche per la funzione di trasferimento $L(s)$.

Osservando poi che risulta:

$$\begin{aligned} |L(j\omega)| &= |L_r(j\omega)e^{-j\omega\tau}| = |L_r(j\omega)| |e^{-j\omega\tau}| = |L_r(j\omega)|, \\ \angle L(j\omega) &= \angle(L_r(j\omega)e^{-j\omega\tau}) = \angle L_r(j\omega) + \angle e^{-j\omega\tau} = \angle L_r(j\omega) - \omega\tau, \end{aligned}$$

si conclude che la presenza del ritardo non altera il diagramma di Bode del modulo, per cui la pulsazione critica si può ricavare direttamente dall'analisi della funzione di trasferimento priva di ritardo. Per quanto riguarda la fase critica, al termine dovuto alla parte razionale della funzione di trasferimento occorrerà sommare un termine pari a:

$$-\omega_c \tau \frac{180^\circ}{\pi}.$$

Si osservi la conversione da radianti a gradi, necessaria per rendere questo termine sommabile alle fasi espresse in gradi.

Esempio

Sia:

$$L(s) = \frac{10}{(1+s)(1+10s)} e^{-s\tau}, \quad \tau > 0.$$

Il diagramma di Bode del modulo è riportato in figura.

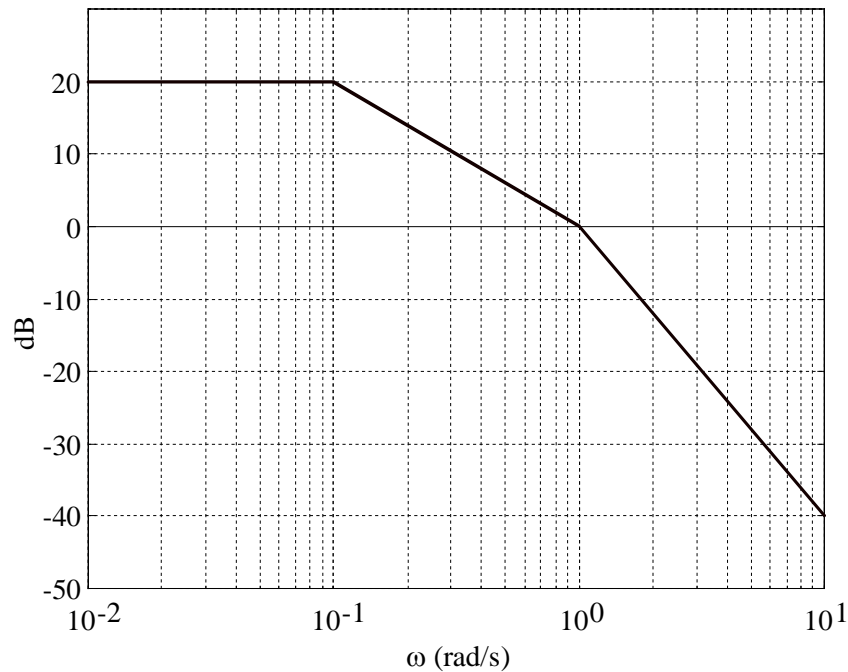


Fig. 85 : Diagramma di Bode asintotico del modulo di L

La pulsazione critica vale approssimativamente $\omega_c = 1$ rad/s.

La fase critica risulta:

$$\varphi_c \approx -\arctan\left(\frac{1}{0.1}\right) - \arctan\left(\frac{1}{1}\right) - \omega_c \tau \frac{180^\circ}{\pi} \approx -84^\circ - 45^\circ - 57^\circ \tau = -129^\circ - 57^\circ \tau.$$

Il margine di fase vale quindi:

$$\varphi_m = 180^\circ - |\varphi_c| = 180^\circ - |-129^\circ - 57^\circ \tau| = 180^\circ - (129^\circ + 57^\circ \tau) = 51^\circ - 57^\circ \tau.$$

Pertanto se, per esempio, $\tau = 0.1$, il margine di fase è positivo ed il sistema in anello chiuso è asintoticamente stabile; se $\tau = 2$, il margine di fase è negativo ed il sistema in anello chiuso non è asintoticamente stabile.

Margine di guadagno

Il margine di fase, oltre a dare un'indicazione binaria sull'asintotica stabilità del sistema in anello chiuso, quantifica anche la **robustezza della stabilità**, ossia il margine di sicurezza con cui possiamo tollerare incertezze sul modello senza compromettere l'asintotica stabilità. Più alto è il margine di fase, più robusto è il sistema. Vi sono tuttavia casi in cui il margine di fase non costituisce un indicatore attendibile, come riportato nell'esempio in figura:

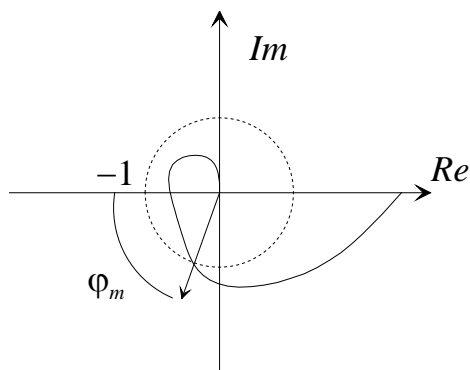


Fig. 86 : Diagramma polare di L ad alto margine di fase ma bassa robustezza

Pur essendo il margine di fase elevato, il diagramma polare passa molto vicino al punto -1 , rendendo il sistema scarsamente robusto a fronte di incertezze di modello.

Un indicatore da usare in congiunzione con il margine di fase è il margine di guadagno. Sempre nell'ipotesi $P_d=0$, si supponga che il diagramma polare di L attraversi il semiasse reale negativo in uno e un solo punto:

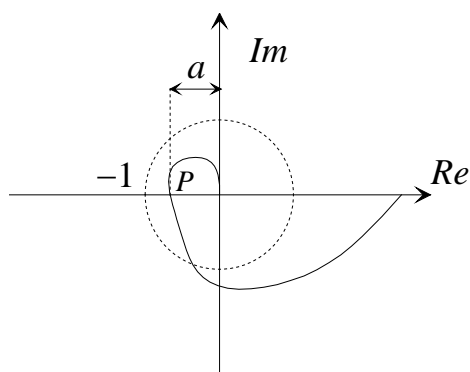


Fig. 87 : Diagramma polare di L con indicazione della distanza del punto P dall'origine

Un indice di robustezza è la distanza del punto P di intersezione dal punto -1 , ovvero la vicinanza del punto P all'origine. Detta allora a la distanza del punto dall'origine, definiamo **margine di guadagno** la quantità:

$$k_m = \frac{1}{a} = \frac{1}{|L(j\omega_p)|}, \quad \text{con } \angle L(j\omega_p) = -180^\circ.$$

Il sistema in anello chiuso è asintoticamente stabile se $k_m > 1$, ed è tanto più robusto quanto maggiore è k_m .

Capitolo 10: Prestazioni dinamiche dei sistemi di controllo

Le prestazioni dinamiche fanno riferimento al comportamento del sistema di controllo durante i transitori, ossia alla modalità con cui le variabili del sistema, ed in particolare la variabile controllata, passano da una condizione di regime ad una nuova, a seguito di variazioni degli ingressi.

Sono di particolare importanza, a questo riguardo:

- la **velocità di risposta**, ovvero la rapidità con cui la variabile controllata segue brusche variazioni (per esempio a scalino) del riferimento;
- lo **smorzamento dei transitori**, ovvero l'assenza o l'irrilevanza di oscillazioni nel transitorio.

Con riferimento ai parametri con cui si era caratterizzata la risposta allo scalino di un sistema dinamico, potremo dire che la velocità di risposta corrisponde ad avere un tempo di salita della risposta allo scalino ridotto mentre lo smorzamento dei transitori corrisponde a sovraelongazione massima e tempo di assestamento contenuti.

Rientrano inoltre nel novero delle prestazioni dinamiche anche la reiezione dei disturbi, siano essi sulla linea di andata o su quella di retroazione, e la moderazione del controllo, ossia la proprietà del sistema di controllo per cui la variabile di controllo non è sottoposta ad eccessive sollecitazioni.

Banda passante

Si consideri un sistema a costante di tempo:

$$H(s) = \frac{1}{1 + sT}, \quad T > 0.$$

E' noto che la risposta allo scalino di questo sistema è tanto più rapida quanto più piccola è la costante di tempo T . Detta $\omega_H = 1/T$ la pulsazione del polo, possiamo equivalentemente dire che la risposta è tanto più veloce quanto più in alta frequenza è la pulsazione del polo.

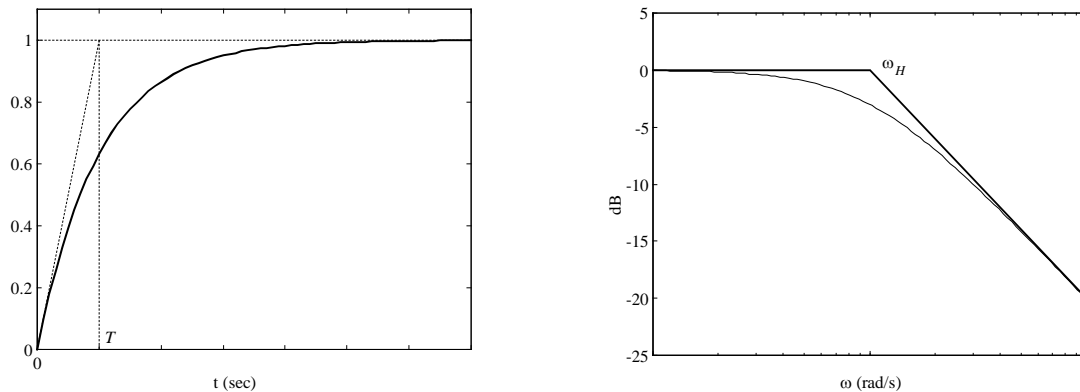


Fig. 88 : Risposta allo scalino e diagramma di Bode del modulo

Pertanto la pulsazione del polo è un buon indice della velocità di risposta del sistema.

Si consideri ora un sistema di controllo in anello chiuso, ed in particolare la funzione di trasferimento dal riferimento y^o alla variabile controllata y :

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

In virtù dell'imposizione dei requisiti statici, $F(s)$ varrà 1, o comunque un valore prossimo a 1, in bassa frequenza (cioè per $s \rightarrow 0$). Inoltre il sistema di controllo sarà progettato in modo tale che la sua funzione di trasferimento rifletta la caratteristica, propria dei sistemi fisici, di avere più poli che zeri. Da queste considerazioni seguono le due caratteristiche fondamentali del diagramma di Bode del modulo di F :

- per $\omega \rightarrow 0$, $|F(j\omega)|_{dB} \approx 0$;
- per $\omega \rightarrow \infty$, $|F(j\omega)|_{dB} \rightarrow -\infty$.

Andamenti plausibili del modulo di F potranno pertanto essere quelli riportati in figura:

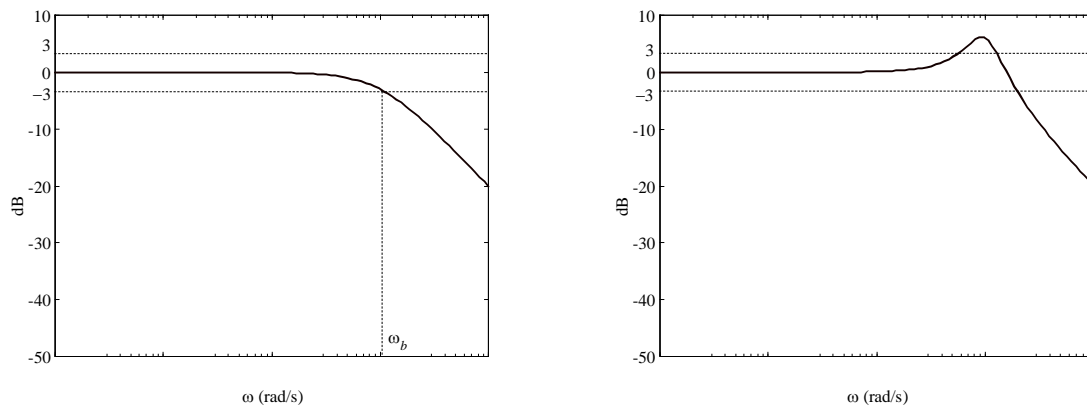


Fig. 89 : Tipici andamenti del diagramma di Bode del modulo di F

Se il diagramma di Bode del modulo di F non supera per nessuna pulsazione il valore 3 dB, si definisce **banda passante** del sistema di controllo l'insieme delle pulsazioni $[0, \omega_b]$, essendo ω_b la pulsazione alla quale il modulo vale -3 dB.

Si osservi che la condizione espressa preliminarmente alla definizione di banda passante esclude la presenza di rilevanti picchi o rigonfiamenti nel diagramma del modulo, e quindi le situazioni rappresentate dal diagramma di destra in Fig. 2. In altre parole il sistema di controllo si comporta da **filtro passabasso**.

Se la pendenza del diagramma di $|F|_{dB}$ dopo la pulsazione ω_b vale -1, il sistema di controllo si comporta in prima approssimazione come un **sistema del primo ordine**, con pulsazione del polo pari a ω_b . Ne consegue che i transitori del sistema in anello chiuso saranno caratterizzati da una costante di tempo approssimativamente pari a:

$$\tau = \frac{1}{\omega_b} .$$

Il sistema sarà quindi tanto più veloce quanto più estesa è la sua banda passante, ossia l'estremo superiore della banda passante è un buon indice della **velocità di risposta** del sistema di controllo.

Velocità di risposta e pulsazione critica

Ci si pone ora l'obiettivo di individuare un indice di velocità di risposta legato alla funzione di trasferimento d'anello $L(s)$, piuttosto che alla funzione di trasferimento in anello chiuso $F(s)$.

Ricordiamo che risulta:

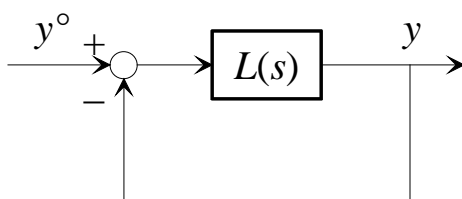


Fig. 90 : Sistema di controllo privo di disturbi

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

Il legame tra le risposte in frequenza è quindi:

$$F(j\omega) = \frac{L(j\omega)}{1 + L(j\omega)},$$

ed in particolare:

$$|F(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|}.$$

Consideriamo ora la seguente **approssimazione**:

$$|F(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|} \approx \begin{cases} 1 & \forall \omega: |L(j\omega)| \gg 1 \\ |L(j\omega)| & \forall \omega: |L(j\omega)| \ll 1 \end{cases}.$$

Tracciato il diagramma di $|L|$, l'approssimazione consiste nell'attribuire a $|F|$ il valore 1 per quelle pulsazioni per cui $|L|$ è decisamente sopra l'asse a 0 dB, o il valore di $|L|$ quando $|L|$ è decisamente sotto l'asse a 0 dB.

Poniamoci ora nelle ipotesi di **applicabilità del criterio di Bode**.

Considerando un tipico andamento di $|L|$, si avrà la situazione riportata in Fig. 91.

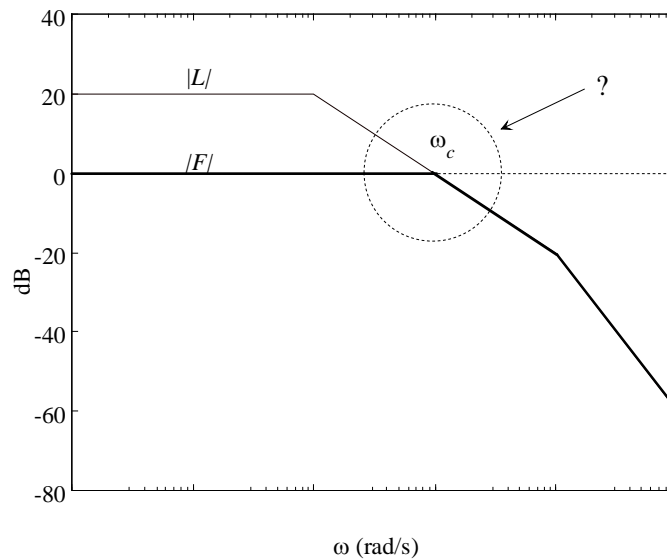


Fig. 91 : Metodo grafico per il tracciamento del diagramma asintotico del modulo di F a partire da quello di L

L' approssimazione consiste allora nel porre:

$$|F(j\omega)| \approx \begin{cases} 1 & \forall \omega \ll \omega_c \\ |L(j\omega)| & \forall \omega \gg \omega_c \end{cases} .$$

L'approssimazione è evidentemente migliore a pulsazioni lontane dalla pulsazione critica ω_c .

Possiamo quindi concludere che F si comporta da filtro passabasso a guadagno unitario, ma non siamo ancora in grado di dire se presenta risonanza (ossia se il modulo di F presenta rigonfiamenti). Di conseguenza non è ancora chiaro se sia più opportuno approssimare F con un sistema del primo ordine o con uno del second'ordine a poli complessi.

Smorzamento e margine di fase

Il calcolo esatto del modulo di F per $\omega=\omega_c$ conduce al seguente risultato:

$$\begin{aligned} |F(j\omega_c)| &= \frac{|L(j\omega_c)|}{|1+L(j\omega_c)|} = \frac{1}{|1+e^{j\varphi_c}|} = \frac{1}{|1+\cos\varphi_c+j\sin\varphi_c|} = \\ &= \frac{1}{\sqrt{1+\cos^2\varphi_c+2\cos\varphi_c+\sin^2\varphi_c}} = \frac{1}{\sqrt{2(1+\cos\varphi_c)}} = \frac{1}{\sqrt{2(1-\cos\varphi_m)}} = \frac{1}{2\sin\left(\frac{\varphi_m}{2}\right)}. \end{aligned}$$

In particolare, per $\varphi_m = 90^\circ$, risulta:

$$|F(j\omega_c)| = \frac{1}{\sqrt{2}} \Rightarrow |F(j\omega_c)|_{dB} = -3,$$

per cui $\omega_c = \omega_b$, ossia la **pulsazione critica** coincide con l'**estremo superiore della banda passante**.

Peraltro per valori inferiori a 90° del margine di fase, ma comunque elevati, il modulo di F in ω_c rimane prossimo al valore precedentemente ottenuto: in particolare rimane inferiore a 1 per valori di φ_m superiori a 60° . Considerando che valori di margine di fase elevato sono di norma associati ad attraversamenti dell'asse delle pulsazioni con pendenza del modulo di L pari a -1 , possiamo concludere che in questo caso ($\varphi_m > 60^\circ$) è adeguata un'approssimazione di F con una funzione di trasferimento del primo ordine a **costante di tempo**, con costante di tempo pari all'inverso della pulsazione critica:

$$F(s) = \frac{1}{1+s/\omega_c}.$$

Il transitorio quindi si assesterà al 99% del valore di regime dopo un tempo pari a circa $4.6/\omega_c$.

Per valori inferiori del margine di fase ($\varphi_m < 60^\circ$) appare più adeguata un'approssimazione del secondo ordine con **poli complessi coniugati** alla pulsazione ω_c :

$$F(s) \approx \frac{\omega_c^2}{s^2 + 2\zeta\omega_c s + \omega_c^2}.$$

Lo smorzamento si può determinare confrontando il valore assunto in ω_c dal modulo della risposta in frequenza dell'approssimante con il valore esatto determinato precedentemente:

$$|F(j\omega_c)| = \frac{1}{2\zeta} = \frac{1}{2\sin\left(\frac{\varphi_m}{2}\right)},$$

da cui:

$$\zeta = \sin\left(\frac{\varphi_m}{2}\right) \approx \frac{\varphi_m}{100},$$

dove l'ultima approssimazione è valida se φ_m è espresso in gradi.

Il transitorio quindi si assesterà al 99% del valore di regime dopo un tempo pari a circa

$$4.6/(\zeta\omega_c).$$

Si osservi che il risultato trovato stabilisce un'importante relazione tra lo smorzamento dei transitori in anello chiuso ed il margine di fase, il quale, quindi, oltre a costituire un indice di robustezza della stabilità, qualifica anche il “grado di stabilità” del sistema.

Esempio

Sia:

$$L(s) = \frac{100}{s(1+0.0025s)}.$$

Si vuole tracciare l'andamento qualitativo della risposta allo scalino della funzione di trasferimento $F(s)$.

Il diagramma di Bode di $|L(j\omega)|$ è riportato in figura:

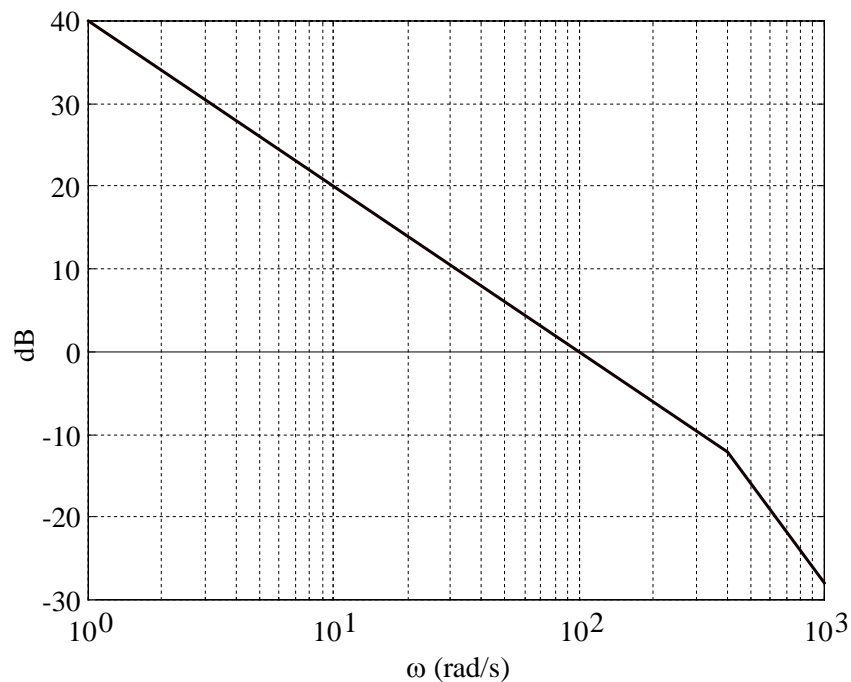


Fig. 92 : Diagramma di Bode asintotico di $|L|$

La pulsazione critica è $\omega_c = 100$ rad/s, mentre il margine di fase risulta:

$$\varphi_m = 180^\circ - |\varphi_c| = 180^\circ - \left| -90^\circ - \arctan\left(\frac{100}{400}\right) \right| = 180^\circ - |-90^\circ - 14^\circ| = 76^\circ.$$

E' quindi possibile procedere all'approssimazione:

$$F(s) \approx \frac{1}{1 + \frac{s}{\omega_c}} = \frac{1}{1 + 0.01s}.$$

La risposta allo scalino approssimata di $F(s)$ è riportata in Fig. 93, dove viene confrontata con la risposta effettiva del sistema di controllo:

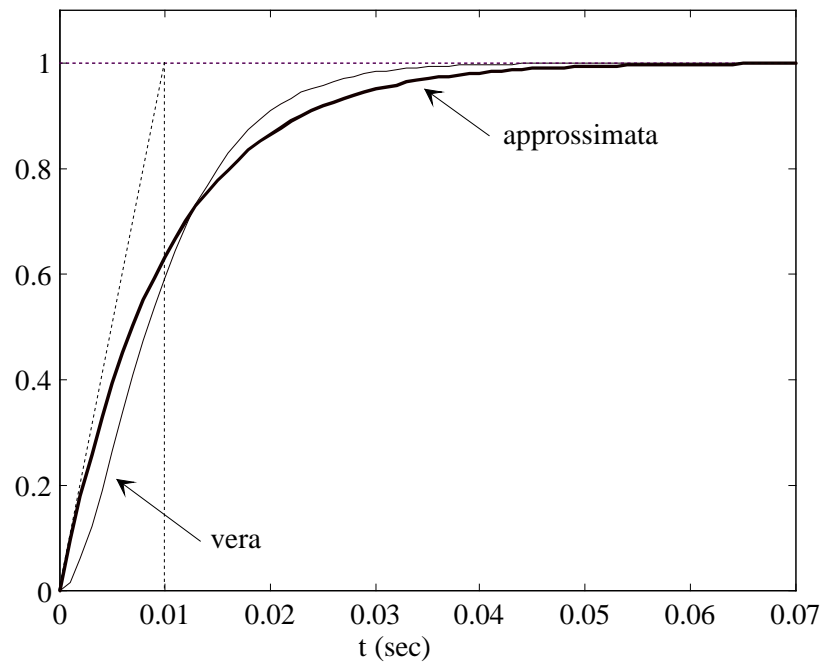


Fig. 93 : Risposta allo scalino di F

Reiezione dei disturbi

Disturbo in linea di andata

Consideriamo un disturbo d in linea di andata:

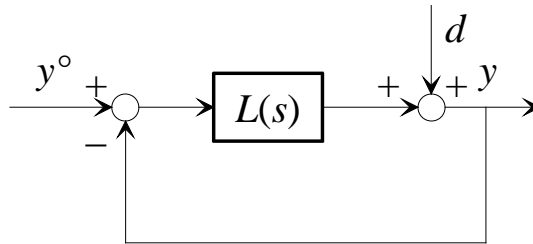


Fig. 94 : Sistema di controllo con disturbo in linea di andata

La funzione di trasferimento dal disturbo d alla variabile controllata y è data da:

$$\frac{Y(s)}{D(s)} = S(s) = \frac{1}{1 + L(s)},$$

dove S prende il nome di **funzione di sensitività**.

