

Controlli Automatici

Alberto Tibaldi

12 giugno 2009

Indice

1	Il problema del controllo automatico	4
1.1	Terminologia e concetti preliminari	4
1.2	Formulazione del problema del controllo automatico	7
1.3	Una struttura generale di controllo automatico	8
1.4	Esempi	8
2	Rappresentazioni e proprietà dei sistemi dinamici	11
2.1	Terminologia e concetti preliminari	13
2.2	Rappresentazione di sistemi dinamici mediante relazioni ingresso- stato-uscita	14
2.3	Esempio “limite”	16
2.4	Il movimento dello stato e dell’uscita nei sistemi LTI	17
2.5	Brevi richiami di algebra lineare	18
2.5.1	Esempio pratico	19
2.6	La stabilità interna	20
2.6.1	Esempio pratico	21
2.7	Scelta dei parametri di simulazione per Simulink	22
2.8	Rappresentazioni di sistemi dinamici mediante relazioni ingresso- uscita: funzione di trasferimento	24
2.8.1	Esempio pratico	25
2.9	La stabilità esterna nei sistemi LTI	26
2.9.1	Esempio pratico	26
2.9.2	Esempio pratico	27
2.9.3	Alcuni risultati aggiuntivi	27
2.9.4	Esempio pratico	28
2.10	Forme e parametri della funzione di trasferimento	30
2.10.1	Guadagno stazionario	32
2.10.2	Puntualizzazione sull’uso dei teoremi della \mathcal{L} -trasformata	33

3	La risposta in frequenza di sistemi LTI	34
3.1	Esistenza di una uscita sinusoidale	34
3.2	Definizione di risposta in frequenza	35
3.3	Esempio pratico	36
3.4	Rappresentazioni grafiche della risposta in frequenza	37
3.4.1	Diagrammi di Bode	37
3.4.2	Diagrammi polari e diagrammi di Nyquist	44
3.4.3	Esempi teorici/pratici	45
3.5	Il criterio di Nyquist per lo studio della stabilità di sistemi retroazionati	47
4	Le caratteristiche dei sistemi di controllo retroazionati con un ingresso e una uscita (SISO)	51
4.1	Una struttura di un sistema di controllo	51
4.2	La stabilità nei sistemi di controllo con retroazione	55
4.3	Sistemi di controllo: la risposta in regime permanente a ingressi polinomiali	57
4.3.1	Concetti preliminari	57
4.3.2	Caso dei sistemi retroazionati	60
4.4	Sistemi di controllo: la risposta in regime permanente a disturbi additivi	65
4.4.1	Disturbi polinomiali	65
4.4.2	Disturbi sinusoidali	67
4.5	La sensibilità alle variazioni parametriche	72
4.6	Introduzione ai problemi di laboratorio	75
4.7	La risposta transitoria di sistemi di controllo con retroazione	78
4.7.1	Caratterizzazione generale nel tempo	79
4.7.2	Parametri della risposta al gradino di sistemi prototipo del secondo ordine	80
4.8	Risposta in frequenza di un sistema di controllo	81
4.8.1	Caratterizzazione generale in frequenza	81
4.8.2	Parametri della risposta in frequenza di sistemi prototipo del secondo ordine	82
4.9	Relazioni tra la risposta al gradino e la risposta in frequenza nei sistemi prototipo del secondo ordine	83
4.10	Curve di T e S a modulo costante	85
4.10.1	Curve di T a modulo costante	85
4.10.2	Curve di S a modulo costante	87
4.11	Indicatori di margini di stabilità	87
4.12	Esempio (traduzione di $\hat{s} \leq \hat{s}_0$)	89
4.13	La carta di Nichols	90

5	Metodi di sintesi basati sull'impiego della risposta in frequenza	94
5.1	Funzioni compensatrici elementari	96
5.1.1	Controllore ad azione proporzionale	96
5.1.2	Controllore ad azione anticipatrice	99
5.1.3	Controllore ad azione attenuatrice	99
5.2	Esempio di progetto di rete integrativa	100
5.2.1	Esempio di progetto di rete anticipatrice	102
5.2.2	Simulazione degli esempi 1, 2, 3	103
5.3	Studio del segno di K_c	106
5.4	Attività del "comando" in funzione del riferimento	107
5.4.1	Analisi	107
5.4.2	Progetto	108
5.5	Attività del comando in funzione di un ingresso sinusoidale	109
5.5.1	Analisi	109
5.5.2	Progetto	110
5.6	Riduzione della complessità delle reti correttrici	110
5.7	Realizzazione di controllori analogici mediante reti RC-attive	111
6	Introduzione al controllo digitale	113
6.1	Struttura di sistema di controllo digitale	113
6.2	Modello matematico sul campionamento	114
6.3	Spettro del segnale campionato	115
6.4	Ricostruzione del segnale di partenza mediante filtro ideale e teorema del campionamento	116
6.5	Ricostruzione mediante ZOH	117
6.6	Legame tra $F^*(s)$ e $F(z)$	118
6.7	Schemi a blocchi di sistemi con campionatori	119
6.8	\mathcal{Z} -trasformata di funzioni contenenti ZOH	122
6.9	Scelta del tempo di campionamento	122
6.10	Metodi di discretizzazione	124
6.10.1	Metodo di invarianza della risposta all'impulso (\mathcal{Z} -trasformata)	124
6.10.2	Metodo di invarianza della risposta al gradino	125
6.10.3	Metodo di trasposizione zeri-poli (matched)	126
6.11	Progetto di controllori digitali per discretizzazione di controllori analogici	126
6.11.1	Progetto di $G_c(z)$ per loop shaping di $G_a(s) = G_c(s)AG_pG_tG_y$	126
6.11.2	Progetto di $G_c(z)$ per loop shaping di $G'_a(s) = G_a(s) \cdot G_h(s)$	127
6.11.3	Progetto di $G_c(z)$ mediante loop shaping di $G_a(z)$	127

Capitolo 1

Il problema del controllo automatico

1.1 Terminologia e concetti preliminari

La terminologia e la notazione sono molto importanti: al fine di avere un linguaggio comune mediante il quale scambiare informazioni, è necessario accordarsi sul differente significato dei termini. Discutendo sul significato di una delle parole chiave del titolo del capitolo, “controllo”, si può, al fine di comprendere meglio di cosa si parla, proporre alcune definizioni proposte da tre differenti dizionari:

- Dizionario 1: unione o insieme delle azioni volte a far assumere a una grandezza un certo valore o una certa successione di valori al variare del tempo; data ad esempio una cisterna, di cui si vuole controllare il livello, si deve in qualche modo “guardare” il livello e aprire/chiusure il collegamento con la rete idrica mediante una valvola, azionata in qualche modo;
- Dizionario 2: dispositivo al quale è affidato il governo del valore di una grandezza fisica. Qua il dizionario potrebbe essere contestabile: identifica l’azione di “controllo” con il controllore, ossia con un dispositivo in grado di effettuare l’azione di controllo.
- Dizionario 3: dispositivo automatico che controlla un altro dispositivo. Doppio errore: come prima non si parla di un dispositivo, e da nessuna parte è detto che un controllo sia automatico, come si vedrà.

In questo ambito, ha sicuramente più senso la parola inglese “control”, che indica per l’appunto l’azione del controllo nel senso che si intende attribuire

in questa trattazione: la possibilità di intervenire nelle azioni del sistema, in modo da poterne modificare alcune proprietà.

Un controllore dovrà decidere quale azione deve fare su di un impianto affinché la grandezza interessata abbia un certo andamento nel tempo. Una grandezza è un ente suscettibile di misura (ossia che può essere misurata).