Possiamo **approssimare** il modulo della risposta in frequenza come fatto prima:

$$|S(j\omega)| = \frac{1}{|1 + L(j\omega)|} \approx \begin{cases} \frac{1}{|L(j\omega)|} & \forall \omega: |L(j\omega)| \gg 1 \\ 1 & \forall \omega: |L(j\omega)| \ll 1 \end{cases},$$

e, nelle ipotesi di applicabilità del criterio di Bode,

$$|S(j\omega)| = \frac{1}{|1 + L(j\omega)|} \approx \begin{cases} \frac{1}{|L(j\omega)|} & \forall \omega \ll \omega_c \\ 1 & \forall \omega \gg \omega_c \end{cases}$$

Tracciato il diagramma di $|L|$, quello di $1/|L|$ si ottiene per ribaltamento rispetto all'asse a 0 dB. L'approssimazione consiste nell'attribuire a $|S|$ il valore 1 per quelle pulsazioni per cui $|L|$ è decisamente sotto l'asse a 0 dB, o il valore di $1/|L|$ quando $|L|$ è decisamente sopra l'asse a 0 dB. Considerando un tipico andamento di $|L|$, si avrà la situazione riportata in figura (con le stesse considerazioni fatte prima circa la validità delle approssimazioni).

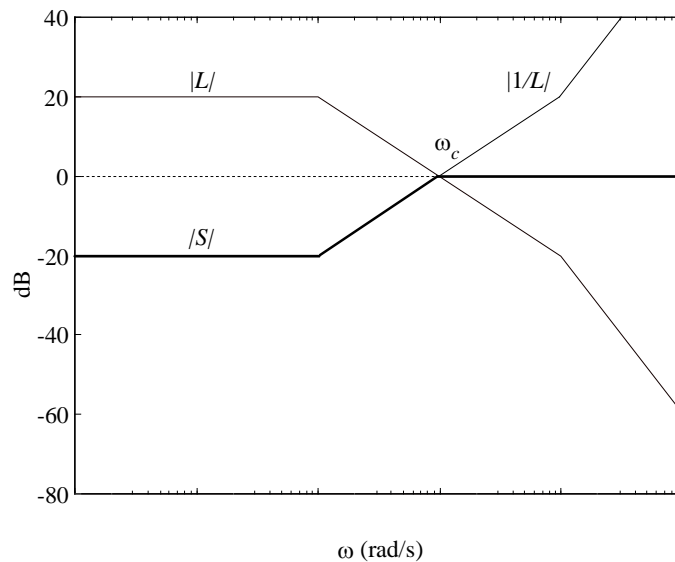


Fig. 95 : Costruzione grafica del diagramma asintotico di $|S|$

Pertanto le componenti armoniche (sinusoidi) del disturbo a pulsazioni inferiori alla pulsazione critica, e quindi interne alla banda passante, sono fortemente attenuate sulla variabile controllata. L'attenuazione è tanto maggiore quanto più alto è il valore del modulo di L sulla banda passante.

Si conclude quindi che per un'efficace reiezione dei disturbi in linea di andata la **banda passante** deve essere **sufficientemente ampia** da contenere le armoniche significative del disturbo.

Disturbo in linea di retroazione

Consideriamo un disturbo n in linea di retroazione:

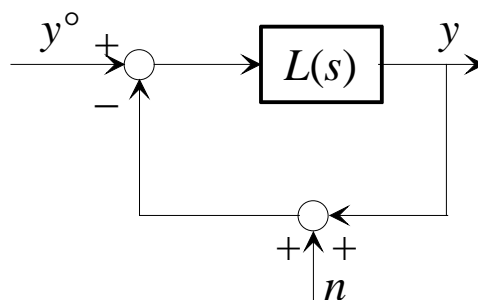


Fig. 96 : Sistema di controllo con disturbo in linea di retroazione

La funzione di trasferimento dal disturbo n alla variabile controllata y è data da:

$$\frac{Y(s)}{N(s)} = -F(s) = -\frac{L(s)}{1 + L(s)},$$

La funzione di trasferimento $F(s)$, che prende il nome di **funzione di sensitività complementare**⁷, è la stessa esistente tra riferimento e variabile controllata. Sappiamo quindi già come ottenerne il modulo a partire da quello di L (si veda la Fig. 4).

Possiamo concludere quindi che tutte le componenti armoniche del disturbo esterne alla banda passante del sistema di controllo vengono attenuate, mentre quelle interne passano. Ne consegue che per un'efficace reiezione dei disturbi in linea di retroazione (che tipicamente si caratterizzano come rumore di alta frequenza), è bene che la **banda passante non** sia **eccessivamente ampia**.

⁷ Si osservi che $F(s) = 1 - S(s)$

Moderazione del controllo

Consideriamo un sistema di controllo in cui mettiamo in evidenza la variabile di controllo:

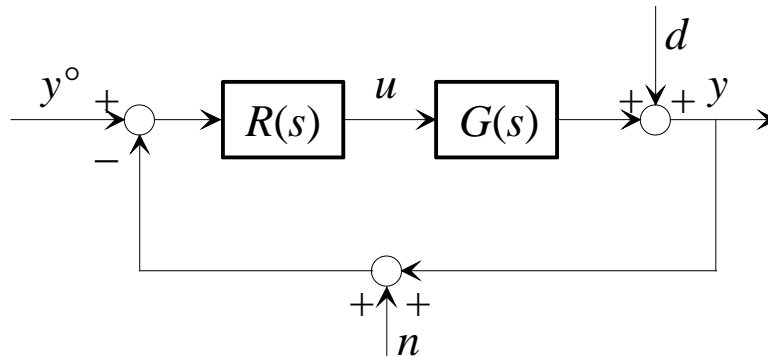


Fig. 97 : Sistema di controllo con in evidenza la variabile di controllo

La funzione di trasferimento dal riferimento y^o alla variabile di controllo u è data da:

$$\frac{U(s)}{Y^o(s)} = Q(s) = \frac{R(s)}{1 + L(s)},$$

dove Q prende il nome di **funzione di sensitività del controllo**.

Si osservi che, a meno del segno, questa è anche la funzione di trasferimento dal disturbo d e dal disturbo n alla variabile di controllo. Poiché il riferimento e i due disturbi hanno caratteristiche armoniche differenti, è bene che la funzione di trasferimento attenui (o non amplifichi) su tutto l'asse delle pulsazioni.

Possiamo approssimare il modulo della risposta in frequenza:

$$|Q(j\omega)| = \frac{|R(j\omega)|}{|1 + L(j\omega)|} \approx \begin{cases} \frac{1}{|G(j\omega)|} & \forall \omega: |L(j\omega)| \gg 1 \quad (\omega \ll \omega_c) \\ |R(j\omega)| & \forall \omega: |L(j\omega)| \ll 1 \quad (\omega \gg \omega_c) \end{cases}.$$

Poiché fuori dalla banda passante ($\omega > \omega_c$) il modulo di Q coincide con quello di R , è bene che il **regolatore** sia progettato in modo che il suo **modulo non assuma valori eccessivi in alta frequenza**.

Per quanto riguarda invece il comportamento in banda passante, dobbiamo ipotizzare una tipologia di risposta in frequenza per G : ipotizziamo un comportamento di tipo filtro passabasso, con un certo guadagno maggiore di 1 e pulsazione di taglio pari a ω_G . E' allora abbastanza evidente che se ω_c è molto maggiore di ω_G , il diagramma di $1/|G|$, cui $|Q|$ coincide in banda, può assumere valori molto elevati.

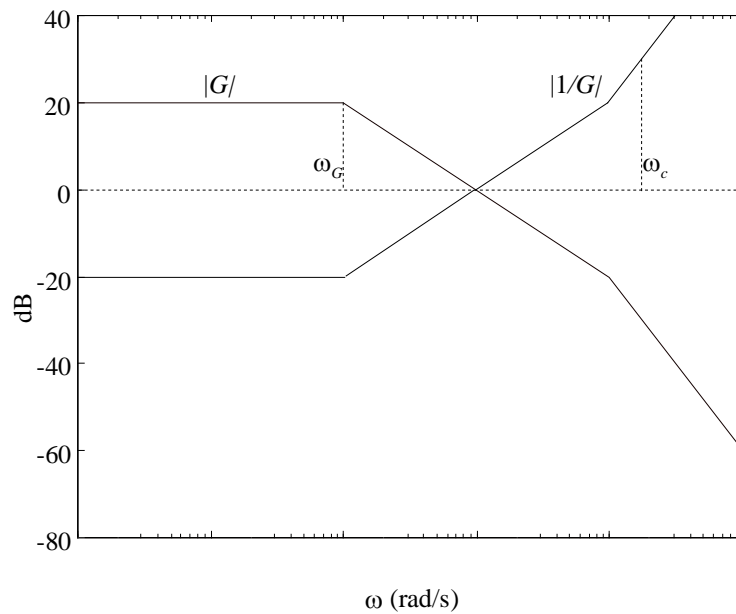


Fig. 98 : Banda in anello aperto e in anello chiuso

Concludiamo allora che per garantire moderazione al controllo la **banda passante non deve essere eccessivamente ampia rispetto alla banda che caratterizza la dinamica in anello aperto.**

Capitolo 11: Prestazioni statiche dei sistemi di controllo

Le prestazioni statiche di un sistema di controllo fanno riferimento al suo comportamento a **transitorio esaurito**, ossia alla situazione in cui il sistema, dopo un transitorio dovuto alla variazione dei suoi ingressi, si è portato in una condizione di regime. In particolare saremo interessati, in questa condizione, all'**errore** tra il segnale di riferimento e la variabile controllata.

Prerequisito del sistema di controllo, necessario per poter parlare di prestazioni statiche, è evidentemente l'**asintotica stabilità** del sistema in anello chiuso.

Prenderemo in considerazione il seguente sistema di controllo, supposto asintoticamente stabile:

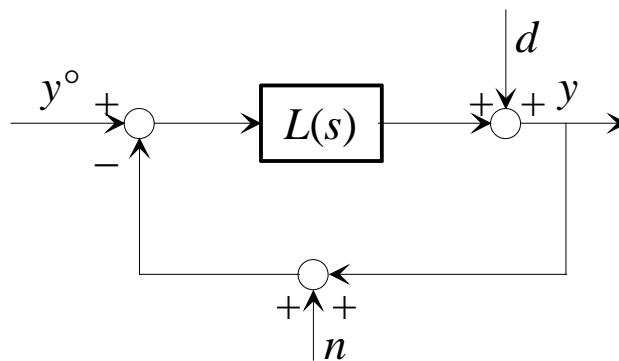


Fig. 99 : Sistema di controllo

Si osservi che con una rielaborazione formale dello schema a blocchi è possibile mettere direttamente in evidenza l'errore tra y^o e y :

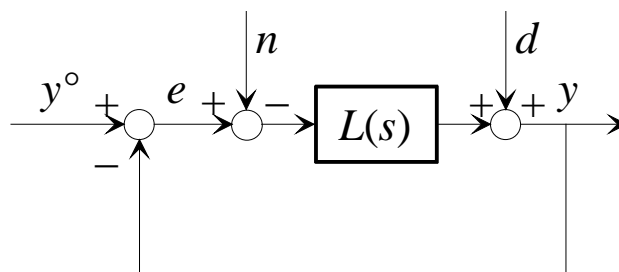


Fig. 100 : Sistema di controllo con in evidenza l'errore

Per lo studio delle prestazioni statiche è sufficiente riferirsi ad un insieme dei segnali di ingresso ristretto ai cosiddetti **segnali canonici**, come lo scalino, la rampa, la parabola ecc.

Infatti, ai fini della valutazione dell'errore a regime, sono del tutto irrilevanti le eventuali variazioni transitorie subite dal segnale di ingresso, del quale riveste interesse solo il comportamento asintotico ($t \rightarrow \infty$).

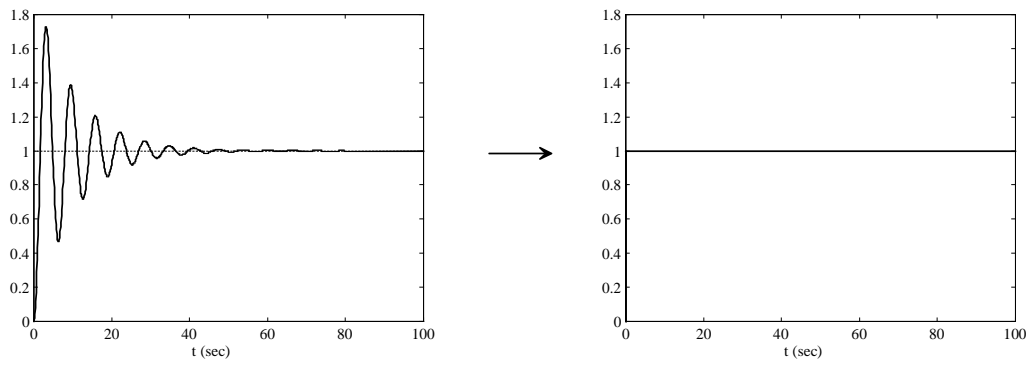


Fig. 101 : Equivalenza tra generici segnali e segnali canonici ai fini della valutazione delle prestazioni statiche

Essendo il sistema di controllo lineare, potremo valutare separatamente l'effetto sull'errore dovuto al segnale di riferimento ed ai disturbi (principio di sovrapposizione degli effetti).

Errore dovuto al segnale di riferimento

La funzione di trasferimento dal riferimento y° all'errore e è la seguente:

$$\frac{E(s)}{Y^\circ(s)} = \frac{1}{1+L(s)}.$$

Scriviamo la funzione di trasferimento d'anello nella seguente forma:

$$L(s) = \frac{\mu_L}{s^{g_L}} \frac{\prod_i (1+sT_i)}{\prod_k (1+s\tau_k)},$$

e calcoliamo il valore limite dell'errore utilizzando il teorema del valore finale (applicabile essendo il sistema asintoticamente stabile):

$$\begin{aligned} e_\infty &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} \left[s \frac{1}{1+L(s)} Y^\circ(s) \right] = \lim_{s \rightarrow 0} \left[s \frac{1}{1 + \frac{\mu_L}{s^{g_L}} \frac{\prod_i (1+sT_i)}{\prod_k (1+s\tau_k)}} Y^\circ(s) \right] = \\ &= \lim_{s \rightarrow 0} \left[s \frac{1}{1 + \frac{\mu_L}{s^{g_L}}} Y^\circ(s) \right] = \lim_{s \rightarrow 0} \left[\frac{s^{g_L+1}}{s^{g_L} + \mu_L} Y^\circ(s) \right] \end{aligned}$$

Se $y^\circ(t) = \text{Asca}(t)$, risulta:

$$e_\infty = \lim_{s \rightarrow 0} \left[\frac{s^{g_L+1}}{s^{g_L} + \mu_L} \frac{A}{s} \right] = \lim_{s \rightarrow 0} \left[A \frac{s^{g_L}}{s^{g_L} + \mu_L} \right] = \begin{cases} A, & g_L < 0 \\ \frac{A}{1+\mu_L}, & g_L = 0 \\ 0, & g_L \geq 1 \end{cases}$$

Se $y^\circ(t) = \text{Aram}(t)$, risulta:

$$e_\infty = \lim_{s \rightarrow 0} \left[\frac{s^{g_L+1}}{s^{g_L} + \mu_L} \frac{A}{s^2} \right] = \lim_{s \rightarrow 0} \left[A \frac{s^{g_L}}{s^{g_L+1} + s\mu_L} \right] = \begin{cases} \infty, & g_L \leq 0 \\ \frac{A}{\mu_L}, & g_L = 1 \\ 0, & g_L \geq 2 \end{cases}$$

Se $y^\circ(t) = \text{Apar}(t)$, risulta:

$$e_\infty = \lim_{s \rightarrow 0} \left[\frac{s^{g_L+1}}{s^{g_L} + \mu_L} \frac{A}{s^3} \right] = \lim_{s \rightarrow 0} \left[A \frac{s^{g_L}}{s^{g_L+2} + s^2\mu_L} \right] = \begin{cases} \infty, & g_L \leq 1 \\ \frac{A}{\mu_L}, & g_L = 2 \\ 0, & g_L \geq 3 \end{cases}$$

Pertanto, per valori negativi del tipo g_L della funzione di trasferimento d'anello, l'errore è sempre infinito, o, tutt'al più, nel caso dell'ingresso a scalino, pari all'ampiezza stessa dello scalino in ingresso: si tratta di situazioni di nessun interesse pratico. Per valori del tipo

maggiori o uguali a zero, si può compilare la seguente tabella:

g_L	$Asca(t)$	$Aram(t)$	$Apar(t)$
0	$\frac{A}{1+\mu_L}$	∞	∞
1	0	$\frac{A}{\mu_L}$	∞
2	0	0	$\frac{A}{\mu_L}$

Si osservi che, quando l'errore assume un valore finito e non nullo, esso è tanto più piccolo quanto maggiore è il valore del guadagno d'anello μ_L .

Esempio

Sia:

$$L(s) = \frac{10}{s} \frac{1+s}{1+10s}.$$

Il sistema in anello chiuso è asintoticamente stabile, come si ricava immediatamente dall'analisi del polinomio caratteristico in anello chiuso.

Poiché il tipo di L vale $g_L = 1$, si ha errore a transitorio esaurito nullo con riferimento a scalino, infinito con riferimento a parabola, mentre con ingresso a rampa, l'errore a regime è pari all'ampiezza della rampa diviso 10.

Errore dovuto al disturbo in linea di andata

Dallo schema di Fig.100 si ottiene la funzione di trasferimento dal disturbo d all'errore e :

$$\frac{E(s)}{D(s)} = -\frac{1}{1+L(s)}.$$

A parte il segno, si tratta della stessa funzione di trasferimento presente tra il riferimento e l'errore. Pertanto tutti i risultati della discussione precedente possono ancora essere utilizzati, pur di tenere conto del cambiamento di segno.

Si possono tuttavia presentare dei casi in cui il disturbo non entra nello schema a blocchi del sistema di controllo come raffigurato in Fig. 2, ossia direttamente in uscita alla funzione di trasferimento del processo. Per poter utilizzare ancora la tabella delle prestazioni statiche, occorre allora "riportare" il disturbo in uscita, considerando uno schema analogo a quello di Fig.2, in cui il disturbo in uscita è tale da dare gli stessi effetti a transitorio esaurito del disturbo effettivo.

Si considerino i seguenti due casi:

- a) Il disturbo entra nel sistema di controllo passando attraverso un sistema di funzione di trasferimento $H(s)$:

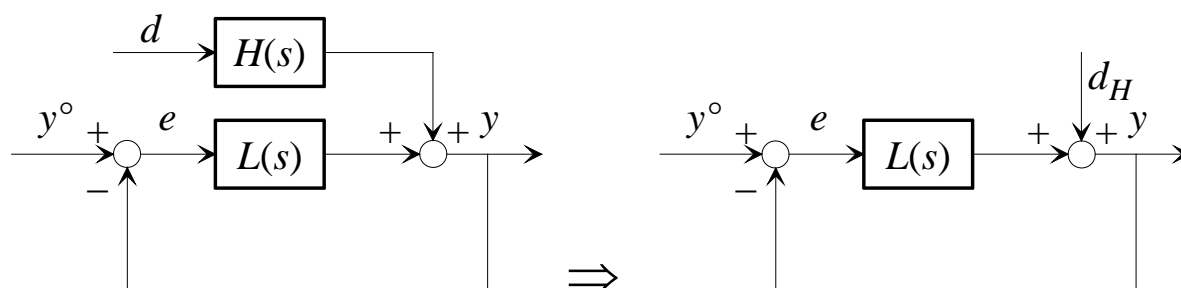


Fig. 102 : Sistema di controllo con disturbo filtrato

Detti μ_H e g_H guadagno e tipo di $H(s)$, il disturbo riportato in uscita, d_H , equivalente agli effetti statici al disturbo effettivo d , avrà trasformata:

$$D_H(s) = \frac{\mu_H}{s^{g_H}} D(s).$$

Si osservi infatti che gli eventuali poli o zeri di H non nell'origine non hanno alcun effetto sul comportamento a regime ($s \rightarrow 0$).

- b) Il disturbo entra nel sistema di controllo a monte del processo, ossia del sistema di funzione di trasferimento $G(s)$ (disturbo di *carico*):

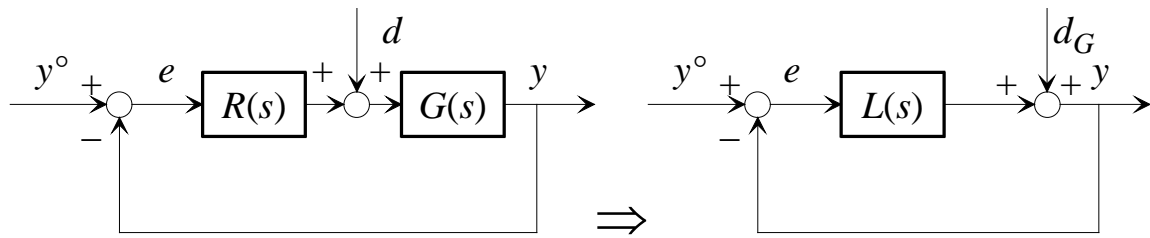


Fig. 103 : Sistema di controllo con disturbo di carico

Detti μ_G e g_G guadagno e tipo di $G(s)$, il disturbo riportato in uscita, d_G , equivalente agli effetti statici al disturbo effettivo d , avrà trasformata:

$$D_G(s) = \frac{\mu_G}{s^{g_G}} D(s).$$

Esempio

Con riferimento alla Fig. 103, sia:

$$R(s) = 5, \quad G(s) = \frac{6}{s} \frac{1+3s}{1+4s}, \quad d(t) = 3\text{sca}(t).$$

La funzione di trasferimento d'anello risulta:

$$L(s) = \frac{30}{s} \frac{1+3s}{1+4s}.$$

Il polinomio caratteristico è:

$$\chi(s) = 30(1+3s) + s(1+4s) = 4s^2 + 91s + 30,$$

ed ha le due radici a parte reale negativa, il che comporta che il sistema in anello chiuso è asintoticamente stabile.

Il disturbo d_G riportato in uscita ha trasformata:

$$D_G(s) = \frac{6}{s} D(s) = \frac{6}{s} \frac{3}{s} = \frac{18}{s^2}.$$

Pertanto:

$$d_G(t) = 18\text{ram}(t).$$

Poiché il tipo della funzione di trasferimento d'anello vale $g_L = 1$, ed il guadagno $\mu_L = 30$, dalla tabella si ottiene:

$$e_\infty = -\frac{18}{30} = -0.6.$$

Errore dovuto al disturbo in linea di retroazione

Dallo schema di Fig.100 si ottiene la funzione di trasferimento dal disturbo n all'errore e :

$$\frac{E(s)}{N(s)} = \frac{L(s)}{1 + L(s)} = F(s) .$$

Calcoliamo il valore limite dell'errore utilizzando il teorema del valore finale:

$$e_{\infty} = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} \left[s \frac{L(s)}{1 + L(s)} N(s) \right] = \lim_{s \rightarrow 0} \left[s \frac{\frac{\mu_L}{s^{g_L}}}{1 + \frac{\mu_L}{s^{g_L}}} N(s) \right] = \lim_{s \rightarrow 0} \left[\frac{\mu_L s}{s^{g_L} + \mu_L} N(s) \right]$$

Se $n(t) = \text{Asca}(t)$, risulta:

$$e_{\infty} = \lim_{s \rightarrow 0} \left[\frac{\mu_L s}{s^{g_L} + \mu_L} \frac{A}{s} \right] = \lim_{s \rightarrow 0} \left[A \frac{\mu_L}{s^{g_L} + \mu_L} \right] = \begin{cases} A \frac{\mu_L}{1 + \mu_L}, & g_L = 0 \\ A, & g_L \geq 1 \end{cases}$$

Se $n(t) = \text{Aram}(t)$, risulta:

$$e_{\infty} = \lim_{s \rightarrow 0} \left[\frac{\mu_L s}{s^{g_L} + \mu_L} \frac{A}{s^2} \right] = \lim_{s \rightarrow 0} \left[A \frac{\mu_L}{s^{g_L+1} + \mu_L s} \right] = \infty, \quad \forall g_L \geq 0 .$$

Se $n(t) = \text{Apar}(t)$, risulta:

$$e_{\infty} = \lim_{s \rightarrow 0} \left[\frac{\mu_L s}{s^{g_L} + \mu_L} \frac{A}{s^3} \right] = \lim_{s \rightarrow 0} \left[A \frac{\mu_L}{s^{g_L+2} + \mu_L s^2} \right] = \infty, \quad \forall g_L \geq 0 .$$

Si può compilare la seguente tabella:

g_L	$\text{Asca}(t)$	$\text{Aram}(t)$	$\text{Apar}(t)$
0	$A \frac{\mu_L}{1 + \mu_L}$	∞	∞
1	A	∞	∞
2	A	∞	∞

Quindi l'errore si mantiene finito solo per disturbo a scalino dove però è pari all'ampiezza del disturbo stesso per tipo maggiore o uguale a 1, e se ne scosta solo leggermente per tipo uguale a zero (si ricorda che μ_L deve essere un numero elevato per garantire errore piccolo sul riferimento e sul disturbo in linea di andata).

E' allora evidente che, in presenza di un trasduttore con errore statico, il sistema di controllo non può garantire a regime una precisione migliore di quella del trasduttore.

Capitolo 12: Progetto del controllore

Consideriamo nuovamente un sistema di controllo in retroazione:

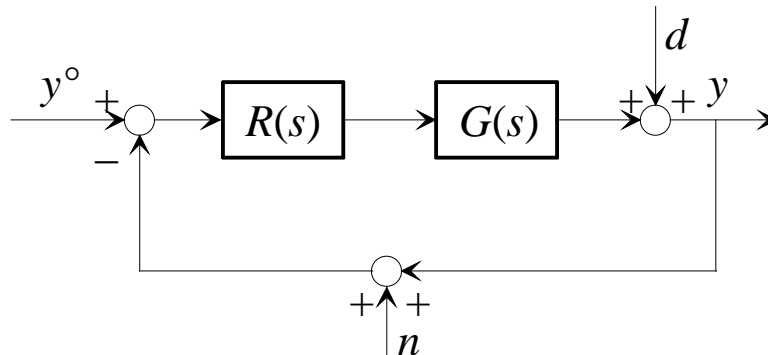


Fig. 104 : Sistema di controllo

Supporremo assegnata la funzione di trasferimento $G(s)$, e ci porremo l'obiettivo della **sintesi** (o **progetto**) della funzione di trasferimento $R(s)$ del controllore, sulla base di un certo numero di specifiche.

Il metodo che seguiremo⁸ sarà basato sul criterio di Bode, e come tale prevederà il rispetto, da parte della funzione di trasferimento d'anello $L(s) = R(s)G(s)$, delle ipotesi necessarie per l'applicabilità del criterio. Ciò comporta, in particolare, che il metodo non sarà applicabile se $G(s)$ ha poli a parte reale positiva. Infatti tali poli, che non possono essere cancellati da corrispondenti zeri nella funzione di trasferimento $R(s)$, si presentano anche in $L(s)$, violando una condizione del criterio.

Le specifiche con le quali tipicamente si imposta il progetto sono le seguenti:

1. Asintotica stabilità

In base al criterio di Bode questa specifica implica:

$$\varphi_m > 0.$$

2. Grado di stabilità e robustezza

La stabilità deve essere garantita con un certo margine rispetto ad inesattezze nella modellistica del sistema sotto controllo. Inoltre si vuole che il sistema di controllo, sollecitato da ingressi canonici (scalino, impulso), esibisca risposte ben smorzate. Entrambe queste condizioni sono approssimativamente soddisfatte se il margine di fase dell'anello è superiore di un valore limite, di volta in volta specificato:

$$\varphi_m > \overline{\varphi_m}.$$

⁸ Per una trattazione analoga a quella qui presentata ma di maggiore respiro, si può consultare il Capitolo 2 del testo: "Esercizi di Controlli Automatici", G.O. Guardabassi e P. Rocco, Pitagora Editrice.

3. Velocità di risposta

Il sistema di controllo deve reagire prontamente a variazioni nel segnale di riferimento o sui disturbi. Sappiamo che, se il margine di fase è sufficientemente elevato, i transitori del sistema in anello chiuso sono governati da una costante di tempo pari all'inverso della pulsazione critica ω_c . Imporre un valore massimo alla durata dei transitori equivale quindi ad imporre un valore minimo alla pulsazione critica:

$$\omega_c \geq \overline{\omega}_c.$$

4. Precisione statica

A seguito dell'imposizione di assegnati segnali canonici agli ingressi, l'errore tra riferimento e variabile controllata deve essere, a transitorio esaurito, nullo o inferiore in modulo ad una assegnata soglia:

$$|e_\infty| \leq \overline{e}_\infty.$$

5. Specifiche aggiuntive

Talvolta possono essere date delle ulteriori specifiche, come l'uso di un controllore dalla struttura assegnata (cioè con un dato numero di poli e zeri), limitazioni all'azione di controllo, attenuazione di specifici disturbi (tipicamente sinusoidali).

Impostazione del progetto

Il progetto viene di norma suddiviso in due fasi:

1. Progetto statico

E' la fase in cui ci si preoccupa di soddisfare la specifica relativa alle prestazioni statiche (errore a transitorio esaurito). Si affronta con l'ipotesi preliminare di essere in grado di rendere il sistema in anello chiuso asintoticamente stabile nella successiva fase di progetto dinamico.

Si fattorizza la funzione di trasferimento del regolatore come segue:

$$R(s) = R_1(s)R_2(s),$$

con:

$$R_1(s) = \frac{\mu_R}{s^{g_R}}, \quad R_2(s) = \frac{\prod_i (1 + sT_i)}{\prod_k (1 + s\tau_k)}.$$

Dal momento che $R_2(0) = 1$, le prestazioni statiche sono unicamente determinate dalle scelte effettuate sui parametri di $R_1(s)$. Per non complicare la successiva fase di progetto, si sceglie:

- il valore minimo del tipo g_R che consente di soddisfare la specifica statica;
- fissato g_R , il valore minimo del guadagno μ_R che consente di soddisfare la specifica.

Può succedere, a valle della scelta del tipo del regolatore, che il valore del guadagno risulti indeterminato: in questo caso è possibile assegnare il guadagno in fase di progetto dinamico.

2. Progetto dinamico

Si determina $R_2(s)$ (poli e zeri del controllore), in modo tale da soddisfare le specifiche dinamiche su margine di fase e velocità di risposta. Si procede per mezzo di una serie, ragionata, di tentativi di sintesi.

Esempio introduttivo

Si consideri il sistema di controllo di figura:

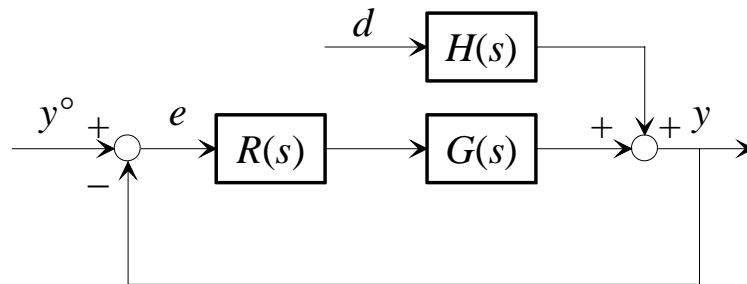


Fig. 105 : Sistema di controllo per l'esempio

dove:

$$G(s) = \frac{50}{(1+0.1s)(1+s)(1+10s)}, \quad H(s) = \frac{5}{1+0.01s}.$$

Si vuole progettare il regolatore $R(s)$ in modo tale che:

- $|e_\infty| \leq 0.025$ per $y^o(t) = 10 \text{ sca}(t)$, $d(t) = \pm \text{sca}(t)$;
- $\omega_c \geq 1 \text{ rad/s}$;
- $\phi_m \geq 60^\circ$.

Progetto statico

Fattorizzata $R(s)$ come:

$$R(s) = R_1(s)R_2(s), \quad R_2(0) = 1,$$

si deve progettare $R_1(s) = \mu_R / s^{g_R}$ in modo da soddisfare la specifica sull'errore a transitorio esaurito. Ipotizzando il sistema in anello chiuso asintoticamente stabile e facendo uso, in virtù della linearità del sistema, del principio di sovrapposizione degli effetti, calcoliamo l'errore dovuto al segnale di riferimento:

$$\begin{aligned} e_\infty^o &= \lim_{s \rightarrow 0} [sE^o(s)] = \lim_{s \rightarrow 0} \left[s \frac{1}{1+L(s)} Y^o(s) \right] = \lim_{s \rightarrow 0} \left[s \frac{1}{1 + \frac{50\mu_R}{s^{g_R}}} \frac{10}{s} \right] = \\ &= \lim_{s \rightarrow 0} \frac{10s^{g_R}}{s^{g_R} + 50\mu_R} = \begin{cases} \frac{10}{1+50\mu_R}, & g_R = 0 \\ 0, & g_R \geq 1 \end{cases} \end{aligned}$$

Il risultato è coerente con le tabelle della precisione statica, pur di tenere conto che il guadagno d'anello vale $50\mu_R$ e che lo scalino ha ampiezza 10.

Passando all'errore dovuto al disturbo, si ha:

$$e_{d\infty} = \lim_{s \rightarrow 0} [sE_d(s)] = \lim_{s \rightarrow 0} \left[s \frac{-H(s)}{1+L(s)} D(s) \right] = \lim_{s \rightarrow 0} \left[s \frac{-5}{1 + \frac{50\mu_R}{s^{g_R}}} \left(\pm \frac{1}{s} \right) \right] =$$

$$= \mp \lim_{s \rightarrow 0} \frac{5s^{g_R}}{s^{g_R} + 50\mu_R} = \begin{cases} \mp \frac{5}{1+50\mu_R}, & g_R = 0 \\ 0, & g_R \geq 1 \end{cases}$$

Anche in questo caso il risultato è coerente con le tabelle, tenendo conto che il disturbo va riportato in uscita (e quindi, agli effetti statici, va moltiplicato per il guadagno di $H(s)$).

Appare evidente che per entrambi i segnali di ingresso (riferimento e disturbo), la specifica sull'errore può essere garantita con regolatore di tipo nullo ($g_R = 0$). Adottando questo valore ed assumendo il guadagno μ_R positivo, si ha:

$$|e_{\infty}| = |e_{\infty}^o + e_{d\infty}| \leq |e_{\infty}^o| + |e_{d\infty}| = \frac{10}{1+50\mu_R} + \frac{5}{1+50\mu_R} = \frac{15}{1+50\mu_R}.$$

La specifica statica sarà quindi senz'altro soddisfatta se:

$$\frac{15}{1+50\mu_R} \leq 0.025 \Rightarrow \mu_R \geq \frac{15-0.025}{1.25} \approx 12.$$

Rispetto al valore limite calcolato dalla precedente disequaglianza, conviene predisporre un opportuno margine di sicurezza, inteso a cautelarsi verso le inevitabili incertezze sui parametri del processo e sul valore dei disturbi. Un valore opportuno potrebbe allora essere $\mu_R = 20$.

Il progetto statico è allora concluso con la scelta della funzione di trasferimento:

$$R_1(s) = 20.$$

Progetto dinamico

La funzione di trasferimento d'anello del sistema può essere scritta come:

$$L(s) = R_1(s)R_2(s)G(s) = R_2(s)L_1(s),$$

con:

$$L_1(s) = R_1(s)G(s) = \frac{1000}{(1+0.1s)(1+s)(1+10s)}.$$

Il primo tentativo da eseguire per il progetto di $R_2(s)$ consiste nel porre banalmente $R_2(s) = 1$. La funzione di trasferimento d'anello coincide, allora, con $L_1(s)$, ed il relativo diagramma di Bode del modulo asintotico è tracciato di seguito:

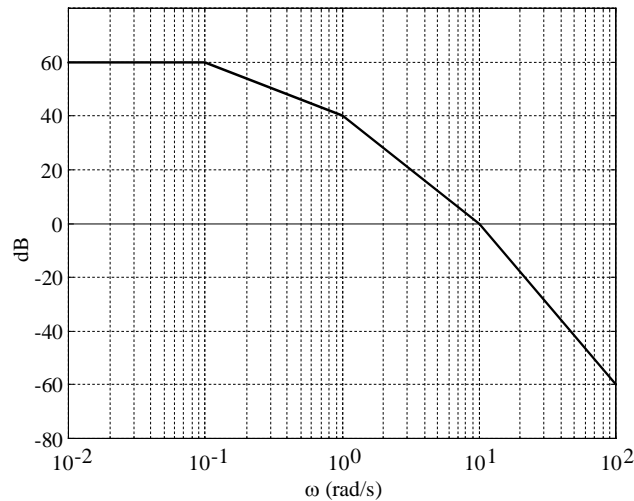


Fig. 106 : Diagramma di Bode asintotico di $|L_1|$

Mentre la pulsazione critica è abbondantemente superiore al valore limite imposto dalla specifica, il margine di fase risulta, come facilmente calcolabile, negativo, per cui il sistema in anello chiuso risulterebbe addirittura instabile.

Per il progetto di $R_2(s)$ si può ricorrere all'osservazione fatta nella Lezione 8 a proposito del criterio di Bode, riguardo ai sistemi a **fase minima**. Si ricorda che, per un sistema a fase minima, il fatto che l'attraversamento dell'asse a 0 dB da parte del diagramma asintotico del modulo della risposta in frequenza di L avvenga con pendenza -1 (-20 dB/decade) garantisce di norma un margine di fase positivo, tanto più prossimo ai 90° , quanto più ampio è il tratto di pendenza -1 .

Conviene allora procedere determinando preliminarmente un opportuno andamento per il diagramma del modulo di L e, a posteriori, risalire all'espressione della funzione di trasferimento del regolatore.

Il **metodo di progetto**, di natura grafica, consiste nell'individuare sull'asse a 0 dB un valore di pulsazione, superiore al limite inferiore richiesto per la pulsazione critica, e per questo punto tracciare un tratto di retta a pendenza -1 , destinato ad essere un tratto del diagramma di $|L|$.

In **bassa frequenza** si può operare come segue:

- il diagramma di $|L|$ deve avere la stessa pendenza di quello di $|L_1|$, altrimenti si modificherebbe il tipo del regolatore in sede di progetto dinamico;
- se il progetto statico si è concluso con un vincolo sul valore del guadagno μ_R , il valore di $|L|$ deve essere maggiore o uguale (di fatto è comodo sceglierlo uguale) a quello di $|L_1|$.

In **alta frequenza**, invece:

- il diagramma di $|L|$ deve avere pendenza maggiore o uguale in modulo (di fatto è comodo sceglierla uguale) a quella di $|L_1|$, altrimenti si perverrebbe al progetto di un regolatore non realizzabile (con più zeri che poli)⁹;

⁹ In realtà il vincolo sulla pendenza assicura la realizzabilità della funzione di trasferimento $R_2(s)$, laddove ciò che interessa è la realizzabilità della funzione di trasferimento $R(s)$ complessiva. In presenza quindi di una funzione di trasferimento $R_1(s)$ di tipo non nullo, il vincolo sulla pendenza potrebbe essere rilassato.

- il valore di $|L|$ deve essere minore o uguale a quello di $|L_1|$, per garantire la moderazione del controllo.

L'applicazione del metodo all'esempio porta, scegliendo come pulsazione critica il valore 2 rad/s, al grafico di $|L|$ di figura 4.

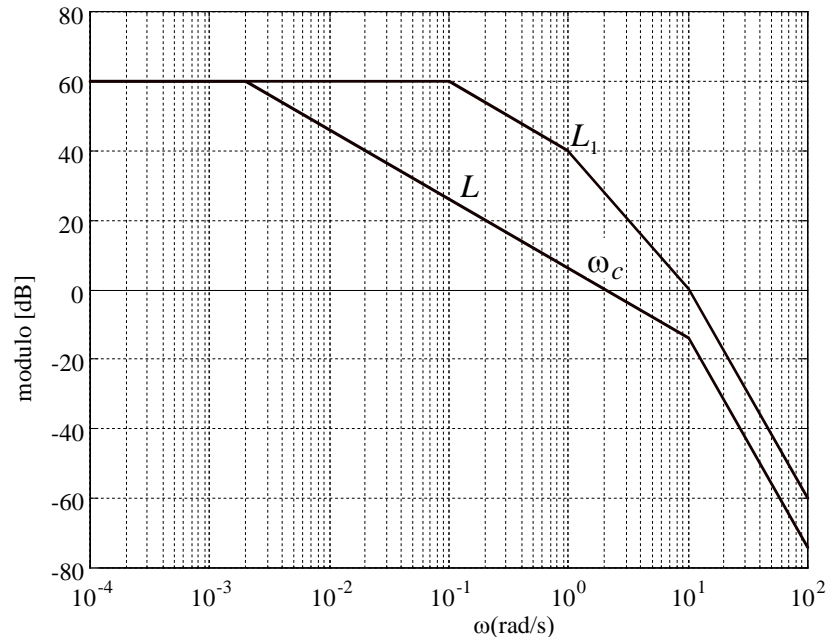


Fig. 107 : Diagramma di Bode asintotico di $|L|$

Il margine di fase risulta:

$$\varphi_m = 180^\circ - \left| -\arctan(2/0.002) - 2 \arctan(2/10) \right| = 180^\circ - \left| -90^\circ - 2 \times 11^\circ \right| = 68^\circ.$$

Le specifiche dinamiche sono quindi soddisfatte, e risulta:

$$L(s) = \frac{1000}{\left(1 + \frac{s}{0.002}\right) \left(1 + \frac{s}{10}\right)^2} = \frac{1000}{(1 + 500s)(1 + 0.1s)^2}$$

Ne consegue:

$$R_2(s) = \frac{L(s)}{L_1(s)} = \frac{1000}{(1 + 500s)(1 + 0.1s)^2} \frac{(1 + 0.1s)(1 + s)(1 + 10s)}{1000} = \frac{(1 + s)(1 + 10s)}{(1 + 500s)(1 + 0.1s)}.$$

Infine:

$$R(s) = R_1(s)R_2(s) = 20 \frac{(1 + s)(1 + 10s)}{(1 + 500s)(1 + 0.1s)}.$$

Si osservi che in alta frequenza il diagramma di $|L|$ non si congiunge con il diagramma di $|L_1|$. Sarebbe stato lecito congiungere i due diagrammi, ma in questo caso si sarebbe ottenuto, come è facile verificare, un controllore di ordine 3 (con 3 poli), più problematico da realizzare.

Esempio di progetto per sistemi di tipo 1

Per il sistema di controllo di figura:

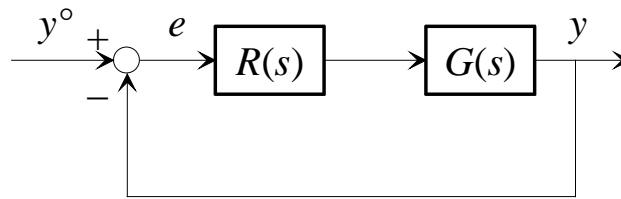


Fig. 108 : Sistema di controllo per l'esempio

dove:

$$G(s) = \frac{10}{s(1+s)},$$

si vuole progettare il regolatore $R(s)$ in modo tale che:

- $e_\infty = 0$ per $y^o(t) = \text{sca}(t)$;
- $\omega_c \geq 2 \text{ rad/s}$;
- $\phi_m \geq 40^\circ$.