Un sistema è un insieme di elementi (volendo, fisici, ma non è detto: un sistema può anche essere economico per esempio), di elementi tra loro interconnessi, di cui si intende ricavare un modello matematico. Un sistema può ad esempio essere una rete elettrica, un'automobile, un circuito idrico, o qualsiasi altra cosa. Prima di tutto, al fine di realizzare un controllo su di un qualche sistema, come ad esempio un'automobile, è necessario modellizzare in maniera anche dettagliata un'automobile, per poi progettare, con tecniche che verranno in seguito discusse, il controllore, in modo che esso abbia la possibilità di intraprendere un'azione.

Esistono due grandi categorie di entità che lavorano con un sistema:

- Cause (ingressi)
- Effetti (uscite)

In un circuito ad esempio le “cause” possono essere i generatori, gli “effetti” le varie uscite del sistema, quali correnti su resistenze, tensioni, o simili.

Un generico sistema \mathcal{P} si può rappresentare come un blocco, nella seguente maniera:

Dove per $u(t)$ si intende una causa, un ingresso del sistema, e per $\nu(t)$ un'uscita. Possono esserci molteplici ingressi e uscite, come si sa anche dalla teoria dei sistemi; si può dire in altri termini che $u(t)$ e $\nu(t)$ siano funzioni vettoriali:

$$u(t) \in \mathbb{R}^m$$

$$\nu(t) \in \mathbb{R}^p$$

Quando si costruisce il modello di un sistema si deve assolutamente fare attenzione a non dimenticare gli ingressi: tutte le cause, al momento della modellizzazione di un sistema, devono assolutamente essere considerate. Discorso differente per quanto concerne le uscite: non tutte le uscite, spesso, saranno fondamentali ai termini dello studio del sistema, dunque ci saranno di fatto alcune uscite che saranno preferite ad altre.

Il problema del controllo si può porre quando si verificano le seguenti condizioni, quando ci sono i seguenti elementi, in un generico sistema \mathcal{P} :

1. Esistono uscite primarie, ossia è richiesto un certo andamento nel tempo di alcune uscite $\nu(t)$. Le uscite dove è fondamentale che alcune uscite assumano dunque particolari valori nel tempo, sono dette primarie. Esse vengono comunemente indicate come $y(t)$.
2. Esistono disturbi: alcune grandezze appartenenti agli insiemi degli ingressi, su cui nessun operatore può agire, influenzano l'insieme delle uscite primarie; in tal caso, essi vengono chiamati "disturbi", e vengono identificati come $d(t)$.
3. Esistono comandi: quando si ha a che fare con cause, ingressi in grado di modificare le uscite primarie sui quali si può agire ad arbitrio, si ha a che fare con un "comando", $c(t)$.
4. Esistono uscite secondarie: appartengono agli insiemi delle uscite, ma in realtà non sono fondamentali, in ambito di controllistica, come le precedenti proprietà: non si desidera che essi abbiano particolari andamenti al variare del tempo. Esse vengono indicate come $\xi(t)$.

Un primo esempio molto banale per quanto riguarda un sistema controllato è un regolatore di tensione, ossia un dispositivo elettronico in grado, a partire da una tensione qualunque in ingresso, di proporre una in uscita pressochè costante. Esaminiamo gli attori in questo tipo di sistema:

1. L'unica uscita primaria è ovviamente la tensione sul carico, poichè la grandezza fisica che si intende regolare è la tensione che il dispositivo elabora, regola;
2. Uno dei disturbi più influenti (se ne possono identificare altri) è la corrente di carico: in un regolatore di tensione, non si sa ovviamente a priori il carico sul quale esso dovrà lavorare e mantenere costante la tensione, dunque la corrente che scorrerà sul carico sarà un disturbo: la corrente sul carico, di fatto, cambia l'uscita primaria. Si può dire, a dispetto di quello che potrebbe sembrare, che la corrente di carico sia un ingresso, nel senso appena introdotto: poichè è una grandezza in grado di modificare l'uscita primaria, è identificabile come disturbo, definito come ingresso non modificabile a libero arbitrio. Altro tipico disturbo in alimentatori di questo tipo è la linea: regolazione di linea e regolazione di carico sono due delle caratteristiche più importanti che un dispositivo di questo tipo deve avere;
3. Esistono diverse uscite secondarie in ogni circuito elettronico; non interesseranno, per ora.

Si può quindi pensare in maniera più elaborata il blocco precedentemente proposto, introducendo tutta la classificazione per ora introdotta:

1.2 Formulazione del problema del controllo automatico

Bisogna proporre una formulazione concettuale, in grado di permettere di capire quale sia l'obiettivo del controllo. Si indica con $y_d(t)$ un'uscita primaria desiderata, in contrapposizione a $y(t)$ che sarà l'uscita primaria effettiva. $y_d(t)$ è una funzione astratta, desiderata idealmente dal progettista, esistente di fatto solo sulla carta, quindi teorica: si tratta della funzione alla quale si vorrebbe che tenda la funzione effettiva, misurabile in uscita al sistema, $y(t)$.

Al fine di quantificare la bontà del controllo, si definisce l'errore sull'uscita come:

$$e^y(t) = y(t) - y_d(t)$$

In questo modo, valutando la differenza tra uscita reale e uscita ideale, è possibile comprendere quanto sia buono il controllo.

Si introduce un insieme di elementi atti a permettere lo studio di un particolare elemento:

- Dato il sistema \mathcal{P} ;
- Dato l'andamento desiderato dell'uscita primaria, $y_d(t)$;
- Dati limiti ammissibili sull'errore $e^y(t)$, tali per cui ad esempio

$$|e^y(t)| < \rho_t \quad \forall t$$

- Date tutte le informazioni possibili riguardo i disturbi $d(t)$ (i disturbi sono o caratterizzabili statisticamente, o, più difficilmente, caratterizzabili puntualmente mediante sistemi di misura; si noti che un misuratore costa, dunque si tende a utilizzare una semplice caratterizzazione statistica);
- Date le misure delle uscite primarie $y(t)$ (che devono assolutamente essere misurabili, altrimenti non ha senso effettuare il controllo);
- Date (eventualmente, in maniera del tutto facoltativa) le misure delle uscite secondarie $\xi(t)$

È possibile decidere l'andamento da imporre al comando $c(t)$ affinché i limiti sull'errore $e^y(t)$ non siano superati, siano rispettati.

Con riferimento a ciò che si è detto, è possibile progettare un dispositivo in grado di fornire gli andamenti richiesti per i comandi $c(t)$ in completa assenza di operatori umani, ottenendo un cosiddetto “controllo automatico”. In un generico controllo, dunque, può o meno esserci un operatore umano, ma se si parla di controllo automatico il sistema deve essere in grado di controllarsi senza alcun operatore.

1.3 Una struttura generale di controllo automatico

Partendo dagli elementi finora introdotti, è possibile disegnare una struttura generale di controllo automatico. Dato un $y_d(t)$ desiderato, ossia un'uscita ideale, di riferimento, e la misura di $y(t)$, si potrebbe ad esempio ideare ciò:

Supponiamo che si vogliano misurare anche i disturbi (introducendo dunque una sistemistica di misura anche per quanto concerne i disturbi), con un trasduttore τ_d ; l'uscita primaria viene trascinata e misurata mediante un trasduttore τ_y , e stessa cosa per quanto riguarda (in questo specifico esempio di struttura) le uscite secondarie $\xi(t)$ con τ_ξ . Si introduce un'unità intelligente, in grado di prendere le decisioni, tenendo conto di tutte le informazioni. Si introduce un attuatore \mathcal{A} atto ad amplificare i segnali decisi.

Questa è la struttura generale di un controllo con retroazione: quando viene misurata l'uscita primaria del sistema, si dice che il controllo sia con retroazione, o in catena chiusa. Si parla di controllo in catena aperta quando non esistono né τ_ξ né τ_y ; di solito non si mette comunque neanche τ_d , a meno che non si intenda effettuare un particolare tipo di controllo, detto “di compensazione diretta dei disturbi”.

Se in presenza di disturbi gli andamenti desiderati delle uscite sono idealmente costanti nel tempo, anziché di controllo si parla di “regolazione”.

1.4 Esempi

Si propongono a questo punto alcuni esempi, atti a chiarire le idee finora proposte.

1. Un tostapane è un esempio di sistema con controllo: l'uscita primaria è il grado di doratura del pane. Questo sistema è ovviamente senza retroazione: non esiste un sistema (che non sia molto costoso) in grado

di misurare il grado di doratura del pane. Non essendovi osservazione, misura dell'uscita, il sistema è in catena aperta.

Per quanto tempo e come dunque devo tenere collegato il tostapane? Su che base si progetta la legge di controllo? La risposta è molto semplice: sull'esperienza dell'operatore, ossia sull'esperienza umana: a seconda di come si è abituati a preparare i toast, si imposterà un timer per più o meno tempo. Tutti i sistemi di controllo a catena aperta sono basati sull'uso dell'esperienza umana, ossia sul fatto che esiste un operatore in grado di controllare manualmente il sistema.

Un esempio di disturbo del sistema è il grado di umidità del pane: cambiando il pane con uno di tipo diverso, di fatto, si potrebbero variare i tempi di cottura, vanificando l'esperienza. Stessa cosa il tempo di vita del tostapane: trattandosi di un sistema tempo-variante, dopo un certo tempo non si comporterà più allo stesso modo, vanificando anche in questo caso l'esperienza.

2. Livello dell'acqua in un serbatoio: l'uscita primaria è il livello dell'acqua, il comando è la posizione della valvola in grado di aumentare o diminuire l'acqua presente nel serbatoio. L'utenza che prende l'acqua dal serbatoio è un disturbo: è di fatto imprevedibile conoscere l'utenza d'acqua in un qualche ambiente, quale ad esempio un condominio.

Esistono fondamentalmente tre strategie con le quali si può realizzare il controllo; se ne propongono le idee, in modo da chiarire ciò che è stato finora detto:

- (a) Regolatore basato sull'operatore umano: un operatore vede il livello e, a seconda della propria esperienza, apre o chiude la valvola in misura tale da regolare il livello;
- (b) Regolatore con sistema tipo galleggiante: come d'altra parte lo sciacquone del bagno, esistono metodi "intelligenti" in grado di regolare in maniera del tutto automatica il livello dell'acqua;
- (c) Regolatore con compensazione diretta dei disturbi: si misura l'utenza e si agisce esclusivamente sulla valvola, di conseguenza a qualsiasi utenza. Si noti che questo, a prima vista, può essere un metodo in grado, al prezzo di introdurre un trasduttore per la misura dell'utenza, di migliorare le prestazioni del sistema. Ciò non è di fatto vero: possono esserci buchi nel serbatoio, o l'utenza potrebbe cambiare improvvisamente abitudini, dunque il controllo di fatto non è funzionante, poichè, per situazioni di questo tipo, l'esperienza

non è sufficiente al fine di regolare il livello: l'esperienza non può prevedere cambi del sistema.

Capitolo 2

Rappresentazioni e proprietà dei sistemi dinamici

Questo capitolo contiene richiami da teoria dei sistemi, dunque di fatto contiene nozioni che dovrebbero essere teoricamente già note a chi legge la trattazione; dal momento che alcuni elementi fondamentali potrebbero non essere noti, e che si tratta comunque di concetti fondamentali da conoscere, conviene prestarvi attenzione.

Introduzione

Abbiamo già parlato del fatto che esiste una realtà fisica che può interessarci, per vari motivi: potrebbe essere necessario introdurre, per sistemi di vario tipo (elettrici, meccanici, pneumatici, idrici, termici...) una modellistica matematica, più o meno semplice.

A seconda delle caratteristiche del sistema, la modellistica risultante potrà essere di vari tipi: se il sistema è dinamico, il modello sarà esprimibile in termini di equazioni differenziali, se è statico in termini di semplici equazioni algebriche; queste equazioni possono essere dotate di determinate caratteristiche, che dipenderanno a loro volta dalle caratteristiche del sistema del quale si introduce la modellizzazione.

La caratteristica più antipatica di un modello sotto il punto di vista della persona che deve costruirlo e/o utilizzarlo, è il fatto che un modello è una rappresentazione di determinate fenomenologie, a meno di alcuni limiti di applicabilità; un buon sistemista deve essere in grado di saper utilizzare il modello più idoneo per le situazioni più idonee, senza tuttavia utilizzarlo fuori dal suo insieme di applicazioni possibili; la legge di Newton ad esempio modella il comportamento delle forze applicabili ad un corpo, ma solo una

volta verificate alcune condizioni fondamentali, quale ad esempio il fatto di trovarsi in un sistema inerziale: se il sistema fosse accelerato, la legge di Newton non potrebbe, da sola, caratterizzare i comportamenti delle forze in un sistema, senza introdurre alcuni termini correttivi.

Di una realtà fisica è possibile introdurre una modellistica fine e accurata, o una più grossolana, in grado di descrivere solo pochi aspetti dell'intero comportamento del sistema. Dove deve arrivare il grado di accuratezza di un modello? Beh, semplice: dipende dall'uso che si intende fare di esso.

Si consideri un esempio molto semplice, perfetto per lo studio dei sistemi elettrici/elettronici: si consideri il seguente resistore:

La legge che notoriamente lega gli andamenti di tensione e corrente in un dispositivo di questo tipo, è la legge di Ohm:

$$v(t) = R \cdot i(t)$$

La legge di Ohm è un modello, ossia una formulazione matematica del comportamento della grandezza “resistenza”. Si è parlato di “resistore”, ossia del componente fisico che più comunemente è usato per simulare un comportamento resistivo. La domanda che ci si potrebbe porre è: la legge di Ohm è sempre rispettata? la risposta è no: a seconda ad esempio della frequenza del segnale di tensione ai capi del resistore, la legge di Ohm potrà essere più o meno rispettata: aumentando eccessivamente la frequenza, ad esempio, la lunghezza d'onda del segnale potrebbe diventare non trascurabile rispetto alla lunghezza fisica del componente, rendendo di fatto il resistore equivalente a una linea di trasmissione, introducendo elementi induttivi o capacitivi ai suoi capi; in tal caso, fissata una certa frequenza, è possibile modellizzare in vari modi un resistore, ad esempio mediante una resistenza in serie a un'induttanza, o una resistenza in parallelo a una capacità, o in altri modi.

Quale modello si deve dunque usare del resistore? Beh, dipende dall'uso che si deve farne, in questo caso a seconda della frequenza di lavoro. Un modello troppo complicato e fine solitamente è inutile: per la continua, nel caso del resistore, utilizzare capacità o induttanze parassite contenute nel dispositivo non ha senso, dal momento che esse non avranno assolutamente influenza, dunque la legge di Ohm è un ottimo modello; d'altra parte utilizzando modelli fini si complicherebbero enormemente i calcoli, che dunque, da fare “manualmente”, diverrebbero pressochè impossibili da risolvere. Utilizzando software di simulazione di sistemi, quale ad esempio Simulink, si può utilizzare un modello qualsiasi, dal momento che i calcoli vengono risolti dal calcolatore, e non dal suo operatore.