Progetto statico

Procedendo come nell'esempio precedente, si ha:

$$e_\infty = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} \left[s \frac{1}{1 + L(s)} Y^o(s) \right] = \lim_{s \rightarrow 0} \left[s \frac{1}{1 + \frac{10 \mu_R}{s} \frac{1}{s^{g_R}}} \right] = \lim_{s \rightarrow 0} \frac{s^{g_R+1}}{s^{g_R+1} + 10\mu_R} = 0, \quad \forall g_R \geq 0$$

In questo caso, anche con regolatore di tipo nullo ($g_R = 0$) si ottiene errore statico nullo. Il tipo della funzione di trasferimento d'anello è infatti la somma del tipo del sistema sotto controllo e del tipo del controllore:

$$g_L = g_G + g_R.$$

Poiché nel presente caso $g_G = 1$, è sufficiente imporre $g_R = 0$ per avere un anello di tipo 1.

Il progetto statico quindi non impone nessun vincolo al regolatore e potremo porre, formalmente:

$$R_1(s) = 1.$$

Progetto dinamico

Poiché:

$$L(s) = R_1(s)R_2(s)G(s) = R_2(s)G(s),$$

conviene tracciare il diagramma del modulo di G per controllare se le specifiche sono già soddisfatte con $R_2 = 1$:

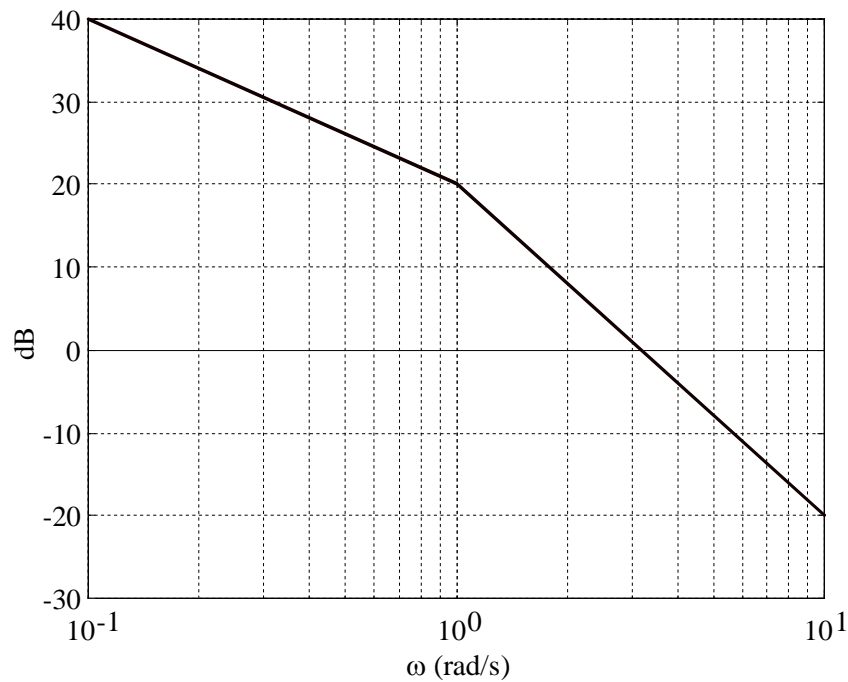


Fig. 109 : Diagramma di Bode asintotico di $|G|$

Si ha $\omega_c = 3$ rad/s, ma:

$$\varphi_m = 180^\circ - |-90^\circ - \arctan(3)| = 180^\circ - |-90^\circ - 71^\circ| = 19^\circ,$$

che non soddisfa la specifica.

La soluzione grafica del progetto dinamico è riportata in Fig. 7: si osservi che in bassa frequenza si è evitato di raccordare i diagrammi del modulo di L e G , in quanto il progetto statico non impone alcun vincolo sul guadagno del regolatore.

Si ottiene $\omega_c = 2$ rad/s, e:

$$\varphi_m = 180^\circ - |-90^\circ - \arctan(2/5)| = 180^\circ - |-90^\circ - 22^\circ| = 68^\circ.$$

Tutte le specifiche sono soddisfatte e risulta:

$$L(s) = \frac{2}{s} \frac{1}{1+s/5} = \frac{2}{s} \frac{1}{1+0.2s},$$

da cui:

$$R(s) = R_2(s) = \frac{L(s)}{G(s)} = \frac{2}{s} \frac{1}{1+0.2s} \frac{s(s+1)}{10} = 0.2 \frac{1+s}{1+0.2s}.$$

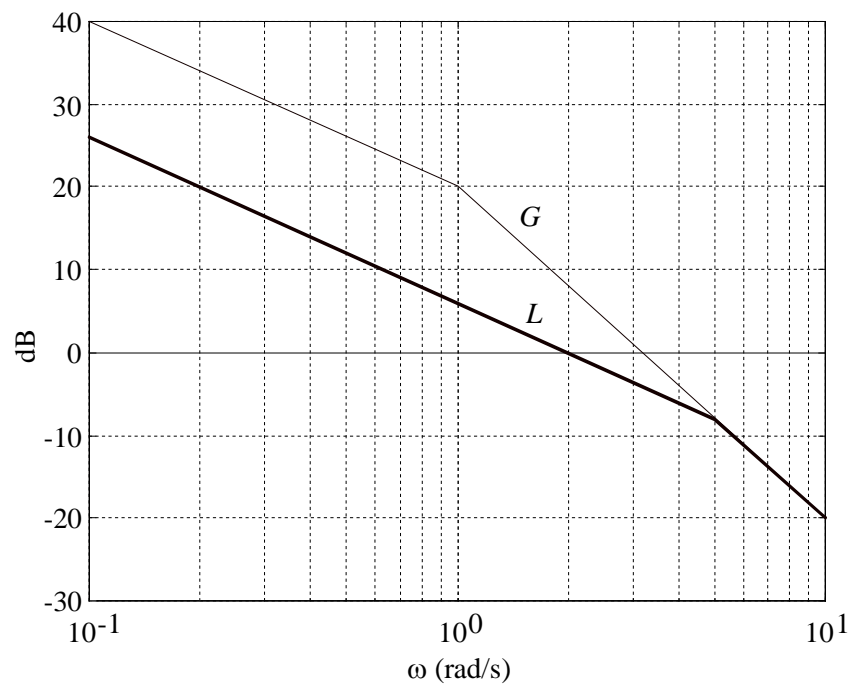


Fig. 110 : Diagrammi di Bode asintotici di $|G|$ e $|L|$

Esempio di progetto per sistemi con zeri a parte reale positiva

Per il sistema di controllo di figura:

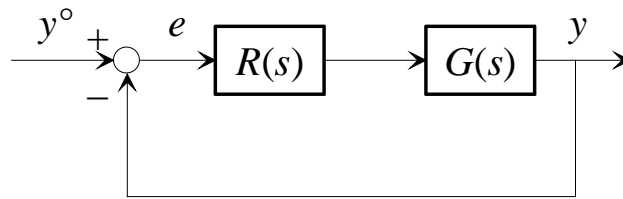


Fig. 111 : Sistema di controllo per l'esempio

dove:

$$G(s) = 10 \frac{1-2s}{1+10s},$$

si vuole progettare il regolatore $R(s)$ in modo tale che:

- $e_{\infty} = 0$ per $y^{\circ}(t) = \text{sca}(t)$;
- $\omega_c \geq 0.1 \text{ rad/s}$;
- $\varphi_m \geq 40^{\circ}$.

Il progetto statico impone la presenza di un integratore nel controllore, mentre non ne vincola in alcun modo il guadagno. Possiamo allora porre:

$$R_1(s) = \frac{1}{s}.$$

Scritta la funzione di trasferimento d'anello come:

$$L(s) = R_2(s)L_1(s),$$

con:

$$L_1(s) = R_1(s)G(s) = \frac{10}{s} \frac{1-2s}{1+10s},$$

tracciamo il diagramma di Bode del modulo di L_1 :

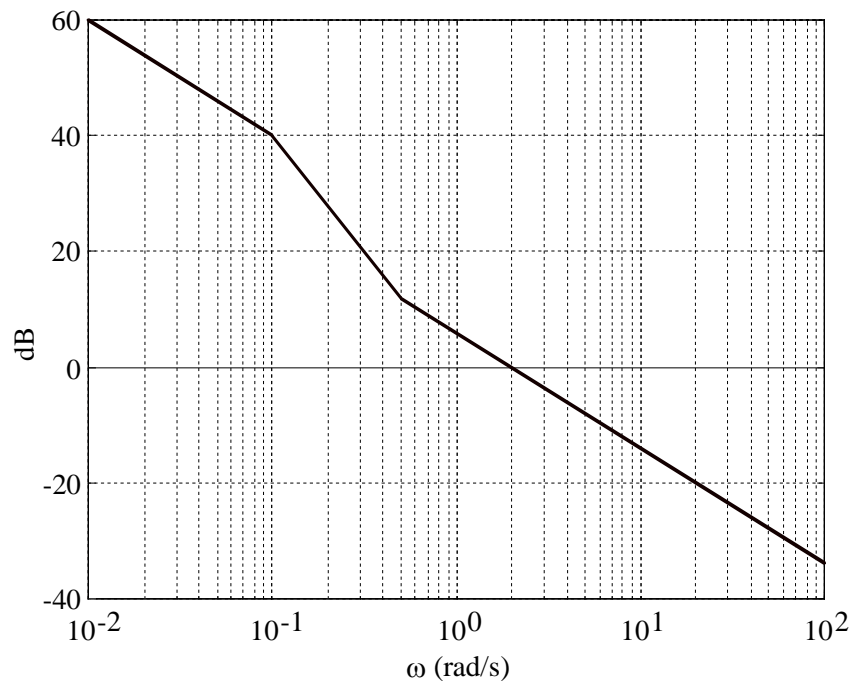


Fig. 112 : Diagramma di Bode asintotico di $|L_1|$

Risulta $\omega_c = 2 \text{ rad/s}$, e:

$$\varphi_m = 180^\circ - \left| -90^\circ - \arctan(2/0.1) - \arctan(2/0.5) \right| = 180^\circ - \left| -90^\circ - 87^\circ - 76^\circ \right| = -73^\circ .$$

Il sistema in anello chiuso sarebbe quindi instabile.

Si osservi che il contributo di fase dello zero alla pulsazione 0.5 rad/s è negativo, essendo lo zero a parte reale positiva. Il sistema non è quindi “a fase minima”, ed il fatto che il diagramma del modulo tagli l’asse a 0 dB con pendenza -1 non implica margine di fase positivo. D’altra parte lo zero non può essere cancellato da un corrispondente polo nel regolatore, pena il manifestarsi di una instabilità interna. L’unico provvedimento utile in questo caso è fare in modo che la pulsazione critica risulti sensibilmente inferiore alla pulsazione dello zero, in modo che il contributo negativo alla fase critica sia meno rilevante.

Si consideri ad esempio la soluzione riportata in figura:

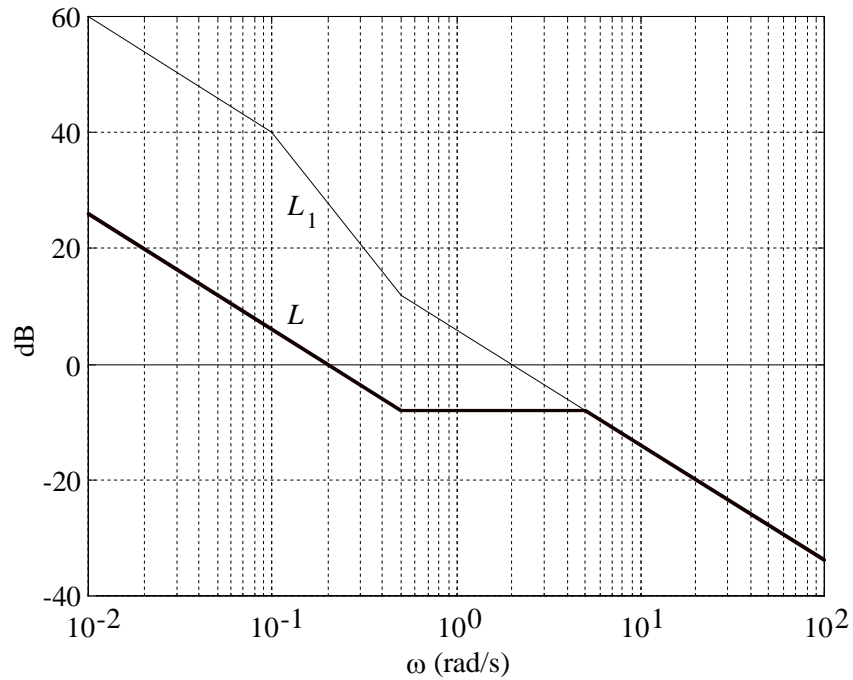


Fig. 113 : Diagrammi di Bode asintotici di $|L_1|$ e $|L|$

Risulta $\omega_c = 0.2$ rad/s, e:

$$\varphi_m = 180^\circ - \left| -90^\circ - \arctan(0.2/0.5) - \arctan(0.2/5) \right| = 180^\circ - \left| -90^\circ - 22^\circ - 2^\circ \right| = 66^\circ .$$

Anche in questo esempio, come nel precedente, i diagrammi non sono stati raccordati in bassa frequenza, non essendoci un vincolo di guadagno dato dal progetto statico.

Risulta quindi:

$$L(s) = \frac{0.2}{s} \frac{1-2s}{1+s/5} = \frac{0.2}{s} \frac{1-2s}{1+0.2s} ,$$

da cui:

$$R_2(s) = \frac{L(s)}{L_1(s)} = \frac{0.2}{s} \frac{1-2s}{1+0.2s} \frac{s}{10} \frac{1+10s}{1-2s} = 0.02 \frac{1+10s}{1+0.2s} ,$$

$$R(s) = R_1(s) R_2(s) = \frac{0.02}{s} \frac{1+10s}{1+0.2s} .$$

Esempio di progetto per sistemi con ritardo

Per il sistema di controllo di figura:

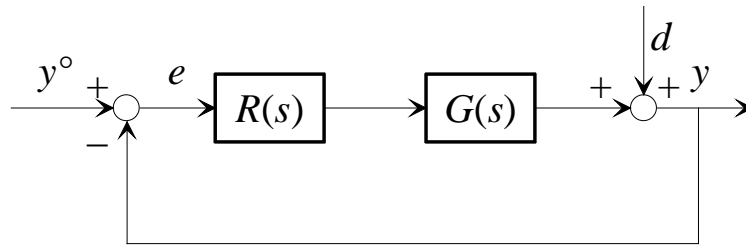


Fig. 114 : Sistema di controllo per l'esempio

dove:

$$G(s) = \frac{e^{-s}}{(1+10s)(1+s)} ,$$

si vuole progettare il regolatore $R(s)$ in modo tale che:

- $|e_{\infty}| \leq 0.15$ per $d(t) = \pm \text{sca}(t)$;
- $\omega_c \geq 0.3 \text{ rad / s}$;
- $\varphi_m \geq 40^\circ$.

Progetto statico

Supposto il sistema di controllo asintoticamente stabile, il ritardo non gioca alcun ruolo nel determinare le prestazioni statiche. Dalle tabelle della precisione statica otteniamo che per funzione di trasferimento d'anello di tipo 0 risulta:

$$|e_{\infty}| = \frac{1}{1 + \mu_L} .$$

Pertanto, scelto un regolatore di tipo 0, si dovrà imporre:

$$\frac{1}{1 + \mu_R} \leq 0.15 \Rightarrow \mu_R \geq 5.67 .$$

Possiamo cautelarci contro eventuali incertezze ponendo $\mu_R = 10$, e quindi concludere che:

$$R_1(s) = 10 .$$

Progetto dinamico

Scritta la funzione di trasferimento d'anello come:

$$L(s) = R_2(s)L_1(s) ,$$

con:

$$L_1(s) = R_1(s)G(s) = 10 \frac{e^{-s}}{(1+10s)(1+s)} ,$$

tracciamo il diagramma di Bode del modulo di L_1 :

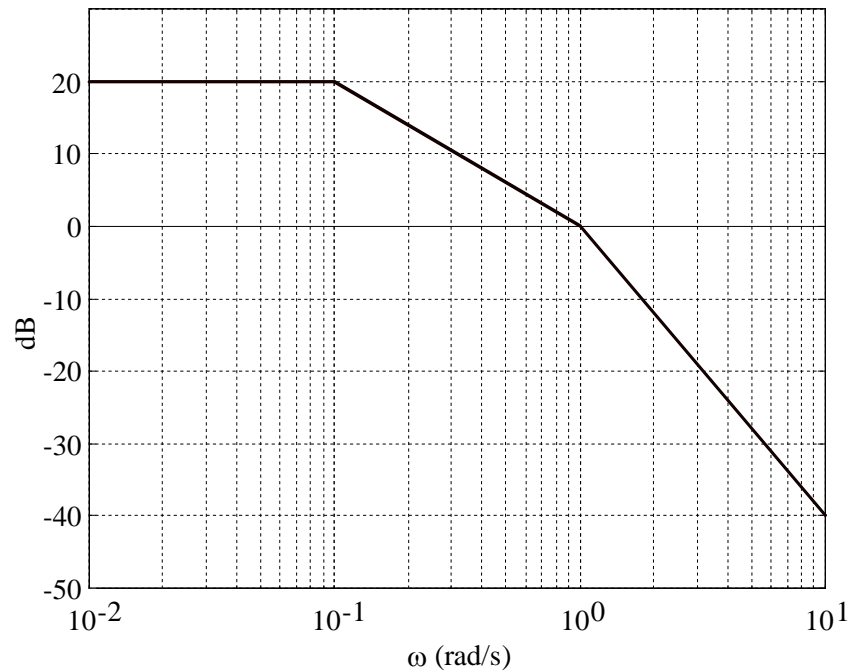


Fig. 115 : Diagramma di Bode asintotico di $|L_1|$

Risulta $\omega_c \approx 1 \text{ rad/s}$, e:

$$\varphi_m \approx 180^\circ - \left| -\arctan(1/0.1) - \arctan(1) - \omega_c \tau \frac{180^\circ}{\pi} \right| = 180^\circ - |-84^\circ - 45^\circ - 1 \times 1 \times 57^\circ| = -6^\circ .$$

A causa della presenza del ritardo il sistema risulterebbe instabile.

Per il progetto si può considerare la soluzione di Fig. 116:

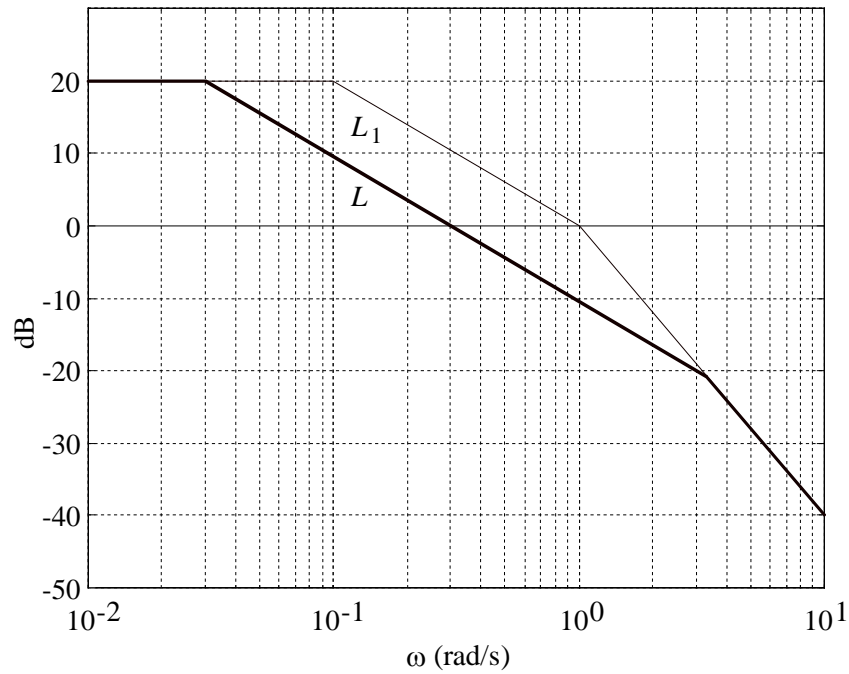


Fig. 116 : Diagrammi di Bode asintotici di $|L_1|$ e $|L|$

Risulta $\omega_c \approx 0.3 \text{ rad/s}$, e:

$$\varphi_m \approx 180^\circ - \left| -\arctan(0.3/0.03) - \arctan(0.3/3) - \omega_c \tau \frac{180^\circ}{\pi} \right| = 180^\circ - |-84^\circ - 6^\circ - 0.3 \times 1 \times 57^\circ| = 73^\circ .$$

Tutte le specifiche sono soddisfatte e risulta:

$$L(s) = 10 \frac{e^{-s}}{(1+s/0.03)(1+s/3)} = 10 \frac{e^{-s}}{(1+33s)(1+0.33s)} ,$$

da cui:

$$R_2(s) = \frac{L(s)}{L_1(s)} = \frac{(1+10s)(1+s)}{(1+33s)(1+0.33s)} ,$$

$$R(s) = R_1(s)R_2(s) = 10 \frac{(1+10s)(1+s)}{(1+33s)(1+0.33s)} .$$

Esempio di progetto per sistemi con disturbo sinusoidale

Per il sistema di controllo di figura:

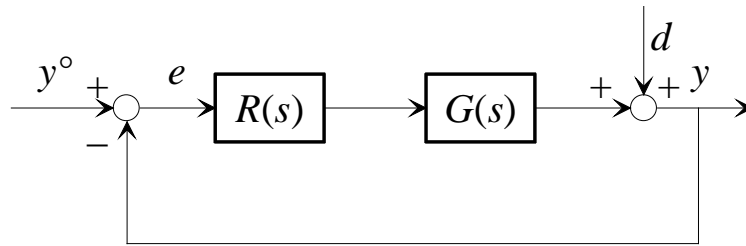


Fig. 117 : Sistema di controllo per l'esempio

dove:

$$G(s) = \frac{10}{s} \frac{1}{1+10s},$$

si vuole progettare il regolatore $R(s)$ in modo tale che:

- $e_{\infty} = 0$ per $y^{\circ}(t) = \text{sca}(t)$;
- un disturbo $d(t) = \sin(\overline{\omega}t)$, con $\overline{\omega} \leq 0.03$, sia attenuato a regime, sull'uscita y , di un fattore almeno pari a 10;
- $\omega_c \geq 0.1 \text{ rad / s}$;
- $\varphi_m \geq 50^\circ$.

Progetto statico

Essendo il sistema sotto controllo di tipo 1, la specifica sull'errore a transitorio esaurito è automaticamente soddisfatta. Poniamo formalmente $R_1(s) = 1$.

Per quanto riguarda invece la specifica sull'attenuazione del disturbo, valida anch'essa a transitorio esaurito, utilizzeremo il teorema della risposta in frequenza, nell'ipotesi di poter rendere il sistema in anello chiuso asintoticamente stabile.

La funzione di trasferimento dal disturbo d all'uscita y risulta:

$$\frac{Y(s)}{D(s)} = \frac{1}{1+L(s)} = S(s).$$

Pertanto, quando $d(t) = \sin(\overline{\omega}t)$, a transitorio esaurito si ha:

$$y(t) = |S(j\overline{\omega})| \sin(\overline{\omega}t + \angle S(j\overline{\omega})).$$

La specifica sull'attenuazione del disturbo si traduce quindi nella seguente condizione:

$$|S(j\overline{\omega})| < \frac{1}{10}, \quad \forall \overline{\omega} \leq 0.03.$$

D'altra parte sappiamo che:

$$|S(j\bar{\omega})| \approx \frac{1}{|L(j\bar{\omega})|},$$

essendo $\bar{\omega} \ll \omega_c$, $\forall \bar{\omega} \leq 0.03$, ossia $|L(j\bar{\omega})| \gg 1$.

Pertanto dovremo semplicemente imporre:

$$|L(j\bar{\omega})| > 10, \quad \forall \bar{\omega} \leq 0.03.$$

Progetto dinamico

Essendo $R_1(s) = 1$, risulta $L(s) = R_2(s) G(s)$. La Fig. 118 riporta i diagrammi di Bode del modulo di G e di L ottenuta secondo il solito criterio di progetto.

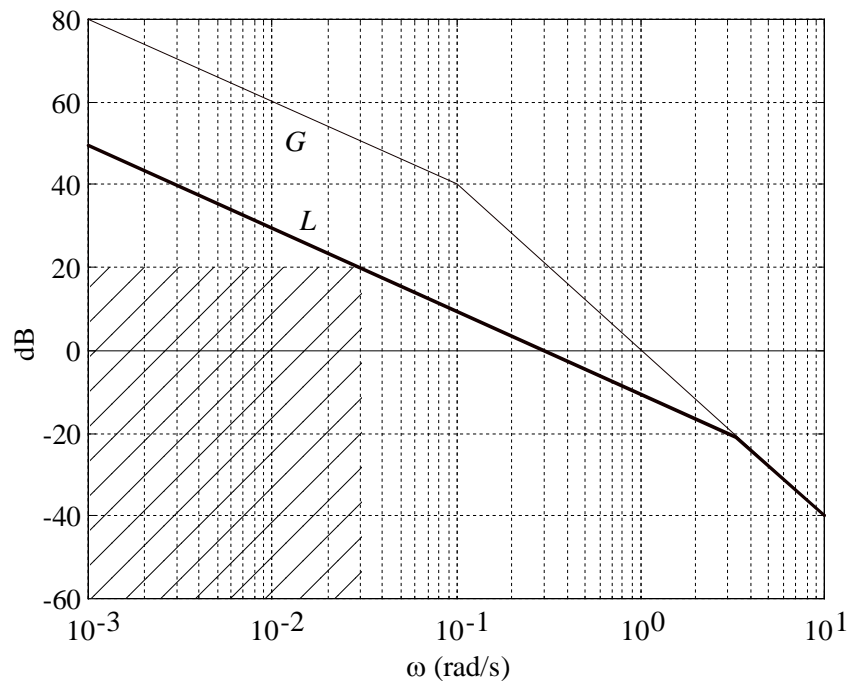


Fig. 118 : Diagrammi di Bode asintotici di $|G|$ e $|L|$

Si osservi che il vincolo sull'attenuazione del disturbo sinusoidale si traduce, graficamente, nell'individuazione di una "zona proibita", ossia di un'area nel piano del diagramma del modulo in cui il diagramma di $|L|$ non può entrare.

Risulta $\omega_c = 0.3$ rad/s, e:

$$\varphi_m \approx 180^\circ - |-90^\circ - \arctan(0.3/3)| = 180^\circ - |-90^\circ - 6^\circ| = 84^\circ.$$

Tutte le specifiche sono soddisfatte e risulta:

$$L(s) = \frac{0.3}{s} \frac{1}{1+0.33s}.$$

da cui:

$$R(s) = R_2(s) = \frac{L(s)}{G(s)} = \frac{0.3}{s} \frac{1}{1+0.33s} \frac{s(1+10s)}{10} = 0.03 \frac{1+10s}{1+0.33s}.$$

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:

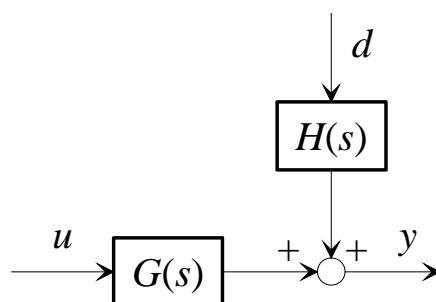


Fig. 119 : Sistema sotto controllo

Uno schema di compensazione potrebbe essere il seguente:

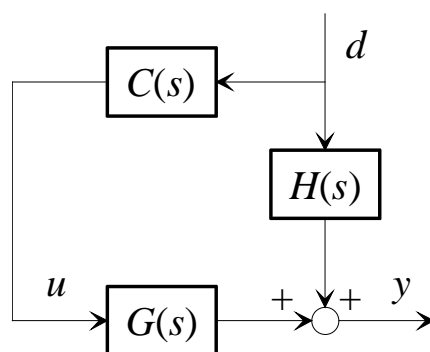


Fig. 120 : Compensazione del disturbo

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 ogniqualevolta $G(s)$ 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:

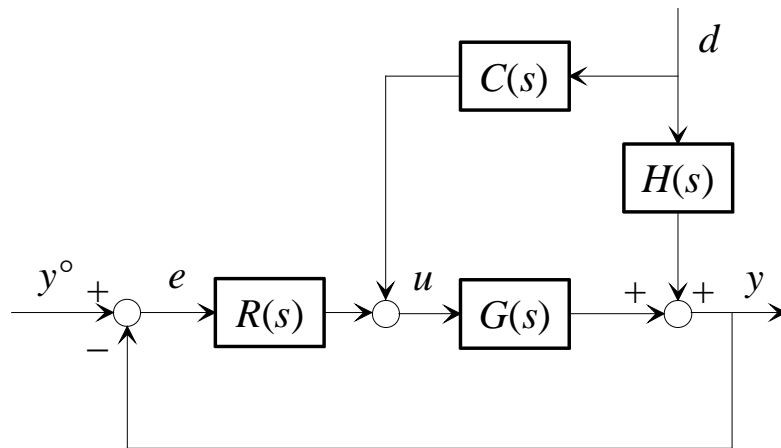


Fig. 121 : Controllo in retroazione con compensazione del disturbo

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**:

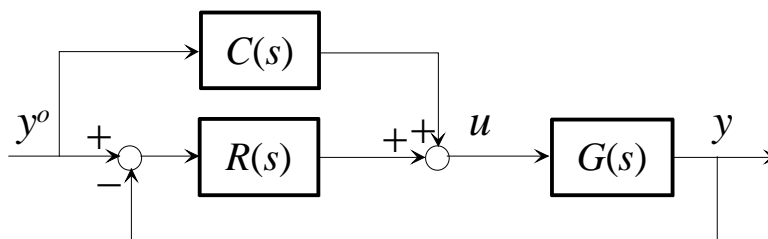


Fig. 122 : Compensazione del riferimento

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:

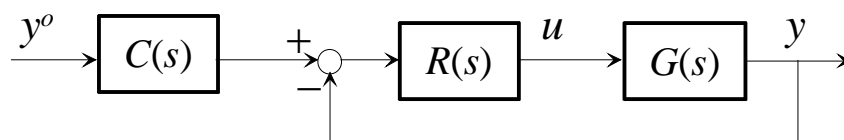


Fig. 123 : Prefiltraggio del riferimento

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:

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

otterremmo 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 $\overline{\omega}_c$. Se fosse prescritta una banda nell'inseguimento del segnale di riferimento di valore superiore $\overline{\omega}_b > \overline{\omega}_c$, si potrebbe progettare un prefiltro avente una funzione di trasferimento di tipo **passaalto**, di espressione:

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

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:

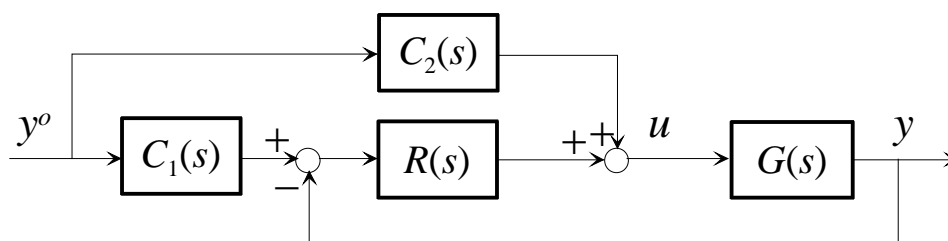


Fig. 124 : Compensazione e prefiltraggio del riferimento

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.

Capitolo 13: Controllori PID

Consideriamo un regolatore che esercita un'azione di controllo dipendente dall'errore attraverso la seguente legge:

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau + K_D \frac{de(t)}{dt}.$$

La legge di controllo è quindi composta da:

- un'azione **P**roporzionale all'errore;
- un'azione **I**ntegrale sull'errore;
- un'azione **D**erivativa sull'errore.

Questo tipo di regolatori prende quindi il nome di **PID**. I tre guadagni che compaiono nella legge di controllo vengono chiamati:

K_P : guadagno proporzionale;

K_I : guadagno integrale;

K_D : guadagno derivativo.

Alternativamente, la legge di controllo si può scrivere come segue:

$$u(t) = K_P \left[e(t) + \frac{1}{T_I} \int_0^t e(\tau) d\tau + T_D \frac{de(t)}{dt} \right],$$

dove:

$$T_I = \frac{K_P}{K_I} : \text{tempo integrale}$$

$$T_D = \frac{K_D}{K_P} : \text{tempo derivativo}$$

Tra le ragioni del vastissimo utilizzo dei regolatori PID nella pratica dell'automazione industriale (i PID sono anche detti **regolatori industriali**), ricordiamo:

- *semplicità di realizzazione* in diverse tecnologie (elettronica, idraulica, pneumatica);
- *efficacia* per la regolazione di un'ampia gamma di processi industriali;
- *standardizzazione* con i relativi vantaggi in termini di *affidabilità* e *economicità*;
- *semplicità di taratura* dei parametri;
- possibilità di *taratura automatica* dei parametri, per mezzo di semplici esperimenti.

Dal caso generale della legge di controllo PID è poi possibile derivare altre leggi di controllo,

annullando una o più delle azioni di controllo. Sono in particolare di interesse le leggi di controllo:

P (proporzionale);

PD (proporzionale-derivativa);

PI (proporzionale-integrale).

Essendo un sistema dinamico lineare e invariante, il regolatore PID può essere rappresentato da una funzione di trasferimento:

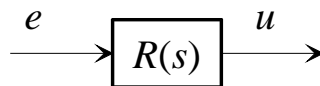


Fig. 125 : Funzione di trasferimento del regolatore PID

$$R(s) = K_P + \frac{K_I}{s} + K_D s = K_P \left(1 + \frac{1}{sT_I} + sT_D \right) = \frac{K_P}{T_I} \frac{1 + sT_I + s^2 T_I T_D}{s} .$$

Il numeratore di $R(s)$ è di grado superiore al denominatore: pertanto, così come scritta, la funzione di trasferimento non è fisicamente realizzabile. Ciò corrisponde all'impossibilità di ottenere dall'errore un segnale che ne costituisca in ogni istante la derivata. Per rendere realizzabile l'azione derivativa occorrerà in effetti aggiungere un polo in alta frequenza, per altro di norma irrilevante ai fini della valutazione delle prestazioni del regolatore PID.

Dall'ultima espressione scritta per $R(s)$ si riconosce che al variare di T_I e T_D gli zeri del regolatore possono essere reali o complessi e coniugati. Imponendo la presenza di due zeri reali e distinti, naturalmente nel semipiano sinistro, il diagramma di Bode del modulo di R assumerà l'andamento tipico riportato in figura:

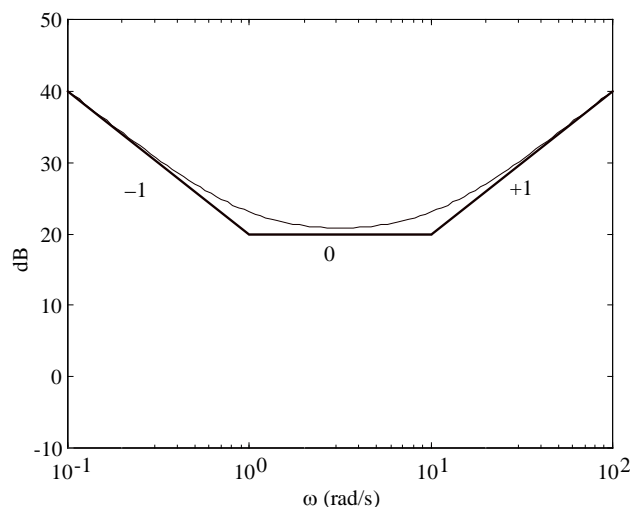


Fig. 126 : Tipico andamento del diagramma di $|R|$

Il progetto del regolatore PID si riduce quindi alla scelta del guadagno e della posizione degli zeri.

Taratura analitica dei regolatori PID

Come tutti i controllori, anche il controllore PID può essere progettato sulla base delle tecniche analitiche viste in precedenza, che fanno uso del modello matematico del sistema sotto controllo scritto in forma di funzione di trasferimento. Tuttavia, nel caso del controllore PID, i gradi di libertà nel progetto sono limitati a 3 (il guadagno e due zeri): è allora opportuno procedere in modo più diretto rispetto alla sintesi per tentativi della funzione di trasferimento d'anello già illustrata, selezionando direttamente la posizione degli zeri (tipicamente in modo da cancellare i poli del processo) e scegliendo il guadagno in modo da soddisfare le specifiche dinamiche.

Esempio

Si consideri lo schema di controllo in figura:

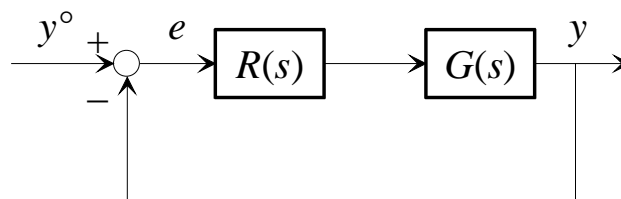


Fig. 127 : Sistema di controllo per l'esempio

dove:

$$G(s) = 0.1 \frac{e^{-3s}}{(1+5s)(1+20s)}.$$

Si vuole progettare il regolatore $R(s)$ nella classe dei regolatori PID in modo tale che:

- $e_{\infty} = 0$ per $y^o(t) = \text{sca}(t)$;
- $\varphi_m \geq 40^\circ$
- ω_c sia la massima possibile.

La specifica statica impone un regolatore di tipo 1, ossia la presenza dell'azione integrale nel regolatore PID. Scritta la funzione di trasferimento come

$$R(s) = \mu_R \frac{(1+sT_1)(1+sT_2)}{s},$$

dove $\mu_R > 0$ è il guadagno, T_1 e T_2 sono le costanti di tempo degli zeri e si è sottintesa la presenza di un polo in alta frequenza introdotto per rendere realizzabile l'azione derivativa, potremo porre:

$$T_1 = 5, T_2 = 20,$$

in modo da cancellare con gli zeri del regolatore i poli del processo.

Si ottiene quindi la funzione di trasferimento d'anello:

$$L(s) = R(s)G(s) = \frac{0.1\mu_R}{s} e^{-3s}.$$

Come è noto, il diagramma di Bode del modulo associato a L ha pendenza -1 su tutto l'asse delle pulsazioni, e taglia l'asse in corrispondenza della seguente pulsazione:

$$\omega_c = 0.1\mu_R.$$

La fase critica risulta quindi:

$$\varphi_c = -90^\circ - \omega_c \times \tau \times \frac{180^\circ}{\pi} = -90^\circ - 0.3\mu_R \frac{180^\circ}{\pi}.$$

Imponiamo il vincolo sul margine di fase:

$$\varphi_m = 90^\circ - 0.3\mu_R \frac{180^\circ}{\pi} \geq 40^\circ \Rightarrow \mu_R \leq \frac{50\pi}{0.3 \times 180^\circ} = 2.91.$$

Scegliendo $\mu_R = 2.9$ si ottiene il regolatore:

$$R(s) = 2.9 \frac{(1+5s)(1+20s)}{s} = 2.9 \frac{100s^2 + 25s + 1}{s} = K_P + \frac{K_I}{s} + K_D s,$$

con

$$K_P = 72.5, K_I = 2.9, K_D = 290.$$

Questo regolatore conferisce al sistema di controllo un margine di fase di circa 40° ed una pulsazione critica di 0.29 rad/s.

Taratura automatica dei regolatori PID

Uno dei vantaggi connessi all'utilizzo dei regolatori PID consiste nella possibilità di effettuare la taratura dei parametri sulla base di semplici prove sperimentali, prescindendo dalla formulazione matematica, non sempre agevole, del sistema sotto controllo.

Tra i numerosi **metodi empirici** per la sintonizzazione dei regolatori PID, ci limitiamo ad accennare ai due tradizionalmente più noti.

Metodo di Ziegler e Nichols in anello chiuso

Il metodo si articola nei seguenti passi:

1. Si chiude l'anello di controllo con il regolatore PID (i cui parametri devono essere sintonizzati), imponendo nulle le azioni integrale e derivativa: $K_I = 0$, $K_D = 0$.

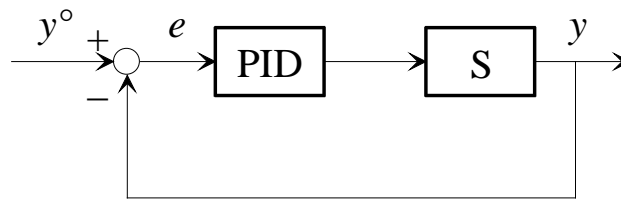


Fig. 128 : Sistema in anello chiuso con regolatore PID

2. Partendo da valori molto piccoli di K_P si effettua un semplice esperimento, consistente nell'applicare un piccolo gradino al segnale di riferimento.
3. Si aumenta progressivamente K_P ripetendo di volta in volta l'esperimento finché non si instaura nell'anello un'**oscillazione permanente**.

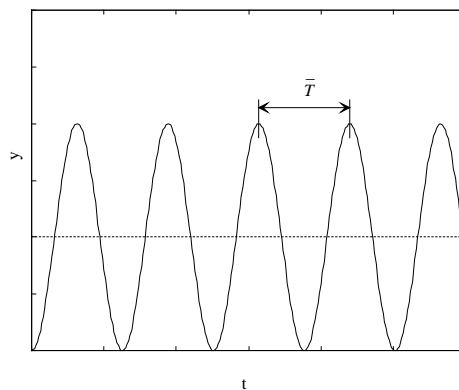


Fig.129 : Oscillazione permanente

4. Detto \bar{K}_P il valore del guadagno proporzionale corrispondente all'oscillazione permanente (*guadagno critico*) e \bar{T} il periodo di tale oscillazione, si tarano i parametri di un regolatore **P**, **PI** o **PID** sulla base della seguente tabella:

	K_P	T_I	T_D
P	$0.5 \bar{K}_P$	–	–
PI	$0.45 \bar{K}_P$	$\frac{\bar{T}}{1.2}$	–
PID	$0.6 \bar{K}_P$	$\frac{\bar{T}}{2}$	$\frac{\bar{T}}{8}$

Il metodo non è sempre applicabile: ci sono infatti sistemi che non generano oscillazioni, anche con guadagni proporzionali elevati. Altre volte può essere pericoloso, o comunque sconsigliabile, portare il sistema al limite di stabilità.

Metodo di Ziegler e Nichols in anello aperto

Il metodo si articola nei seguenti passi:

1. Si applica una variazione a scalino all'ingresso del sistema sotto controllo.

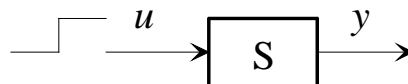


Fig.130 : Perturbazione a scalino

2. Si traccia la tangente alla risposta nel **punto di flesso**:

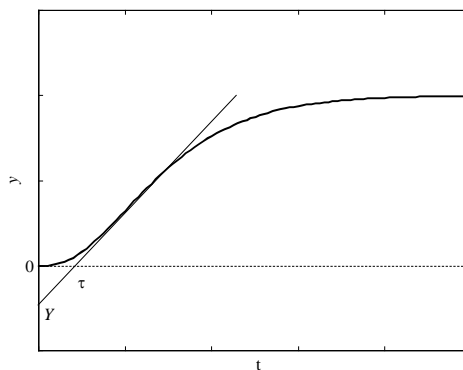


Fig.131 : Metodo della tangente nel punto di flesso

3. Si individuano graficamente le intercette τ e Y della tangente sugli assi t e y , rispettivamente.
4. Si tarano i parametri di un regolatore **P**, **PI** o **PID** sulla base della seguente tabella:

	K_P	T_I	T_D
P	$\frac{1}{Y}$	–	–
PI	$\frac{0.9}{Y}$	3τ	–
PID	$\frac{1.2}{Y}$	2τ	0.5τ

Il metodo non è ovviamente applicabile se la risposta allo scalino non presenta flesso o se la risposta presenta oscillazioni. Inoltre non sempre è possibile operare sul processo in anello aperto, o perturbare bruscamente il suo ingresso.

Capitolo 14: Schemi avanzati di controllo

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.

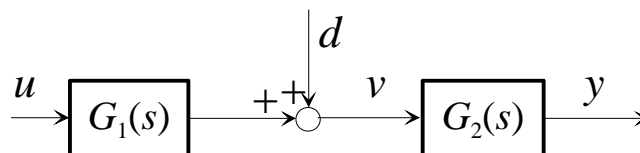


Fig. 132 : Sistemi in cascata

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 controelettromotrice. 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**.

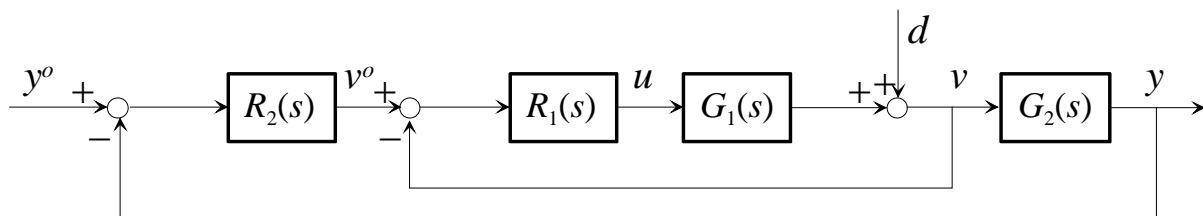


Fig. 133 : Schema di controllo in cascata

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 $G_2(s)$.