2.1 Terminologia e concetti preliminari

Esaminiamo le principali classificazioni effettuabili per sistemi di vario genere, in modo da introdurre i concetti fondamentali per la comprensione dei concetti successivi:

- Precedentemente abbiamo imparato a distinguere sistemi a tempo continuo da quelli a tempo discreto, differenziandoli per il dominio nel quale variano i parametri temporali (in un sistema a tempo continuo la variazione è in \mathbb{R} , in un sistema a tempo discreto in \mathbb{N}); ciò che il calcolatore fa, molto spesso, è discretizzare il problema continuo (utilizzando ad esempio un sistema di campionamento). Si noti che esistono anche sistemi misti, ossia parzialmente a tempo continuo e parzialmente a tempo discreto. Un processo di discretizzazione avviene ad esempio nel caso dell'uso di processori: quando si progetta un controllore digitale, ossia un sistema di controllo basato sull'uso di processori, il problema viene discretizzato, e trattato come problema discreto da calcolatori: un processore effettua il controllo a tempo discreto.
- Altro criterio di classificazione è quello che li suddivide in sistemi con o senza memoria, ossia sistemi statici o dinamici. Restando nell'ambito dei sistemi elettrici, un circuito puramente resistivo è un esempio di circuito senza memoria, ossia di sistema statico; dualmente, un circuito contenente elementi reattivi, quali induttori o condensatori, è un circuito con memoria.
- Un criterio di classificazione molto importante riguarda la causalità del sistema: un sistema è causale se l'uscita dipende esclusivamente da tempi passati o presenti, non da tempi futuri dell'ingresso. Quest'osservazione potrebbe sembrare molto bizzarra, ma di fatto bisogna prestarvi notevole attenzione: al momento del progetto, il progettista potrebbe avere la tentazione di realizzare, senza volerlo, un sistema non causale; ciò non è fisicamente realizzabile, dunque, se anche il progetto dovesse portare a un qualche risultato, questo potrebbe essere presumibilmente privo di significato fisico;
- Stazionarietà: se le caratteristiche del sistema sono costanti al trascorrere del tempo, il sistema è detto "stazionario", o "tempo-invariante". Un sistema elettronico reale ad esempio invecchia nel tempo, dunque le sue modalità di funzionamento cambiano, e così cambiano anche le sue caratteristiche. Si dice per questo motivo che il sistema sia non-stazionario. Spesso la stazionarietà non si studia dall'inizio del tempo

fino a oggi, bensì in intervalli di tempo interessanti: mesi, anni. Se il sistema nell'arco di alcuni anni non cambia le proprie caratteristiche, si può dire, in prima approssimazione, che sia stazionario;

- Linearità: il mondo reale non è quasi mai lineare; i conseguenti modelli dei sistemi, dunque, saranno non lineari. Nell'arco dei secoli sono tuttavia state sviluppate tecniche di analisi e progetto soprattutto nell'ambito di sistemi lineari, dunque quello che solitamente si fa è descrivere, mediante ad esempio processi di linearizzazione, un sistema non lineare come lineare nell'intorno di un certo punto di funzionamento. Si sappia che negli ultimi anni si sta sviluppando la tendenza di introdurre tecniche di progetto anche per quanto riguarda sistemi non lineari, a seconda del grado di non linearità del sistema (un sistema esponenziale di sicuro non è semplice da studiare quanto un sistema parabolico o cubico).

Il metodo per determinare il fatto che un sistema sia lineare o meno è basato sullo studio della sovrapposizione degli effetti: un sistema è lineare se e solo se vi è applicabile il principio di sovrapposizione degli effetti.

2.2 Rappresentazione di sistemi dinamici mediante relazioni ingresso-stato-uscita

Come si sa dalla teoria dei sistemi, è possibile dare, per un sistema una rappresentazione del tipo:

$$\begin{cases} \dot{x} = f(x; u) \\ y = g(x; u) \end{cases}$$

Dove:

$$x \in \mathbb{R}^n; \quad u \in \mathbb{R}^m; \quad y \in \mathbb{R}^p$$

Un generico sistema non lineare, dunque, si può linearizzare, e se ne può estrarre una rappresentazione lineare di questo tipo:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

Si propone a questo punto un esercizio particolare, molto utile soprattutto in vista dell'uso di Simulink. L'esercizio che si intende fare, nella fattispecie, è dare una rappresentazione grafica del sistema, mediante uno schema a

blocchi, proprio come Simulink, ossia il simulatore del comportamento di sistemi dinamici che si utilizzerà come riferimento per la trattazione.

Si presenta lo schema a blocchi delle equazioni di stato in forma matriciale appena scritte:

A partire dall'unico ingresso, u , esso viene moltiplicato per la matrice B ; ciò confluisce in un nodo di somma, in cui arriverà un altro segnale, Ax , ancora da ricavare; sommati Ax e Bu , al fine di ottenere x da \dot{x} , è necessario integrare, considerando anche le condizioni iniziali; in uscita dal sistema di integrazione si avrà dunque lo stato x , che, al fine di essere introdotto nel precedente nodo sommatore, andrà "reazionato" dopo essere stato moltiplicato per la matrice A . Al fine di ottenere l'uscita, non manca molto: è necessario moltiplicare l'ingresso, preso per un altro ramo, per la matrice D , e sommarvi x , preso dall'integratore, dopo avervi moltiplicato C .

Quella appena ottenuta è una rappresentazione grafica delle equazioni di stato-uscita.

Esiste un modo duale di procedere: come si sa, le soluzioni delle equazioni di ingresso-stato-uscita possono essere ricavate sia nel dominio del tempo sia nel dominio della trasformata di Laplace s ; il passaggio successivo, dunque, potrebbe essere quello di trasformare secondo Laplace le equazioni, ottenendo:

$$sX(s) - x(0) = AX(s) + BU(s)$$

$$Y(s) = CX(s) + DU(s)$$

Questa si può riscrivere come:

$$X(s) = (sI)^{-1} \cdot [AX(s) + BU(s) + x(0)]$$

Questo tipo di soluzione, alla quale forse non si è abituati, è comoda per questo motivo: lo schema a blocchi risultante è il seguente:

Trattando così l'equazione, per quanto sia un sistema inusuale, il risultato finale è del tutto uguale al precedente!

Simulink propone diversi metodi, a partire da diversi "blocchi" che ha a disposizione; si propongono dunque le principali soluzioni che esso può proporre, a partire dalla conoscenza delle equazioni caratterizzanti il sistema:

1. Esiste un blocco che, definite in MATLAB le matrici A, B, C, D , vengono riconosciute da Simulink il quale, a fronte di un particolare ingresso, può simulare il sistema. Questo tipo di metodo è molto semplice, ma molto poco flessibile: si può ottenere soltanto l'uscita del sistema, in funzione dell'ingresso;

2. Si può simulare disegnando uno dei due schemi a blocchi proposti, dove i coefficienti moltiplicativi vengono presi dalle matrici definite nel workspace di MATLAB; si noti che, in Simulink, i guadagni vengono modellizzati mediante il blocco “Gain”; questo metodo di lavoro è molto più flessibile del precedente, dal momento che è possibile recuperare qualsiasi variabile interessata;
3. Si può considerare uno schema a blocchi ricavato mediante la scrittura diretta delle equazioni di stato e di uscita “una a una”, non considerando dunque più le matrici e la notazione ad essa collegate, bensì ciascuna funzione ed equazione, una per volta. Questo metodo è il più scomodo e complicato, dal momento che richiede più lavoro rispetto ai casi precedenti, ma è anche quello che riesce meglio a visualizzare il sistema relazionandolo con il sistema reale, rendendolo molto semplice da interpretare.