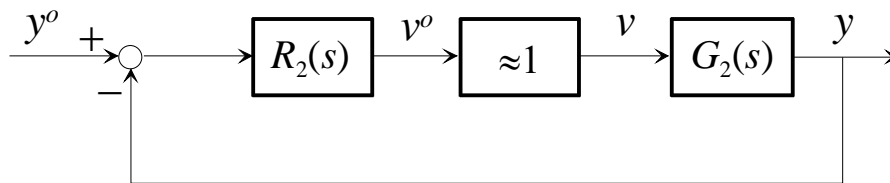


Fig. 134 : Schema per il progetto del regolatore esterno

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:

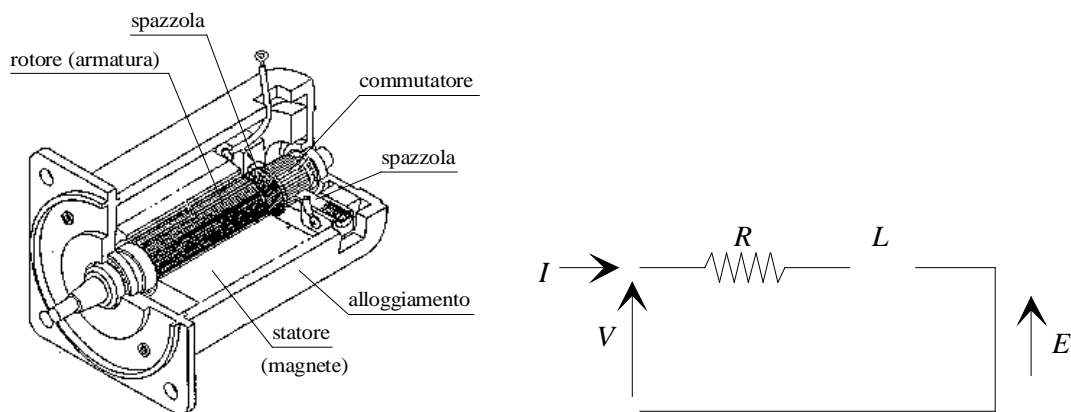


Fig. 135 : Motore a corrente continua e schema del circuito d'armatura

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:

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

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

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

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

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

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:

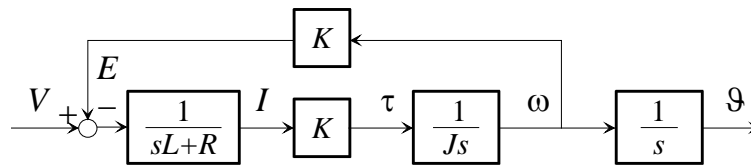


Fig. 136 : Schema a blocchi del motore a corrente continua

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

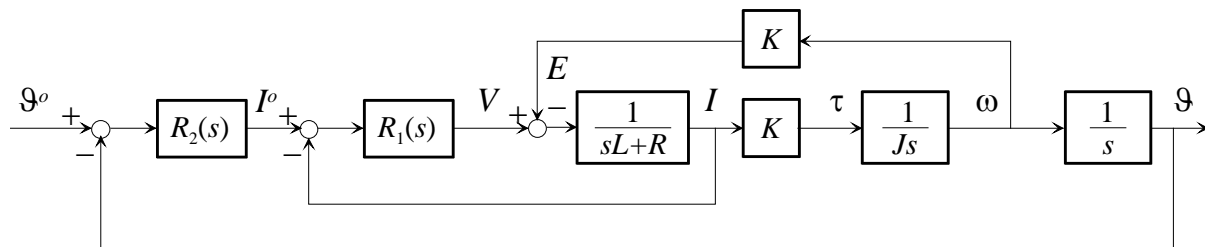


Fig. 137 : Controllo in cascata del motore a corrente continua

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 transitori 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:

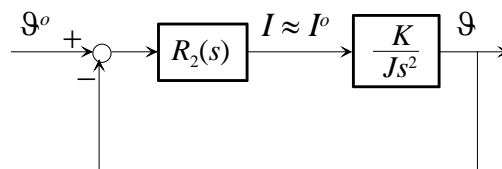


Fig. 138 : Schema per il progetto del controllore esterno (di posizione)

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:

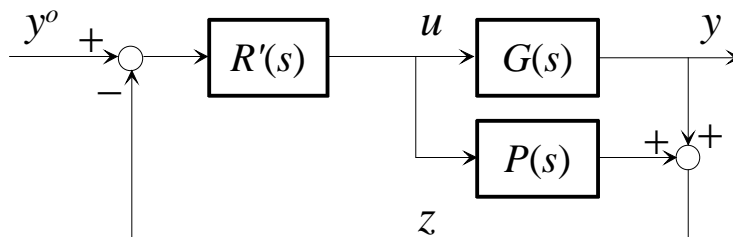


Fig. 139 : Schema di controllo con predittore di Smith

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:

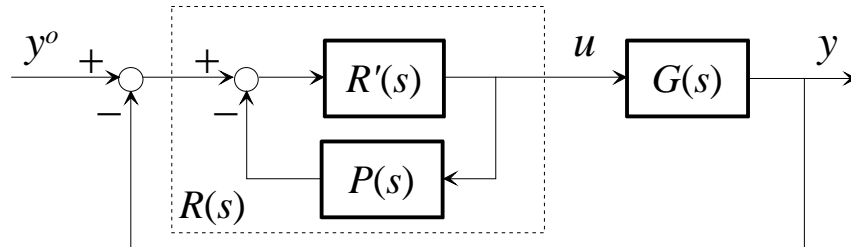


Fig. 140 : Interpretazione dello schema di controllo con predittore di Smith

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:

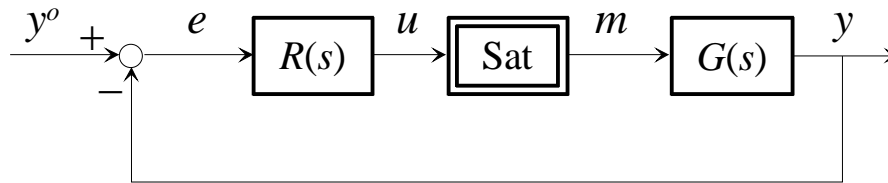


Fig. 141 : Sistema di controllo con saturazione

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

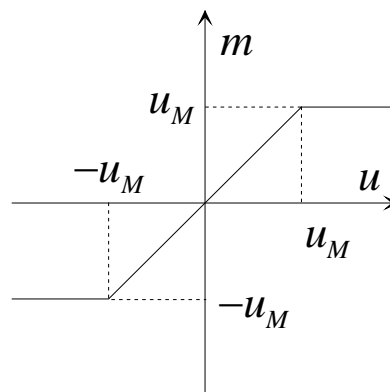


Fig. 142 : Caratteristica di saturazione

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 $H(s)$.

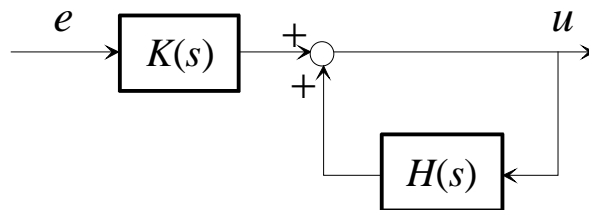


Fig. 143 : Realizzazione alternativa di $R(s)$

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:

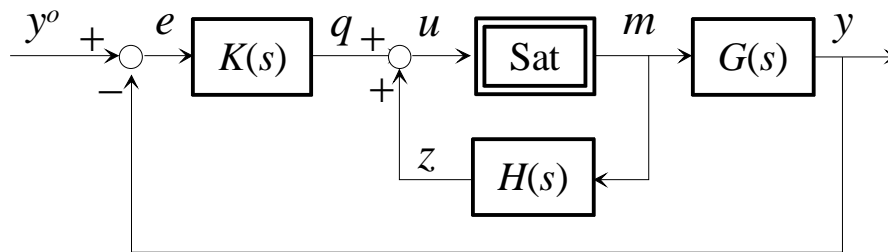


Fig. 144 : Realizzazione anti windup del regolatore

Se invece la variabile m non è accessibile, si dovrebbe riprodurre all'interno del regolatore, una caratteristica, indicata con $\hat{\text{Sat}}$ in Fig. 145, uguale a quella fisica dell'attuatore, e porla sulla linea d'andata dello schema a blocchi che realizza il regolatore, secondo lo schema seguente:

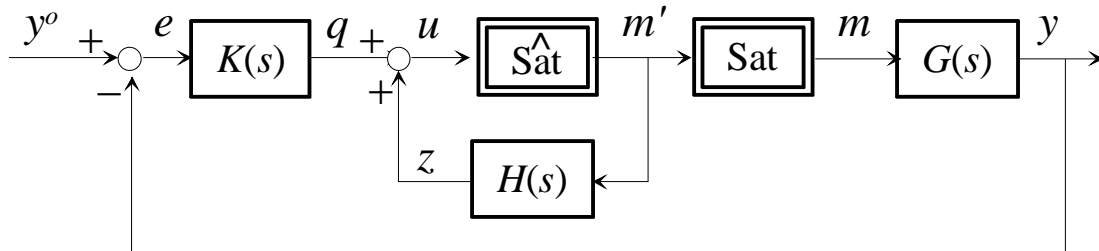


Fig. 145 : Realizzazione anti windup del regolatore senza misura di m

Facendo riferimento alla Fig. 144 (ma le stesse considerazioni valgono per lo schema di Fig. 145), 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:

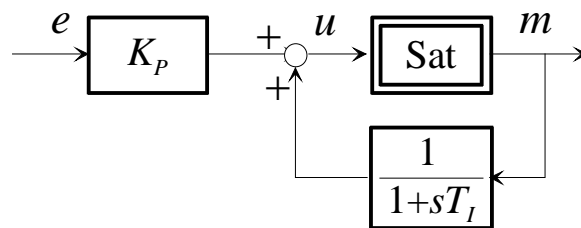


Fig. 146 : Realizzazione anti windup del regolatore PI

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 $G(s)$.

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:

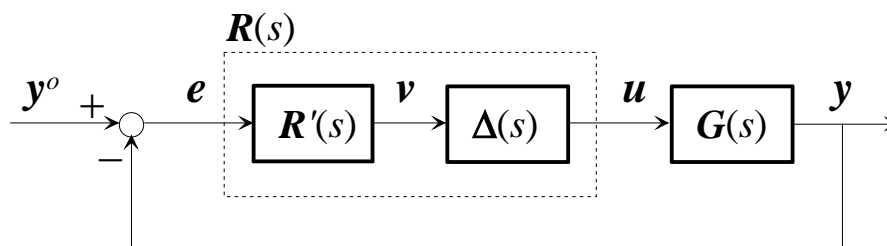


Fig. 147 : Sistema di controllo con disaccoppiatore

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

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

Il sistema di funzione di trasferimento $\Delta(s)$, 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 $G(s)\Delta(s)$ **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 $G(s)\mathbf{R}(s)\mathbf{R}'(s)$ 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 $\Delta(s)$.

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:

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

quella del disaccoppiatore. Imponiamo che il prodotto $G(s)\Delta(s)$ 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 $G_{di}(s)$.

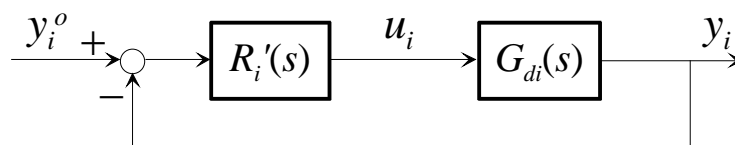


Fig. 148 : Problema di controllo monovariabile

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:

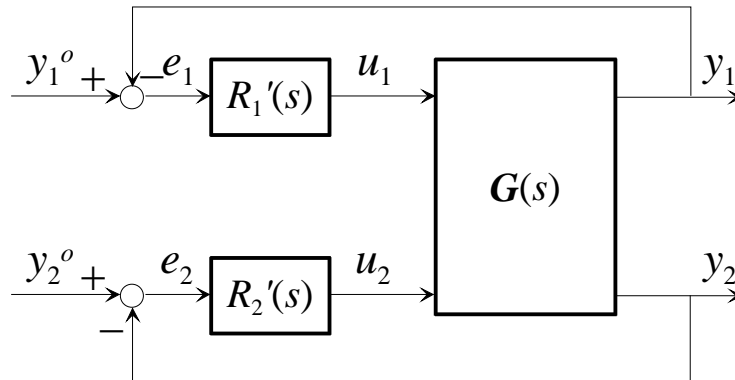


Fig. 149 : Sistema di controllo decentralizzato

A questo punto, si potrebbe condurre il progetto di $R'_i(s)$ sulla base dell' i -esimo 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_1^o in quanto influente per la discussione. Lo schema corrisponde alle equazioni:

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

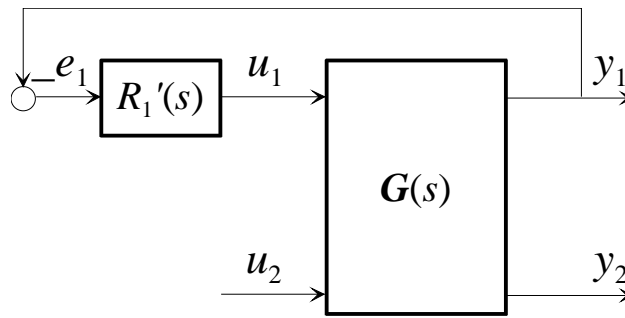


Fig. 150 : Sistema sotto controllo per il secondo regolatore

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 $G(s)$, si passa ad un altro accoppiamento permutando opportunamente righe e colonne di $G(s)$. 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:

$$\triangleleft = G(0) \otimes (G(0)^{-1})^T,$$

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 Λ più diagonale, con elementi sulla diagonale prossimi a 1.

Per $m=2$, posto:

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

si ottiene facilmente:

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

con:

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

Si osservi che se $\lambda=1$, Λ 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.

Capitolo 15: Luogo delle radici

Si consideri il sistema retroazionato di Fig. 151, la cui funzione di trasferimento d'anello è $L(s)$.

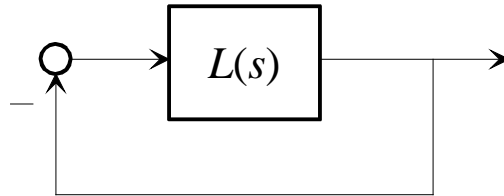


Fig. 151 : Sistema retroazionato

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:

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

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

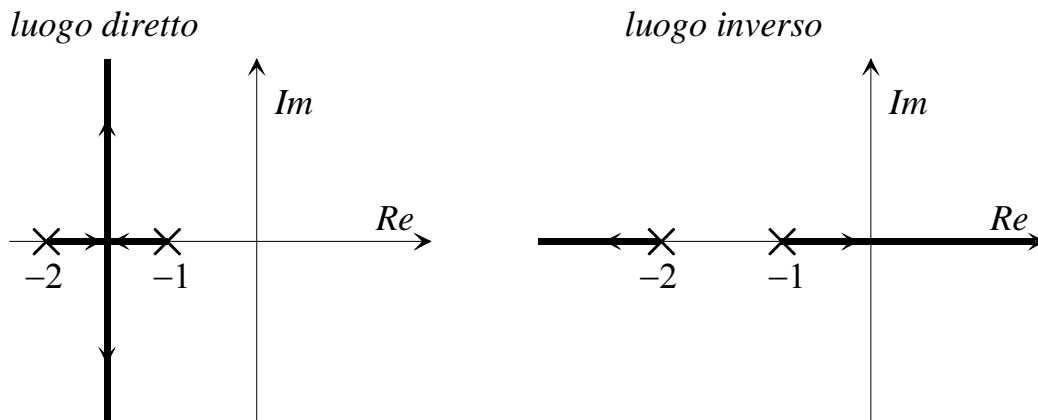


Fig. 152 : Luoghi delle radici per l'esempio introduttivo

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 enunciarne 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:

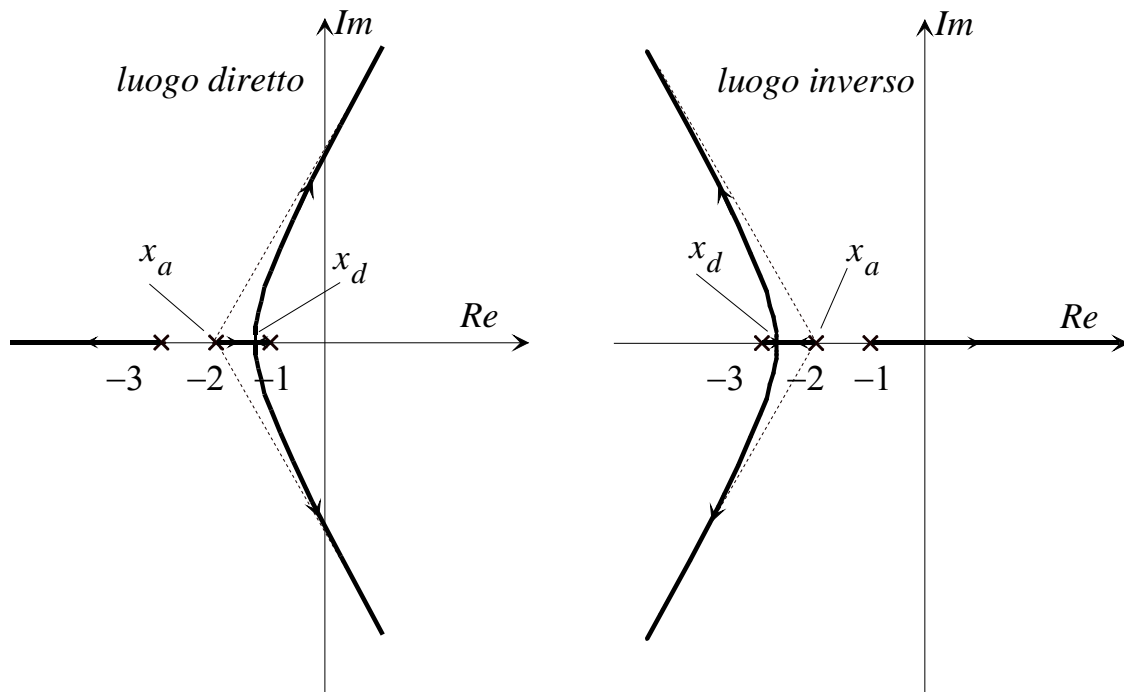


Fig. 153 : Luoghi delle radici per l'esempio 1

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:

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:

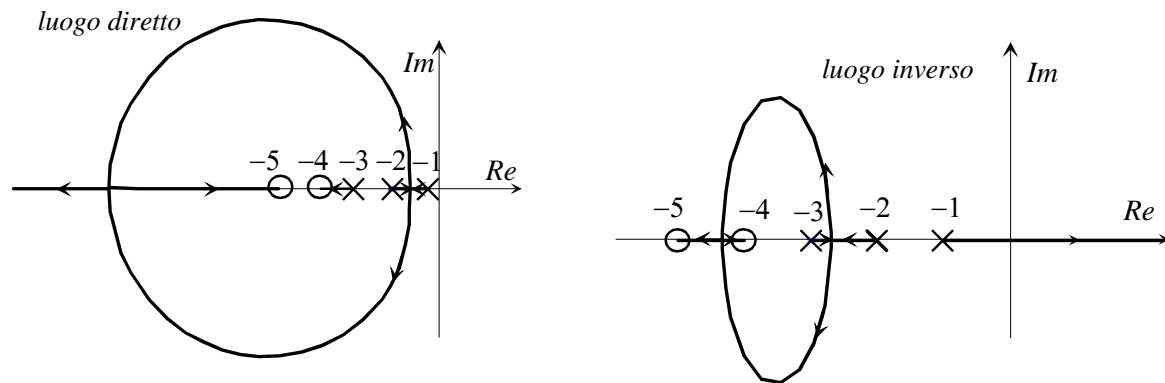


Fig. 154 : Luoghi delle radici per l'esempio 2

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:

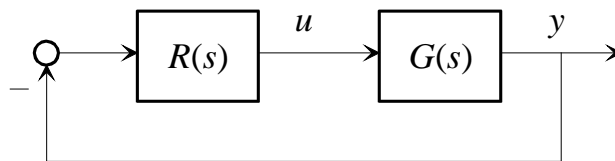


Fig. 155 : Sistema di controllo in anello chiuso

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:

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

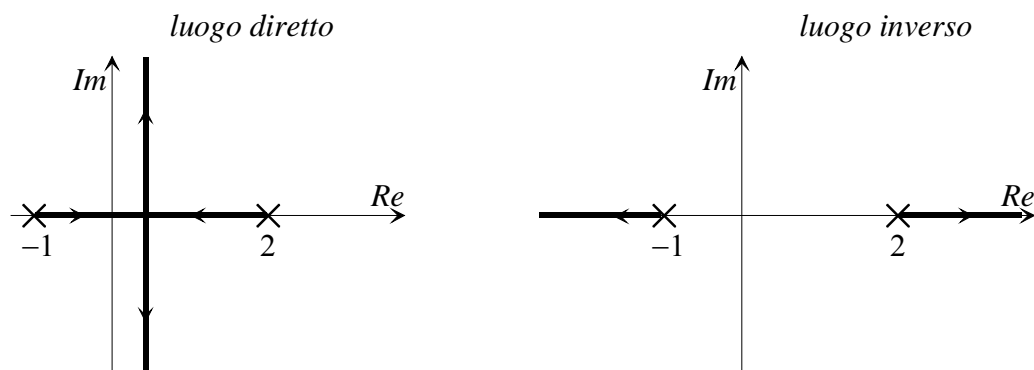


Fig. 156 : Luoghi delle radici con regolatore proporzionale

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 .

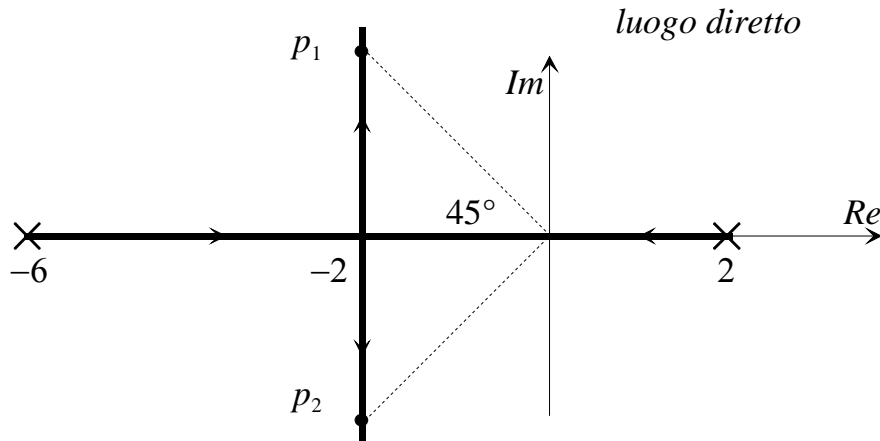


Fig. 157 : Luogo delle radici modificato

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):

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

Capitolo 16: Assegnamento degli autovalori

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.

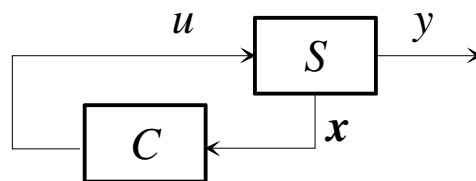


Fig. 158 : Assegnamento degli autovalori con informazione completa

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.