2.3 Esempio “limite”

Dato il seguente risonatore RLC, ricavarne lo schema a blocchi nella terza versione:

Innanzitutto, si scrivono le equazioni caratteristiche dei componenti:

$$v_L = L \frac{di_L}{dt}; \quad i_C = C \frac{dv_C}{dt}; \quad v_R = R \cdot i$$

Si scrivono dunque le equazioni costitutive del circuito:

$$u(t) = v_L + v_R + v_C$$

Si identificano le variabili di stato come:

$$x_1 = v_C; \quad x_2 = i_L$$

Si può a questo punto scrivere le equazioni di stato come:

$$\begin{cases} \dot{x} = \frac{1}{C}x_2 \\ \dot{x}_2 = \frac{1}{L}[u - x_1 - Rx_2] \\ y = Rx_2 \end{cases}$$

Da qui si potrebbe o estrapolare le quattro matrici, o disegnare direttamente uno schema a blocchi a partire da queste equazioni. Il risultato finale sarebbe semplicemente il seguente:

Si noti che in realtà il significato matematico di questo schema a blocchi è differente rispetto a quello presentato in ambito della rappresentazione grafica

delle equazioni di stato: precedentemente si gestivano contemporaneamente tutte le integrazioni, ora esse vengono gestite una a una manualmente.

Un cenno a una possibile applicazione futura: un risultato che potrebbe interessarci, per un sistema, è la funzione di trasferimento. Come si potrebbe procedere, al fine di calcolarla? Beh, ora come ora, si possono suggerire due vie:

1. Ricavare analiticamente o mediante MATLAB la funzione:

$$H(s) = C(sI - A)^{-1}B + D$$

Dove la matrice $(sI - A)^{-1}$ si può calcolare come:

$$(sI - A)^{-1} = \frac{1}{\det(sI - A)} \text{adj} \{sI - A\}$$

2. Analizzare lo schema a blocchi precedentemente ricavato, semplificarlo mediante teoremi dell'algebra degli schemi a blocchi, e ricavare in maniera semplice e intuitiva, per sistemi anche molto complicati, l'espressione della funzione di trasferimento.

2.4 Il movimento dello stato e dell'uscita nei sistemi LTI

Si possono sostanzialmente utilizzare due tipi di approcci, al fine di determinare espressioni in forma chiusa del movimento dello stato e dell'uscita per sistemi LTI:

- Utilizzare la formula di Lagrange, ossia un approccio basato rigorosamente su di una risoluzione nel dominio del tempo. A scopi didattici è molto interessante poichè fornisce un metodo di risoluzione analitico di equazioni differenziali, senza far intervenire altri operatori. Essa dice che la soluzione di una generica equazione differenziale del primo ordine ha la seguente forma:

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau$$

Dove t_0 è il tempo di accensione del sistema, ossia il tempo a partire dal quale il sistema inizia a fornire una risposta all'ingresso. A e B sono gli operatori di stato precedentemente introdotti, parlando di sistemi linearizzati.

- Uso della trasformata di Laplace: dato il sistema di partenza lo si trasforma, come precedentemente fatto, nel suo equivalente nel dominio della variabile complessa s ; in questo dominio vengono effettuate le varie operazioni di tipo algebrico, quindi si antitrasforma nel dominio del tempo. Nella stragrande maggior parte dei casi, questo è l'approccio più consigliato da seguire.

2.5 Brevi richiami di algebra lineare

In questa sezione verranno riprese alcune nozioni fondamentali per lo studio della teoria dei controlli automatici.

Data la ben nota rappresentazione ingresso-stato-uscita:

$$\dot{x} = Ax + Bu$$

$$y = Cx + Du$$

Si vuole ricordare il fatto che il determinante della matrice $(sI - A)$ (fondamentale per il calcolo della matrice inversa, come visto precedentemente) è un polinomio di grado n monico, dove n è il numero degli stati presenti nel sistema. Dal momento che è monico, il coefficiente della potenza di grado più alto sarà sempre unitario.

La seguente equazione:

$$\det(sI - A) = 0$$

Ammette soluzioni eventualmente nel campo complesso \mathbb{C} , eventualmente anche non distinte tra loro (ovviamente se esiste una soluzione complessa deve esistere anche la sua complessa coniugata). Alcune nomenclature:

- La matrice $(sI - A)$ è detta “matrice caratteristica”;
- Il determinante della matrice caratteristica è detto “polinomio caratteristico”, e sarà indicato con $d(s)$;
- Le radici del polinomio caratteristico sono dette “autovalori” per la matrice A .

Il polinomio caratteristico $d(s)$ si può fattorizzare nel seguente modo:

$$d(s) = (s - s_1)^{n_1} \cdot (s - s_2)^{n_2} \cdot \dots \cdot (s - s_r)^{n_r}$$

Dove $n_1 + n_2 + \dots + n_r = n$, ossia è pari al numero degli stati del sistema.

Come $(sI - A)$ è una matrice, anche la sua inversa, $(sI - A)^{-1}$ lo sarà; essa, come già accennato, si può costruire mediante la formula di inversione basata sul calcolo del reciproco del determinante di $(sI - A)$, moltiplicante l'aggiunta della medesima. Quello che si avrà nella pratica è dunque il reciproco di un polinomio, che moltiplicherà una matrice di funzioni; ciascuna funzione, in via definitiva, sarà una funzione razionale, strettamente proprie (supponendo che il sistema sia, come è ovvio, fisicamente realizzabile).

Può capitare un fatto estremamente interessante: se tutti gli elementi della matrice aggiunta hanno una o più radici in comune con il polinomio caratteristico, si introducono cancellazioni tra gli elementi interni alla matrice aggiunta e gli elementi del polinomio caratteristico. Si noti che questo fatto fornisce informazioni globali sul sistema solo se le cancellazioni avvengono per tutti gli elementi della matrice: in questo modo è possibile semplificare per ogni funzione razionale una radice del denominatore. Si può avere, per ciascuna funzione, qualcosa di questo genere:

$$(sI - A)^{-1} = \frac{Q_i(s) \cdot \gamma(s)}{\gamma(s)M(s)}$$

Dove $Q_i(s)$ è il polinomio che differenzia ciascuna funzione, $\gamma(s)$ è una parte del polinomio comune al denominatore.

Una volta effettuata questa semplificazione, se essa vale per tutti gli elementi della matrice (o quantomeno quelli interessati, se si utilizza solo parte degli ingressi), il polinomio caratteristico diviene il “polinomio minimo”.

2.5.1 Esempio pratico

Si consideri il seguente esempio, atto a migliorare la comprensione della spiegazione appena fornita; dato il sistema caratterizzato dalla seguente matrice di stato:

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

Si determini la matrice $(sI - A)^{-1}$.

Data A , si calcola $(sI - A)$ come:

$$(sI - A) = \begin{bmatrix} s - 2 & 0 \\ 0 & s - 2 \end{bmatrix}$$

La matrice $(sI - A)^{-1}$ si calcola quindi come:

$$(sI - A)^{-1} = \frac{1}{(s - 2) \cdot (s - 2)} \cdot \begin{bmatrix} s - 2 & 0 \\ 0 & s - 2 \end{bmatrix} = \frac{1}{s - 2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Vi è stata una cancellazione, che ha portato a una modifica del polinomio caratteristico della matrice da $(s - 2)^2$ a $(s - 2)$. Per questo, il polinomio da considerare sarà il polinomio minimo, e non quello caratteristico, per alcuni criteri, come si vedrà tra breve.

2.6 La stabilità interna

Si parla della stabilità interna ogni qual volta si studi la stabilità definita a partire dal movimento libero dello stato. Nella fattispecie, si fornisce la seguente definizione di stabilità interna:

- Un sistema LTI è internamente stabile se e solo se il movimento libero dello stato è limitato.
- Un sistema LTI è asintoticamente stabile se e solo se il movimento libero dello stato tende a convergere a 0, per $t \rightarrow \infty$.
- Un sistema LTI è instabile se e solo se il movimento libero dello stato non è limitato.

Queste definizioni qualitative sono fondamentali al fine di comprendere il significato vero e proprio di stabilità: se un sistema ha un movimento dello stato non convergente a zero ma limitato tra alcuni valori esso è “semplicemente stabile”, o “stabile”. Se si aggiunge la convergenza a zero del movimento dello stato, il sistema è detto “asintoticamente stabile”. Un sistema asintoticamente stabile è stabile, poichè è limitato, ma un sistema stabile non lo è per forza anche asintoticamente.

Da un punto di vista pratico, esistono alcune condizioni, basate sullo studio delle caratteristiche degli autovalori della matrice A , in grado di determinare immediatamente le proprietà di stabilità interna del sistema.

- Se tutti gli autovalori hanno parte reale strettamente minore di zero, ossia se:

$$\operatorname{Re} \{ \lambda_i \} < 0 \quad \forall i$$

Il sistema è asintoticamente stabile.

- Il sistema è semplicemente stabile se e solo se tutti gli autovalori di A hanno parte reale non positiva, e quelli con parte reale nulla sono radici semplici del polinomio minimo.

- Il sistema è instabile se esiste almeno un autovalore di A con parte reale positiva, o a parte nulla radice non semplice del polinomio minimo.

Si noti che lo studio della stabilità va fatto sul polinomio minimo, non assolutamente sul polinomio caratteristico: considerare il polinomio minimo permette di considerare, anziché la molteplicità algebrica degli autovalori, la molteplicità geometrica, legata alla dimensione dell'autospazio relativo all'autovalore.

2.6.1 Esempio pratico

Data la matrice A :

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -5 \end{bmatrix}$$

Determinare le proprietà di stabilità del sistema.

Si calcola $(sI - A)$:

$$(sI - A) = \begin{bmatrix} s & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & s + 5 \end{bmatrix}$$

Da qui, si calcola $(sI - A)^{-1}$:

$$(sI - A)^{-1} = \frac{1}{s^2 \cdot (s + 5)} \cdot \begin{bmatrix} s(s + 5) & 0 & 0 \\ 0 & s(s + 5) & s \\ 0 & 0 & s^2 \end{bmatrix}$$

La matrice è triangolare, dunque gli autovalori coincidono con i termini presenti sulla colonna principale della matrice. Come si può vedere, ciascun elemento della matrice è moltiplicato per s , dunque è possibile effettuare una cancellazione tra tutti i componenti della matrice e una s al denominatore, ottenendo:

$$(sI - A)^{-1} = \frac{1}{s \cdot (s + 5)} \cdot \begin{bmatrix} s + 5 & 0 & 0 \\ 0 & s + 5 & 1 \\ 0 & 0 & s \end{bmatrix}$$

Il sistema non è instabile, dal momento che vi è una cancellazione tale da proporre un polinomio minimo diverso da quello caratteristico, che nella fattispecie elimina la molteplicità doppia nella radice in zero, rendendo il sistema semplicemente stabile.

2.7 Scelta dei parametri di simulazione per Simulink

Al fine di essere in grado di simulare un sistema, con Simulink, è buona cosa sapere alcune cose riguardo esso. Simulink è un tool in grado di simulare un sistema mediante l'uso di metodi numerici per l'integrazione di equazioni differenziali. Questi metodi, implementati in alcune funzioni di MATLAB, devono essere impostati nella maniera corretta, in modo da evitare di ottenere effetti sgradevoli al momento della visualizzazione dell'uscita del sistema.

I parametri fondamentali da conoscere sono sostanzialmente due:

- Passo di integrazione: come già detto, i metodi utilizzati per l'integrazione delle equazioni differenziali sono numerici; a seconda del passo di integrazione, si troveranno funzioni più o meno simili a funzioni che si vedrebbero in un sistema a tempo continuo. Le cose da soddisfare per ottenere una buona simulazione sono sostanzialmente le seguenti: visualizzare l'uscita a regime, in modo da farla sembrare continua, ossia da non far sembrare "spezzate" curve che dovrebbero essere molto lisce. Un effetto sgradevole potrebbe ad esempio essere il seguente:

Al fine di evitare un esempio di questo tipo, bisogna impostare un passo di integrazione in modo da rendere trascurabili le lunghezze delle spezzate sullo schermo. Per rendere trascurabili le lunghezze delle spezzate, un'idea può essere quella di scegliere, come passo di integrazione, qualcosa che sia molto più piccolo del periodo della sinusoide (nel caso si abbia ad esempio un segnale sinusoidale); dato ad esempio un segnale sinusoidale, dunque, a partire dalla pulsazione ω se ne calcola la frequenza ν , normalizzando ω per 2π ; di ciò si calcola il reciproco, ottenendo il periodo della sinusoide. Si può scegliere di avere un passo di integrazione 20, 50, 100, 150 volte minore del periodo della sinusoide. Vi sono due indicazioni particolari a questo punto: un tempo di integrazione troppo lungo provoca l'effetto visivo sgradito appena visualizzato; un tempo di integrazione troppo corto, dove ossia si considerano istanti temporali discreti troppo vicini tra loro, provoca un grosso aumento della complessità computazionale della soluzione dell'equazione differenziale, aumento che di fatto non fornisce vantaggi. Buoni compromessi sono tempi di integrazione dalle 50 alle 150 volte minori rispetto al periodo della sinusoide.

- Potrebbe capitare qualcosa del genere:

In questo caso l'errore è più banale, richiede spiegazioni più semplici: la durata di simulazione è troppo lunga rispetto al periodo della sinusoide

finale in uscita; per questo motivo, capita che vengono visualizzate, con pessima qualità, molte sinusoidi, rendendo la simulazione incomprensibile. Al fine di risolvere il problema, è sufficiente cercare una periodicità nell'uscita, e usare come tempo di simulazione un tempo pari a quattro o cinque volte il periodo ricavato.

- Considerazione aggiuntiva può essere fatta riguardo la durata del transitorio, al fine di aggiungere condizioni per quanto concerne il tempo di simulazione: al fine di vedere l'uscita del sistema a regime, una buona idea può essere quella di considerare, dato un sistema approssimabile a un sistema del primo ordine, la costante di tempo τ del sistema, definita come:

$$\tau = \left| \frac{1}{\operatorname{Re}\{\lambda_i\}} \right|$$

Dove λ_i si può intendere sia come autovalore della matrice A (che si ridurrà a uno scalare), sia come posizione del polo della funzione di trasferimento del sistema.

In teoria dopo un $t = 3\tau$ si dovrebbe avere una garanzia di rientrare nei margini del 95 % rispetto all'uscita a regime; introducendo un periodo pari a $t = 5\tau$, si hanno garanzie di convergenza ancora maggiori.

Nel caso il sistema fosse del secondo ordine, a poli complessi, l'espressione generica della funzione di trasferimento $H(s)$ sarebbe:

$$H(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

Quali criteri si usano in questo caso? Beh, si può da un lato sfruttare τ , per quanto essa venga comunemente definita solo in ambito di sistemi a poli reali, dall'altro osservare la pulsazione naturale ω_n del sistema, e capire che il tempo di simulazione può essere ad esempio pari a quattro o cinque oscillazioni di quella frequenza.

Una domanda finale potrebbe essere: quale metodo numerico utilizzare? Vengono proposte diverse "ode", tra i vari metodi possibili nel menù a tendina di Simulink. Esiste un'ulteriore differenziazione, ossia il fatto che il passo sia variabile o costante: nel caso di passo costante, tutte le indicazioni finora introdotte sono valide: si sceglie, mediante i criteri precedentemente introdotti, un passo di integrazione definito dall'utente, quindi si lancia la simulazione. Un buon metodo di integrazione potrebbe essere il "Runge-Kutta".