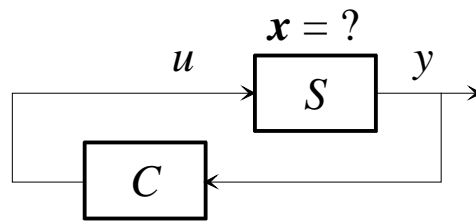


Fig. 159 : Assegnamento degli autovalori con informazione parziale

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.

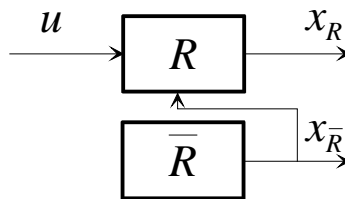


Fig. 160 : Parti raggiungibile e non raggiungibile

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

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

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

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

Sia $u(t) = \mathbf{K}\mathbf{x}(t)$ con $\mathbf{K} = [k_1 \quad k_2 \quad \cdots \quad k_n]$ 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:

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

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

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

$$= \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1 + k_1 & -a_2 + k_2 & \cdots & \cdots & -a_n + k_n \end{bmatrix}.$$

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:

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

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:

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

mentre quella del sistema in forma canonica di controllo è:

$$\begin{aligned}\hat{\mathbf{K}}_r &= [\hat{\mathbf{B}} \quad \hat{\mathbf{A}}\hat{\mathbf{B}} \quad \hat{\mathbf{A}}^2\hat{\mathbf{B}} \quad \dots] = [\mathbf{TB} \quad \mathbf{TAT}^{-1}\mathbf{TB} \quad \mathbf{TAT}^{-1}\mathbf{TAT}^{-1}\mathbf{TB} \quad \dots] = \\ &= \mathbf{T}[\mathbf{B} \quad \mathbf{AB} \quad \mathbf{A}^2\mathbf{B} \quad \dots] = \mathbf{TK}_r.\end{aligned}$$

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:

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

Conclusione: 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:

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

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:

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

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

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

Pertanto la forma canonica di controllo del sistema è:

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

cui corrisponde la matrice di raggiungibilità:

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

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:

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

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:

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

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

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

Si può facilmente verificare che gli autovalori della matrice $A+BK$ 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 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 B_i la i -sima colonna di B , la coppia (A, B_i)

risultati 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 K_i in modo che gli autovalori della matrice $A+B_iK_i$ siano in posizioni desiderate del piano complesso.

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

$$K_r = [B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B],$$

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 K di dimensioni $m \times n$ per cui gli autovalori della matrice $A+BK$ 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.

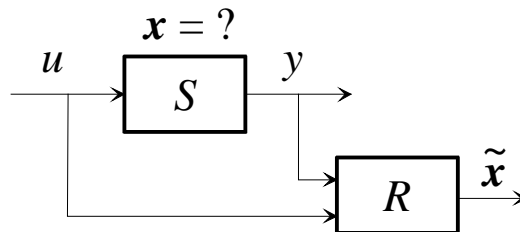


Fig. 161 : Stima dello stato

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:

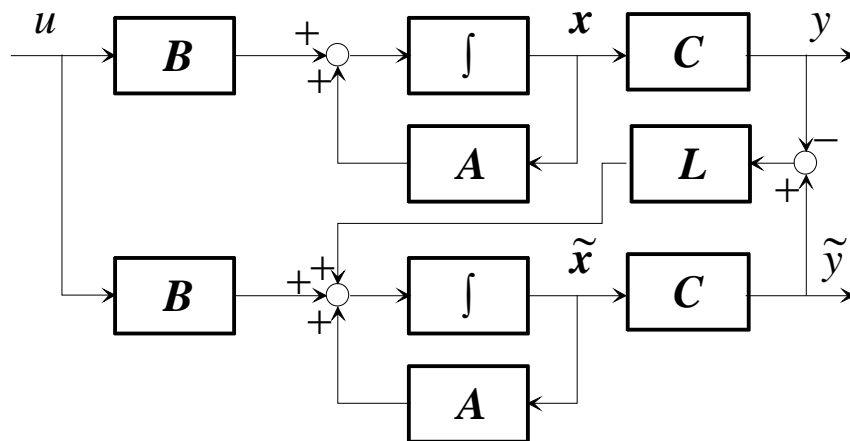


Fig. 162 : Sistema con ricostruttore

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 ¹⁰:

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

e posto:

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

si ricava:

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

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

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

si ottengono le relazioni:

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

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

¹⁰ 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:

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

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:

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

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:

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

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

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

e quindi ovviamente:

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

Si può facilmente verificare che gli autovalori della matrice $\mathbf{A} + \mathbf{LC}$ 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:

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

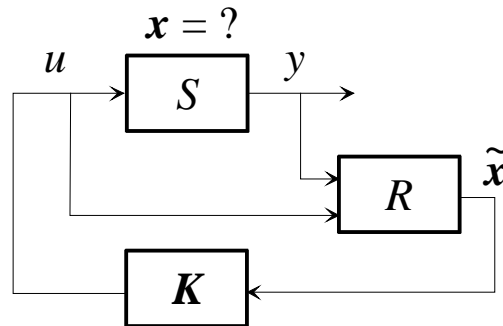


Fig. 163 : Legge di controllo sulla stima dello stato

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

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

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

$$y(t) = Cx(t)$$

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

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

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

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

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

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

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

$$\dot{x}(t) = (A + BK)x(t) - BK\epsilon(t)$$

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

La matrice dinamica del sistema in anello chiuso è quindi:

$$A_c = \left[\begin{array}{c|c} A + BK & -BK \\ \hline 0 & A + LC \end{array} \right],$$

e risulta triangolare a blocchi.

Ne consegue che gli autovalori della matrice sono la riunione degli autovalori delle due sottomatrici, $A+BK$ e $A+LC$, sulla diagonale. Sappiamo che se la coppia (A, 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} :

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

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

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 $\mathbf{A} + \mathbf{BK} + \mathbf{LC}$.

Si osservi che, poiché dalla stabilità delle matrici $\mathbf{A} + \mathbf{BK}$ e $\mathbf{A} + \mathbf{LC}$ non si evince la stabilità della matrice $\mathbf{A} + \mathbf{BK} + \mathbf{LC}$, 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:

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

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

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.

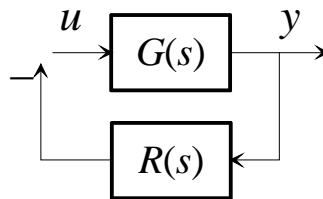


Fig. 164 : Connessione tra sistema sotto controllo e controllore

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.

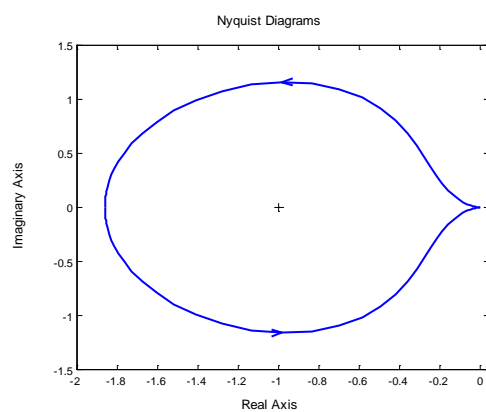


Fig. 165 : Diagramma di Nyquist associato a L

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:

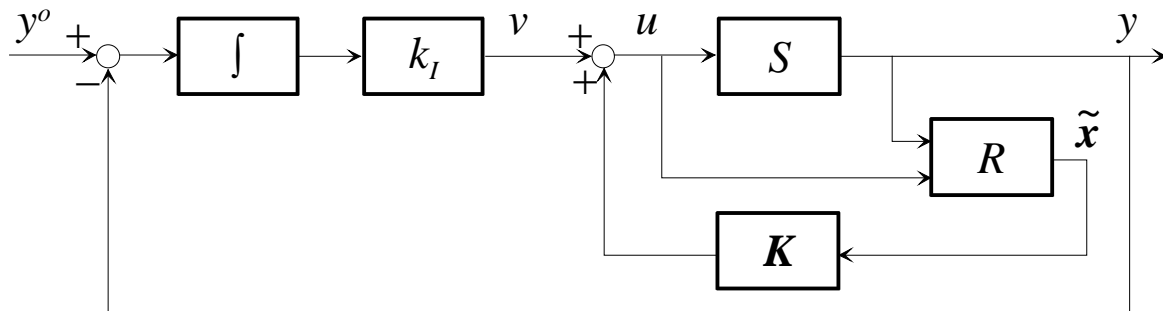


Fig. 166 : Introduzione dell'integratore

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:

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

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 $(\mathbf{F}, \mathbf{G}_u)$. 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 risolubile 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:

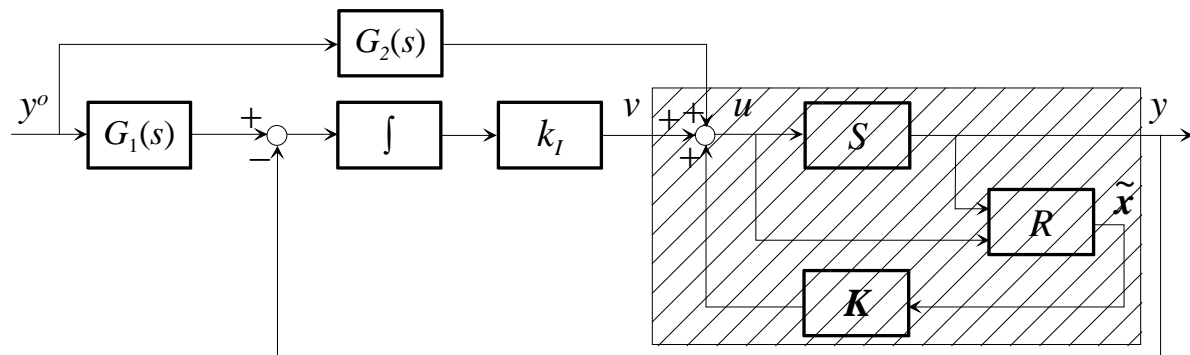


Fig. 167 : Introduzione degli elementi in feedforward

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 .

Capitolo 17: Sistemi dinamici a tempo discreto

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.

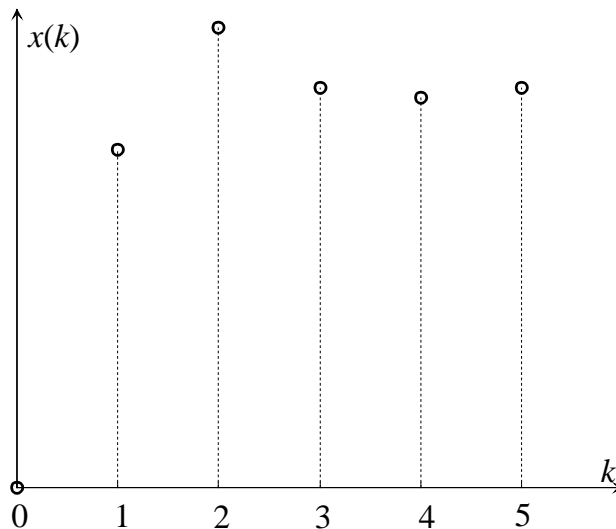


Fig. 168 : Segnale a tempo discreto

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.

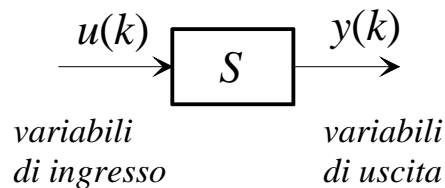


Fig. 169 : Sistema a tempo discreto

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:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{f}(\mathbf{x}(k), \mathbf{u}(k), k) \\ \mathbf{y}(k) &= \mathbf{g}(\mathbf{x}(k), \mathbf{u}(k), k) \end{aligned}$$

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:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned}$$

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:

$$\begin{aligned} x(1) &= -(1/2)^3 = -1/8 \\ x(2) &= -(-1/8)^3 = 1/512 \\ x(3) &= -(1/512)^3 = -1/134217728 \end{aligned}$$

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:

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

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:

$$x_1(k+1) = c(k+1) = \alpha y(k) = \alpha(c(k) + i(k) + u(k)) =$$

$$= \alpha x_1(k) + \alpha x_2(k) + \alpha u(k)$$

$$x_2(k+1) = i(k+1) = \beta(c(k+1) - c(k)) = \beta(\alpha(c(k) + i(k) + u(k)) - c(k)) =$$

$$= \beta(\alpha - 1)x_1(k) + \beta\alpha x_2(k) + \beta\alpha u(k)$$

$$y(k) = x_1(k) + x_2(k) + u(k)$$

Il sistema è LTI, con matrici:

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

$$\mathbf{C} = [1 \quad 1], \quad D = 1$$

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:

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

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

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

\vdots

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:

$$\bar{x} = (I - A)^{-1} B \bar{u}$$

Inoltre risulta:

$$\bar{y} = \mu \bar{u}$$

con:

$$\mu = C(I - A)^{-1} B + D$$

guadagno statico del sistema.

E' utile ricordare che, effettuando un cambiamento di variabili di stato:

$$\hat{x}(k) = T x(k), \quad \det T \neq 0$$

le matrici del sistema si trasformano esattamente come a tempo continuo:

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

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:

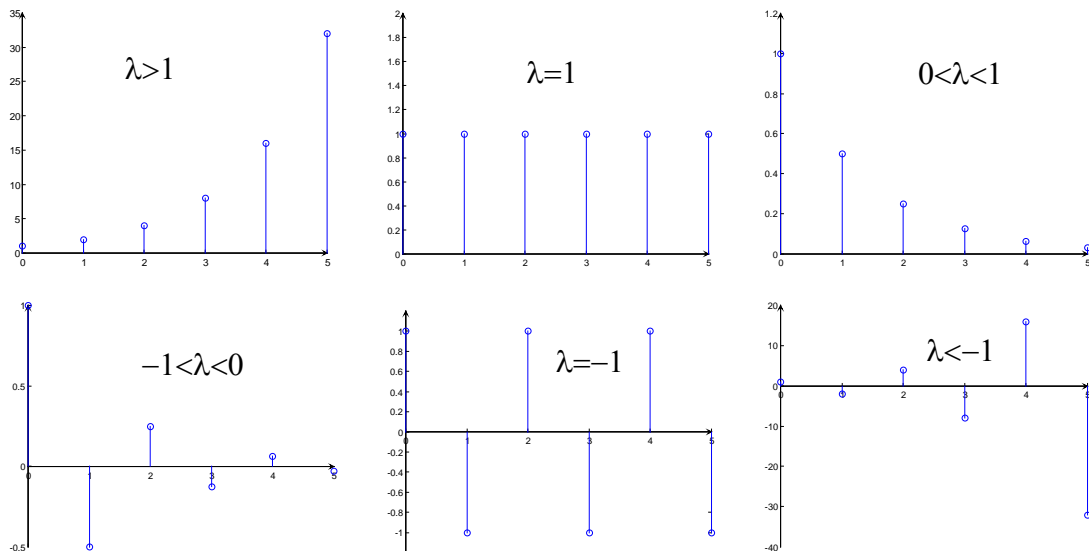


Fig. 170 : Esponenziale a tempo discreto

Naturalmente, accanto ad un autovalore complesso $\lambda_i = \rho_i e^{j\vartheta_i}$ ci sarà anche il coniugato e la combinazione dei due termini darà luogo ad un termine reale del tipo $\rho_i^k \cos(\vartheta_i k + \varphi_i)$, 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 \mathbf{A} hanno modulo minore di 1;
- stabile:** se e solo se tutti gli autovalori di \mathbf{A} hanno modulo minore o uguale a 1 e ne esistono a modulo uguale a 1;
- instabile:** se e solo se esistono autovalori di \mathbf{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 \mathbf{T} , la matrice \mathbf{A} del sistema si trasforma secondo una relazione di similitudine ($\hat{\mathbf{A}} = \mathbf{TAT}^{-1}$). 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 \mathbf{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:

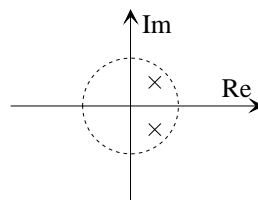


Fig. 171 : Autovalori per l'esempio economico

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:

$$|z|^2 = \frac{|1+x+jy|^2}{|1-x-jy|^2} = \frac{x^2 + y^2 + 2x + 1}{x^2 + y^2 - 2x + 1} < 1 \quad \Leftrightarrow \quad x < 0.$$

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:

$$\begin{array}{ccc} 27 & 9 & 0 \\ 27 & 1 & 0 \\ 8 & 0 & \\ 1 & & \end{array}$$

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:

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

Siamo interessati allo studio di eventuali equilibri che si hanno quando l'ingresso è nullo: $\mathbf{u}(k) = \bar{\mathbf{u}} = 0$. 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:

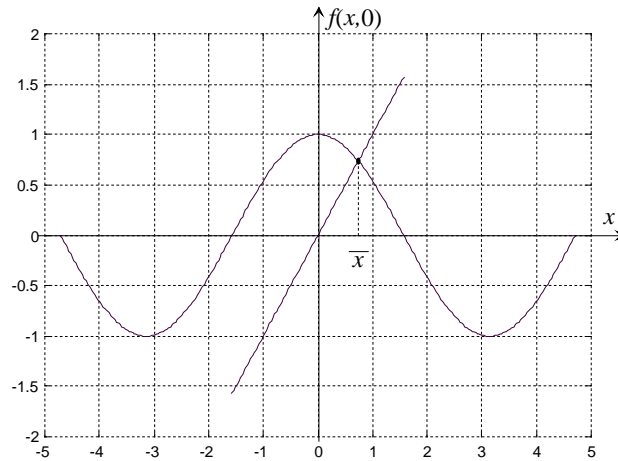


Fig. 172 : Ricerca grafica dello stato di equilibrio

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.

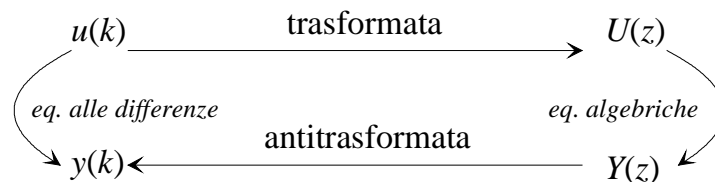


Fig. 173 : Approccio nel dominio delle trasformate

Si consideri dunque una generica funzione reale $v(k)$, definita per k intero ≥ 0 , tale che, per almeno un $r \in \mathbb{R}$, $r > 0$, 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:

$$v(0) = \lim_{z \rightarrow \infty} V(z) = 1$$

$$v(\infty) = \lim_{z \rightarrow 1} [(z-1)V(z)] = \lim_{z \rightarrow 1} \left[(z-1) \frac{z}{z-a} \right] = 0 \quad \Leftrightarrow \quad |a| < 1$$

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:

$v(k)$	$V(z)$
$\text{imp}(k)$	1
$\text{sca}(k)$	$\frac{z}{z-1}$
$\text{ram}(k)$	$\frac{z}{(z-1)^2}$
$\text{par}(k)$	$\frac{z}{(z-1)^3}$
a^k	$\frac{z}{z-a}$
ka^k	$\frac{az}{(z-a)^2}$

dove $\text{par}(k) = k(k-1)/2$, $k \geq 0$.

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:

$$V(z) = \frac{N(z)}{D(z)} = \beta_0 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots \Rightarrow \begin{aligned} v(0) &= \beta_0 \\ v(1) &= \beta_1 \\ v(2) &= \beta_2 \end{aligned}$$

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

$$\begin{aligned} z\mathbf{X}(z) &= \mathbf{A}\mathbf{X}(z) + \mathbf{B}U(z) \\ \mathbf{Y}(z) &= \mathbf{C}\mathbf{X}(z) + \mathbf{D}U(z) \end{aligned} \Rightarrow \begin{aligned} \mathbf{X}(z) &= (z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}U(z) \\ \mathbf{Y}(z) &= [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}]U(z) \end{aligned}$$

Si è ottenuto:

$$\mathbf{Y}(z) = \mathbf{G}(z)\mathbf{U}(z), \quad \mathbf{G}(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D}$$

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:

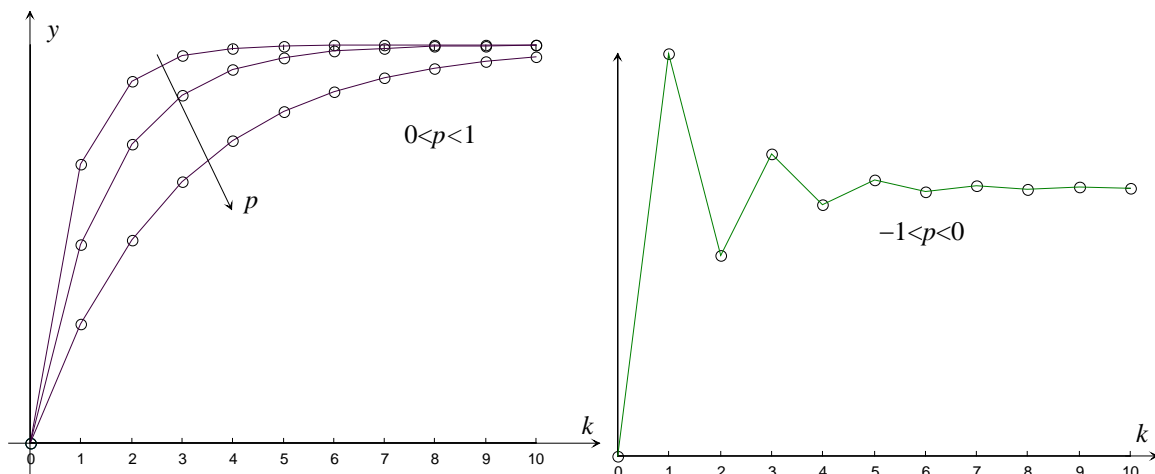


Fig. 174 : Risposta allo scalino di un sistema del primo ordine

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:

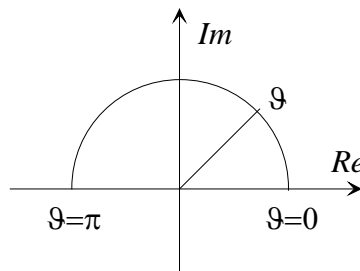


Fig. 175 : Luogo dei punti per il calcolo della risposta in frequenza

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.