Una classe di metodi numerici molto indicati per segnali a costanti di tempo molto variabili è quella dei cosiddetti “metodi stiff”; Simulink ne propone due, le “ode15s” e “ode23s”. Queste sono implementazioni che richiedono l’uso di un passo variabile di integrazione, definito dal software al momento del calcolo della soluzione; ciò che si può fare, se si desidera una certa qualità, è impostare (tra i parametri) un passo di simulazione minimo e uno massimo, in modo da pilotare il calcolatore a far convergere la soluzione solo a patto che siano rispettati questi vincoli, in modo da ottenere una buona soluzione.

2.8 Rappresentazioni di sistemi dinamici mediante relazioni ingresso-uscita: funzione di trasferimento

Precedentemente è stato analizzato un modello basato sull’uso di variabili di stato per quanto riguarda generici sistemi dinamici; si cambia obiettivo a questo punto, introducendo una rappresentazione basata sulla relazione diretta tra ingresso e uscita di un sistema: la funzione di trasferimento.

Al fine di introdurre tutti gli elementi del caso, si parte da una definizione preliminare.

Definizione: matrice di trasferimento

Per matrice di trasferimento di un sistema LTI definito dalle matrici A , B , C , D , si intende la trasformata di Laplace (che verrà anche definita come \mathcal{L} -trasformata) della matrice $h(t)$ delle risposte agli impulsi.

Teorema

La matrice di trasferimento $H(s)$ di un generico sistema LTI (A , B , C , D) è data da:

$$H(s) = C(sI - A)^{-1}B + D$$

Questa espressione si può ricavare semplicemente considerando le rappresentazioni in variabili di stato del sistema, trasformandole mediante la trasformata di Laplace ed evidenziando i movimenti liberi e forzati per stato e uscita. Per ricavare $H(s)$ è necessario considerare il solo movimento forzato dell’uscita, eccitando il sistema con un impulso, la cui trasformata di Laplace è notoriamente la costante 1.

Alcune osservazioni riguardo la matrice sulla quale si basa il calcolo della funzione/matrice di trasferimento: $(sI - A)^{-1}$; se $D = 0$, la funzione razionale fratta è strettamente propria; in generale, dato un elemento $H_{i,j}(s)$ della matrice di trasferimento, esso ha una forma del tipo:

$$H_{i,j}(s) = \frac{b_n s^n + b_{n-1} s^{n-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}$$

Se il sistema è proprio, dunque $D = 0$, b_n sarà uguale a 0.

D'ora in avanti, per comodità notazionale, si considererà una matrice di trasferimento costituita da un solo elemento; invece di un generico $H_{i,j}(s)$, si indicherà questa con $H(s)$.

Per una generica funzione di trasferimento si introducono alcune nozioni particolari:

- Per “grado relativo” della funzione di trasferimento si definisce la differenza di grado da denominatore a numeratore;
- $s = z_i$ è uno “zero” della funzione di trasferimento $H(s)$ se:

$$H(z_i) = 0$$

Ossia se è un valore della variabile complessa di Laplace s tale per cui la funzione di trasferimento va a 0.

- $s = p_i$ è un “polo” per la funzione di trasferimento $H(s)$ se:

$$H(p_i) \longrightarrow \infty$$

Ossia se è un valore della variabile s tale per cui la funzione di trasferimento tende ad acquisire valori elevatissimi, modellizzabili come infinito.

2.8.1 Esempio pratico

Si consideri la seguente funzione di trasferimento:

$$H(s) = \frac{N_H(s)}{D_H(s)} = 2 \frac{s^3 + 3s^2 - s - 3}{(s-1)(s+2)(s+1)}$$

Si può dire a priori quali siano i poli della funzione di trasferimento? La risposta è no: bisognerebbe prima controllare il fatto che non esistano cancellazioni zeri-poli nella funzione di trasferimento; se le radici del denominatore

(in forma “comoda”, poichè fattorizzate) sono anche radici del numeratore, si dovrebbero elidere, eliminando di fatto un polo. Nel caso di $s = -2$ ad esempio si può stare tranquilli, poichè sostituendo -2 al numeratore esso non si annulla. -1 purtroppo è un caso diverso: esso annulla sia numeratore sia denominatore, dunque non è un polo.

Osservazione aggiuntiva: se per $s \rightarrow \infty$ la funzione tende a 0, si suol dire che essa abbia “degli zeri all’infinito”. Questo perchè il fatto che per valori estremamente alti, modellizzabili mediante infinito, della variabile s , la funzione tende ad annullarsi.

La funzione di trasferimento rappresenta un sistema dinamico mediante una relazione ingresso-uscita. Lo stato, come si può vedere, non fa parte di questo tipo di rappresentazione. Dal momento che, inoltre, la funzione di trasferimento rappresenta una modellizzazione del solo movimento forzato dell’uscita, si può dire che, al momento di calcolarla, le condizioni iniziali vadano sempre considerate nulle.

2.9 La stabilità esterna nei sistemi LTI

Si è precedentemente parlato di stabilità interna, ossia stabilità del movimento libero dello stato. Di fatto, quando si parla di stabilità, esistono altre definizioni, nella fattispecie quella di “stabilità esterna” o “BIBO-stabilità” (BIBO : Bounded Input, Bounded Output).

Si consideri la seguente definizione.

Definizione

Un sistema LTI (A, B, C, D) si dice “esternamente stabile” se il movimento forzato dell’uscita rimane limitato in presenza di tutti i possibili ingressi limitati.

Questa definizione è interessante, ma non può essere utilizzata al fine di studiare la stabilità di un sistema: non è assolutamente possibile provare per un sistema tutti i possibili ingressi limitati, per poi capire che essi “funzionino sempre”. Può essere al contrario più utile il contrario, ossia verificare che, per un dato ingresso limitato, il sistema restituisca un’uscita non limitata, dimostrando di essere di fatto non BIBO-stabile.

2.9.1 Esempio pratico

Dato un sistema del tipo:

Ossia un banale integratore, il sistema è stabile o instabile? Beh, si provi a introdurre ad esempio una costante in ingresso, tipico segnale limitato. L'uscita del sistema è l'integrale della costante, ossia una retta, segnale notoriamente illimitato.

2.9.2 Esempio pratico

Si consideri il seguente circuito:

Questo sistema è esternamente stabile o instabile? Beh, si sa che il regolatore reale ha una curva del tipo:

Se il risonatore è tuttavia ideale, il Q , ossia il valore dell'ordinata in prossimità della pulsazione di risonanza, è infinito; utilizzando dunque come ingresso una sinusoide (segnale notoriamente limitato) vibrante alla pulsazione di risonanza, l'uscita tenderà a divergere.

2.9.3 Alcuni risultati aggiuntivi

Un sistema LTI (A, B, C, D) è stabile esternamente se e solo se la sua parte raggiungibile e osservabile è asintoticamente stabile, ossia se tutti gli autovalori hanno parte reale negativa.

Teorema

Una rappresentazione LTI (A, B, C, D) è esternamente stabile se e solo se tutti i poli della funzione di trasferimento $H(s)$ hanno parte reale negativa. Esempio semplice e già visto, è l'integratore; la funzione di trasferimento dell'integratore ha una forma del tipo:

$$H(s) = \frac{1}{s}$$

Dal momento che l'unico polo è $s = 0$, il sistema è esternamente instabile. Nel caso del risonatore:

$$H(s) = \frac{sL}{1 + sL^2C}$$

Come si può vedere, si ha uno zero nell'origine, un polo negativo, un polo positivo. Quest'ultimo causa di fatto l'instabilità del sistema.

Stabilità asintotica interna implica stabilità esterna: l'insieme dei poli e quello degli autovalori non coincide, a meno che il sistema non sia completamente osservabile e raggiungibile, dunque potrebbero esservi cancellazioni zeri-poli.

2.9.4 Esempio pratico

Si propone a questo punto un esempio pratico molto corposo, in grado di mostrare come il perfezionamento di un modello matematico di un sistema possa cambiarne completamente le caratteristiche.