Capitolo 18: Controllo digitale

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):

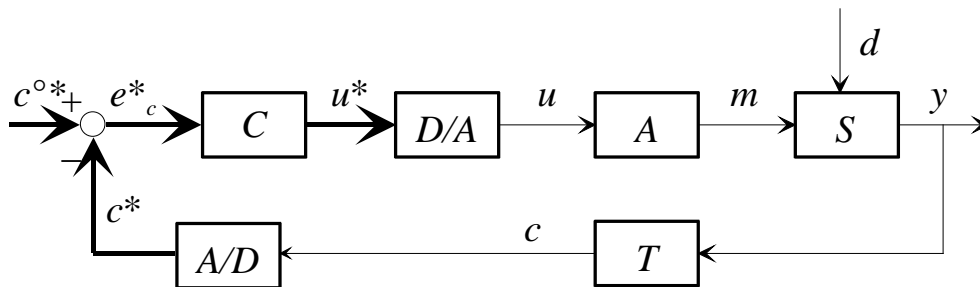


Fig. 176 : Sistema di controllo digitale

Uno schema alternativo in cui si converte in digitale l'errore è il seguente:

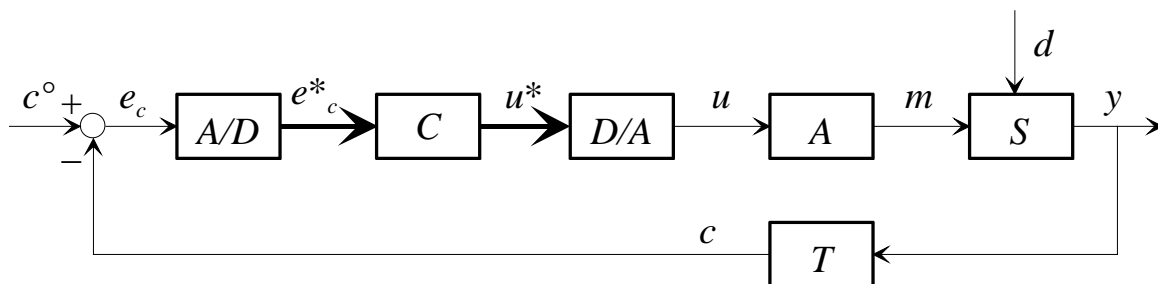


Fig. 177 : Sistema di controllo digitale con conversione in digitale dell'errore

Supporremo 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 $v(t)$, $t \in \mathbb{R}$. 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:

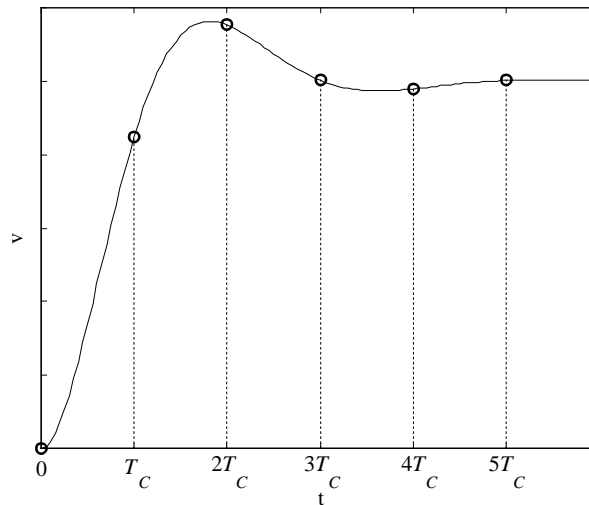


Fig. 178 : Campionamento

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:

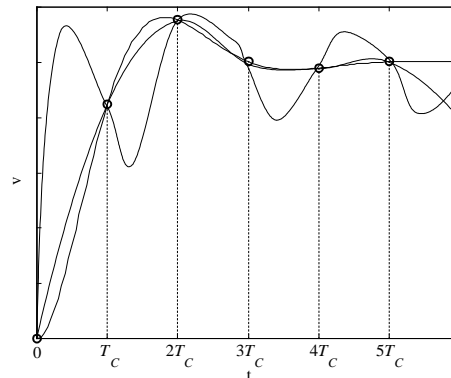


Fig. 179 : Segnali analogici che danno luogo allo stesso segnale campionato

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:

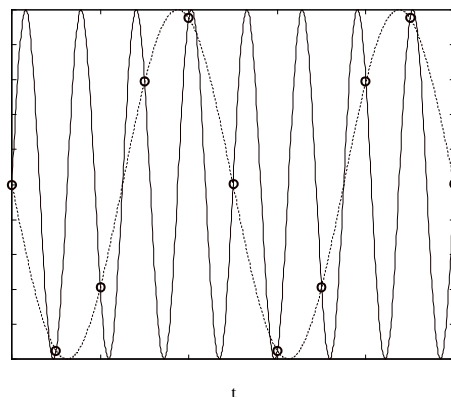


Fig. 180 : Aliasing

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 sinusoide. 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 sinusoide di periodo \bar{T} con tempo di campionamento $T_C = 3/4 \bar{T}$:

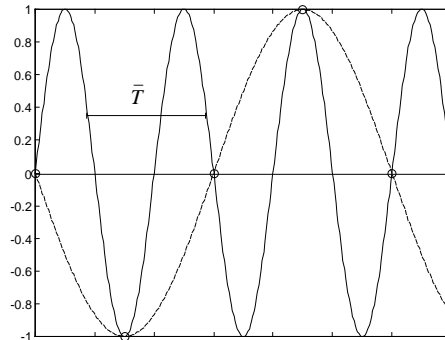


Fig. 181 : Campionamento di una sinusoide

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 sinusoide 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.

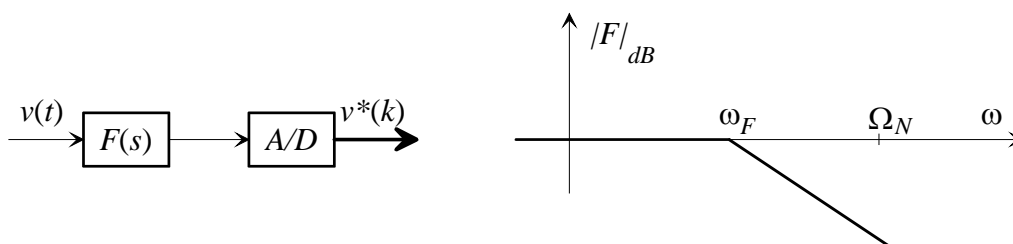


Fig. 182 : Filtro anti-aliasing

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.

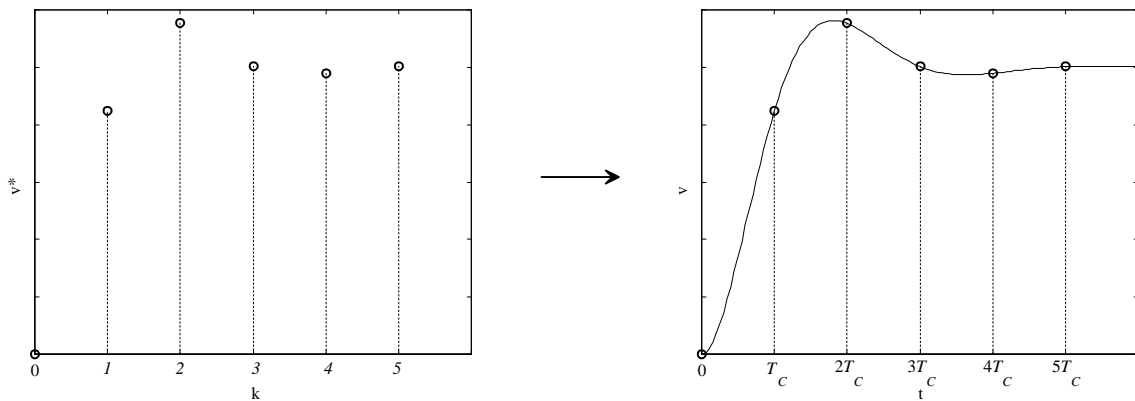


Fig. 183 : Conversione D/A

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:

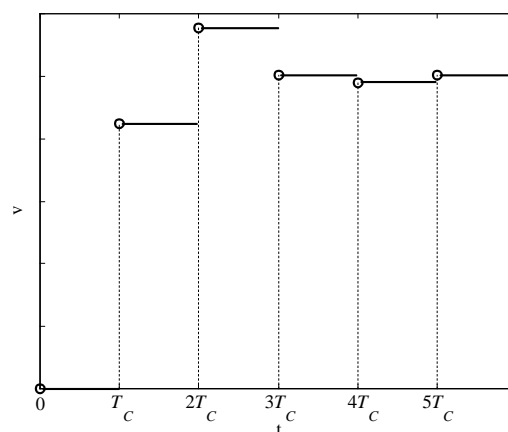


Fig. 184 : Mantenitore di ordine zero

Il dispositivo che realizza questa operazione si chiama **Mantenitore di ordine zero** o **ZOH** (**Z**ero **O**rdere **H**old).

Consideriamo ora la sequenza di un campionatore ed uno ZOH:

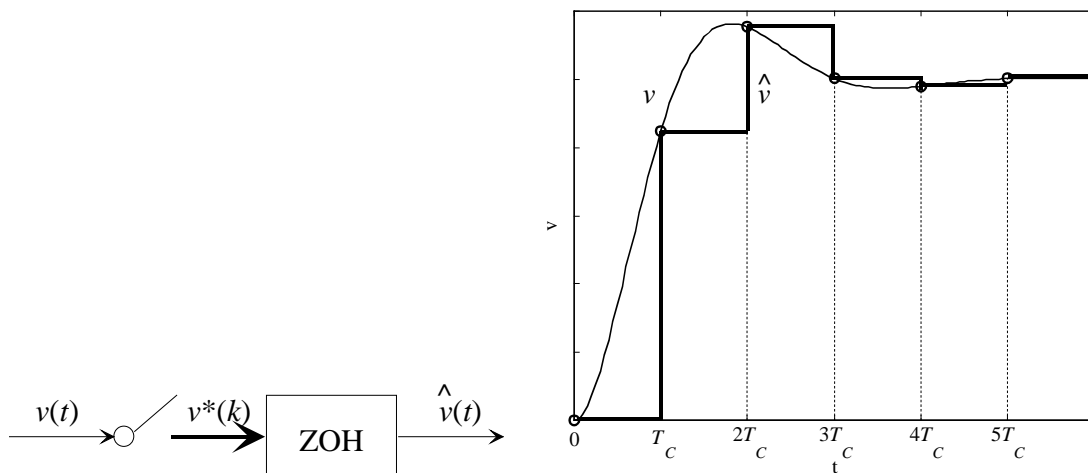


Fig. 185 : Campionatore e mantentore

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:

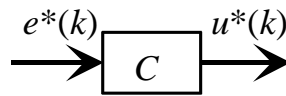


Fig. 186 : Regolatore digitale

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 mantenitore 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 $e^*(k)$:

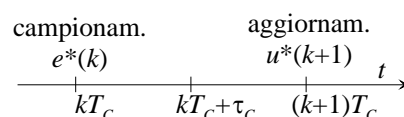


Fig. 187 : Temporizzazione nel caso di regolatore strettamente proprio

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:

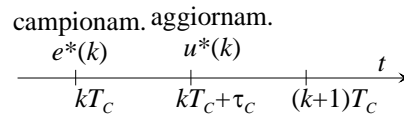


Fig. 188 : Temporizzazione nel caso di regolatore proprio non strettamente

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 :

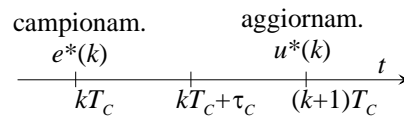


Fig. 189 : Temporizzazione alternativa nel caso di regolatore proprio non strettamente

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

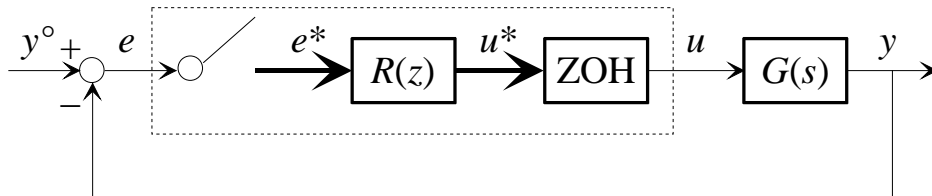


Fig. 190 : Punto di vista analogico

La parte tratteggiata è un sistema esternamente a tempo continuo. Dato un regolatore analogico ci proporremo di realizzarlo in tecnologia digitale, con un'opportuna scelta di $R(z)$ e del periodo di campionamento.

Punto di vista digitale

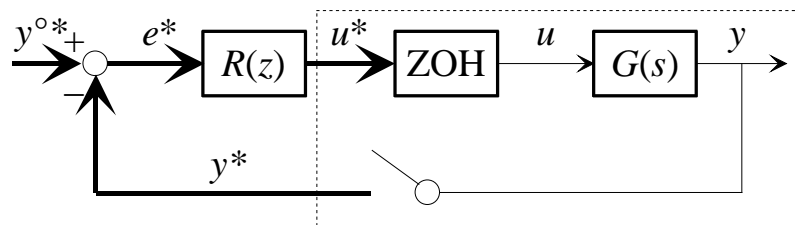


Fig. 191 : Punto di vista digitale

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 $R^o(s)$ e di volerlo realizzare mediante un regolatore a tempo discreto $R(z)$ con un tempo di campionamento T .

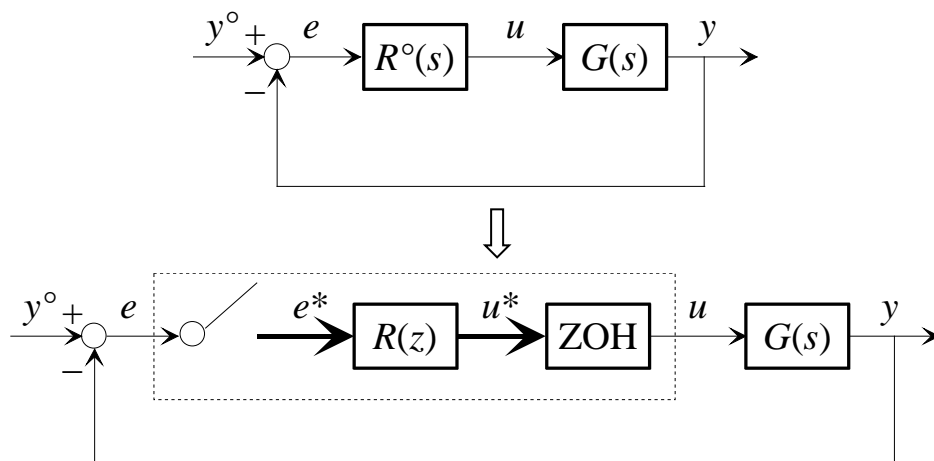


Fig. 192 : Realizzazione digitale del controllore

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**:

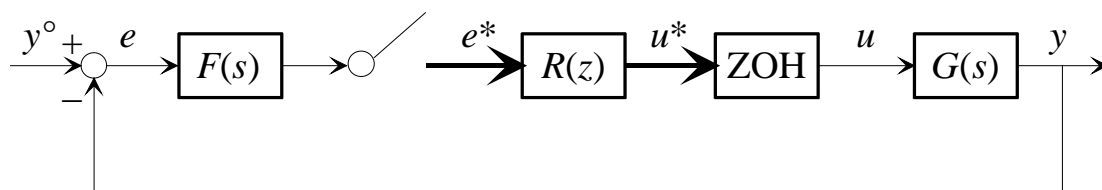


Fig. 193 : Inserimento di un filtro antialiasing nell'anello

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):

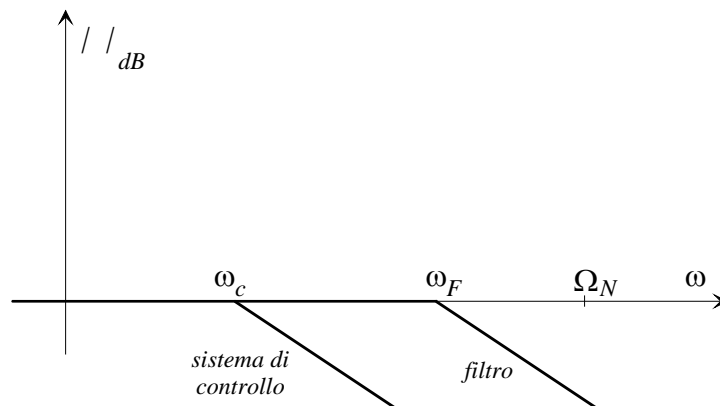


Fig. 194 : Progetto del filtro antialiasing

Si osservi a questo punto che il fatto stesso di realizzare il regolatore in tecnologia digitale con campionatori e mantenitori 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:

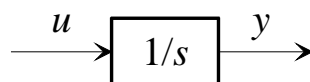


Fig. 195 : Sistema da integrare numericamente

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:

$$\begin{aligned} \alpha = 0 \quad s &= \frac{z-1}{T} && \text{Eulero in avanti (o Eulero esplicito)} \\ \alpha = 1 \quad s &= \frac{z-1}{Tz} && \text{Eulero all'indietro (o Eulero implicito)} \\ \alpha = \frac{1}{2} \quad s &= \frac{2}{T} \frac{z-1}{z+1} && \text{Trapezi o Tustin} \end{aligned}$$

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:

$$R^o(s) = k_p \left(1 + \frac{1}{sT_I} \right).$$

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**.

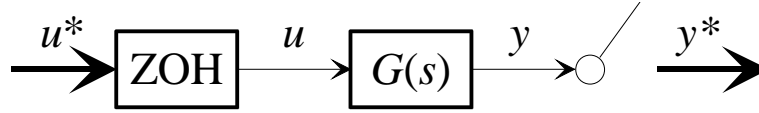


Fig. 196 : Sistema a segnali campionati

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:

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

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:

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}_0 + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)d\tau.$$

Sceghieremo ora come istante iniziale t_0 un generico istante di campionamento e come istante t il successivo istante di campionamento, ossia $t_0 = kT$, $t = kT + T$:

$$\mathbf{x}(kT + T) = e^{\mathbf{A}T}\mathbf{x}(kT) + \int_{kT}^{kT+T} e^{\mathbf{A}(kT+T-\tau)}\mathbf{B}u(\tau)d\tau.$$

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:

$$\mathbf{x}^*(k+1) = e^{\mathbf{A}T}\mathbf{x}^*(k) + \left[\int_{kT}^{kT+T} e^{\mathbf{A}(kT+T-\tau)}\mathbf{B}d\tau \right] u^*(k).$$

Eseguito il cambiamento di variabili:

$$\sigma = kT + T - \tau$$

otteniamo:

$$\mathbf{x}^*(k+1) = e^{AT} \mathbf{x}^*(k) + \left[\int_0^T e^{A\sigma} \mathbf{B} d\sigma \right] \mathbf{u}^*(k)$$

In definitiva, il sistema a tempo discreto si lascia descrivere dalle seguenti equazioni:

$$\begin{aligned} \mathbf{x}^*(k+1) &= \mathbf{A}^* \mathbf{x}^*(k) + \mathbf{B}^* \mathbf{u}^*(k) \\ \mathbf{y}^*(k) &= \mathbf{C} \mathbf{x}^*(k) + \mathbf{D} \mathbf{u}^*(k) \end{aligned},$$

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:

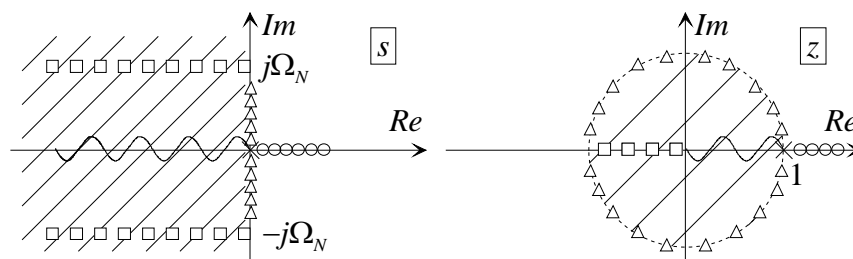


Fig. 197 : Trasformazione di campionamento

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 esposte costituisce 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:

$$\mathbf{A}^* = e^{\mathbf{A}T} = e^{-pT}, \quad \mathbf{B}^* = \int_0^T e^{\mathbf{A}\sigma} \mathbf{B} d\sigma = \int_0^T e^{-p\sigma} d\sigma = \left[\frac{e^{-p\sigma}}{-p} \right]_0^T = \frac{1 - e^{-pT}}{p},$$

e dalla funzione di trasferimento:

$$G^*(z) = \mathbf{C}(\mathbf{zI} - \mathbf{A}^*)^{-1} \mathbf{B}^* + \mathbf{D} = \frac{\mathbf{CB}^*}{z - \mathbf{A}^*} = \frac{1 - e^{-pT}}{z - e^{-pT}}.$$

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:

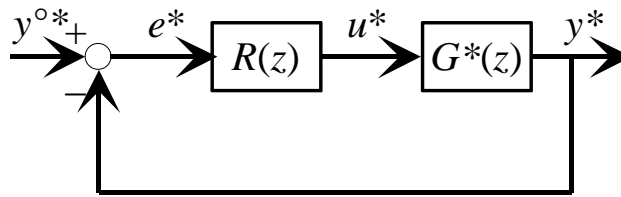


Fig. 198 :Sistema di controllo a tempo discreto

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 transitori 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 $B(z)$:

$$\boxed{N(z_i) = 0, \quad |z_i| \geq 1 \quad \Rightarrow \quad B(z_i) = 0}$$

- Ogni polo p_i di $G^*(z)$ ($D(p_i)=0$), a modulo maggiore o uguale a 1 deve essere anche radice di $A(z)-B(z)$:

$$D(p_i)=0, \quad |p_i| \geq 1 \Rightarrow A(p_i)=B(p_i)$$

3) Precisione statica

Per avere errore nullo a regime con riferimento a scalino, $F^*(z)$ deve avere guadagno unitario ($F^*(1)=1$):

$$A(1)=B(1)$$

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:

$$F^*(z) = \frac{B(z)}{z^{n_A}}$$

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 $y^{o*}(k)=sca(k)$, dopo n_A passi y^* non cambia più valore.

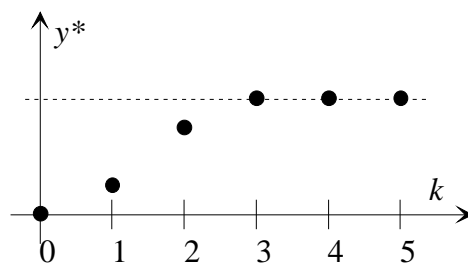


Fig. 199 :Controllo in tempo finito

Se si vuole in particolare che l'uscita raggiunga il valore di regime nel **tempo minimo**, sarà sufficiente porre:

$$F^*(z) = \frac{1}{z^{\delta_G}}$$

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:

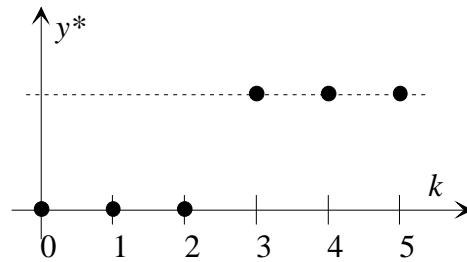


Fig. 200 : Controllo in tempo minimo

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:

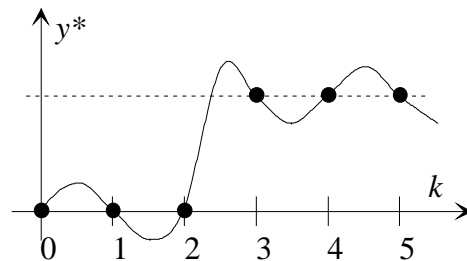


Fig. 201 : Oscillazioni nascoste

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).