Si consideri il seguente circuito basato sull'uso di un amplificatore operazionale:

Per chi ha studiato i corsi di elettronica, questo è un tipico caso di amplificatore non invertente. Dall'elettronica, si sa che se il guadagno dell'operazionale è estremamente elevato, si può dire che:

$$\frac{v_u}{v_e} = 1 + \frac{R_2}{R_1}$$

Ora ci si potrebbe porre la domanda da controllisti: questo oggetto è stabile? Beh, di fatto, ora come ora, questo sistema non è neanche dinamico: sembra che il guadagno sia costante per qualsiasi frequenza.

Quest'affermazione non è vera: il modello finora utilizzato è una prima approssimazione del sistema reale. In corrente continua, volendo, esso potrebbe funzionare, in banda audio non è detto, a radiofrequenza di sicuro no. L'ipotesi che limita estremamente il range di validità del blocco, è quella secondo cui $A \rightarrow \infty$ sia sempre verificata.

Chi conosce l'elettronica, sa che nei datasheet, tra gli altri parametri, se ne forniscono alcuni riguardanti il comportamento in frequenza di $|A|$:

È fondamentale prestare attenzione ai limiti di validità del modello: A_0 potrebbe valere 10^5 , 10^6 , 10^7 , ma solo per una banda abbastanza ristretta, con un polo in un intorno dei 10 Hz; il secondo polo è di solito nell'ordine dei MHz. Si può dire che il guadagno a catena aperta dell'amplificatore operazionale, con un'approssimazione migliore, abbia un'espressione del tipo:

$$A(s) = \frac{A_0}{\left(1 + \frac{s}{\omega_1}\right) \left(1 + \frac{s}{\omega_2}\right)}$$

Un legame ingresso-uscita più dettagliato potrebbe dunque essere simile a questo:

v_e si ricava a partire da v_u , moltiplicando mediante il fattore di partizione di tensione:

$$v_e = V^+ - V^-; \quad V^+ = v_e; \quad V^- = v_u \cdot \frac{R_1}{R_1 + R_2}$$

Si può dire dunque che:

$$v_u = A(s)v_e$$

Dunque, $H(s)$ avrà una forma del tipo:

$$H(s) = \frac{v_u(s)}{v_e(s)}$$

Si apre una parentesi per una regola generale, che poi verrà meglio analizzata in seguito: data, in forma generale, una rete del tipo:

Un risultato semplice afferma che la funzione di trasferimento risultate è:

$$G_{ry}(s) = \frac{Y(s)}{P(s)} = \frac{G(s)}{1 \pm G(s)H(s)}$$

Ossia, al numeratore si introduce la funzione di trasferimento sul ramo che collega direttamente ingresso e uscita, mentre al denominatore, a seconda del fatto che la reazione sia positiva o negativa (entrambi + o + e - al sommatore), la “funzione di anello”:

$$G_a(s) = G(s)H(s)$$

Dove, in senso generale, per funzione di anello si intende quella funzione ottenuta come prodotto di tutto ciò che sta tra il nodo E e il nodo F .

Nel caso specifico di questo sistema:

$$H(s) = \frac{A(s)}{1 + A(s)\frac{R_1}{R_1+R_2}}$$

Si noti un fatto interessante: se $A(s) \rightarrow \infty$, si ritrova il guadagno precedente, ossia:

$$H(s) = H = 1 + \frac{R_2}{R_1}$$

Cosa si può dire a questo punto sulla stabilità esterna di questo sistema? Sostituendo $A(s)$ nell'espressione generale della funzione di trasferimento, si troveranno polinomi di ordine non superiore al secondo, come si può osservare “a occhio”. Si può utilizzare la regola di Cartesio per determinare che la parte reale sia sempre negativa, dal momento che sui polinomi di secondo ordine il segno dei coefficienti fornisce condizioni necessarie e sufficienti. Questo sistema, dunque, è BIBO-stabile (stabile esternamente).

2.10 Forme e parametri della funzione di trasferimento

In questo paragrafo verranno considerate le tre principali forme nelle quali è possibile rappresentare una generica funzione di trasferimento; verranno eventualmente introdotte alcune osservazioni riguardo la loro utilità in diversi ambiti.

Forma rapporto di polinomi

Si considera la prima forma, comunemente detta “rapporto di polinomi”:

$$H(s) = \frac{b_n s^n + b_{n-1} s^{n-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}$$

Alcune osservazioni: un sistema è in “forma minima” se esso è completamente raggiungibile e osservabile. Altro modo di dire ciò, è considerare il fatto che gli insiemi di poli e di autovalori coincidano. Si immagina dunque, data una generica funzione di trasferimento, che essa sia espressa in forma minima.

Fattorizzazione zeri-poli

La fattorizzazione zeri-poli, ossia quella che mette in evidenza le singole radici di numeratore e denominatore della funzione di trasferimento, ha un’espressione di questo tipo:

$$H(s) = K_\infty \frac{\prod_i (s - z_i) \prod_j (s^2 + 2\xi_j \omega_{n,j} s + \omega_{n,j}^2)}{\prod_k (s - p_k) \prod_l (s^2 + 2\xi_l \omega_{n,l} s + \omega_{n,l}^2)}$$

Dove si identificano, per quanto riguarda i poli complessi coniugati:

ω_n come la lunghezza del vettore, ϑ la sua fase, e lo smorzamento ξ come il coseno di ϑ :

$$\xi = \cos(\vartheta)$$

Forma fattorizzata a costanti di tempo

Esiste una terza forma, frequentemente utilizzata soprattutto nell’ambito della rappresentazione mediante diagrammi di Bode, per le funzioni di trasferimento:

$$H(s) = K_{st} \cdot \frac{\prod_i \left(1 - \frac{s}{z_i}\right) \prod_j \left(1 + 2\xi_j \frac{s}{\omega_{n,j}} + \frac{s^2}{\omega_{n,j}^2}\right)}{\prod_k \left(1 - \frac{s}{p_k}\right) \prod_l \left(1 + 2\xi_l \frac{s}{\omega_{n,l}} + \frac{s^2}{\omega_{n,l}^2}\right)}$$

Per radici reali negative, le costanti di tempo sono:

$$\tau_i = -\frac{1}{z_i}$$

$$\tau_k = -\frac{1}{p_k}$$

In linea di principio si parla di costanti di tempo relative a poli e zeri reali e nel semipiano sinistro del dominio di Laplace.

Osservazione fondamentale: a seconda dei gusti e delle necessità, utilizzare liberamente una di queste rappresentazioni; fondamentale è non mischiarle: solo una di esse, a propria discrezione, deve essere usata per volta, mai più di una, onde evitare problemi di vario tipo.

Esempio pratico

Una buona cosa, per quanto riguarda le varie forme nelle quali si può rappresentare una funzione di trasferimento, è essere in grado di passare da una a un'altra. Qua si propone un esempio di come si deve procedere.

Data la funzione di trasferimento in forma “rapporto di polinomi”:

$$H(s) = \frac{10s^4 + 200s^3 + 1400s^2 + 4000s + 3840}{s^6 + 16s^5 + 86s^4 + 176s^3 + 105s^2}$$

Si può ricavare, ad esempio mediante MATLAB, la forma “fattorizzazione zeri-poli”:

$$H(s) = 10 \cdot \frac{(s+2)(s+4)(s+6)(s+8)}{s^2(s+1)(s+3)(s+5)(s+7)}$$

Raccogliendo i coefficienti, si può calcolare semplicemente anche la forma in costanti di tempo:

$$H(s) = 10 \frac{2 \times 4 \times 6 \times 8}{1 \times 3 \times 5 \times 7} \cdot \frac{\left(1 + \frac{s}{2}\right) \left(1 + \frac{s}{4}\right) \left(1 + \frac{s}{6}\right) \left(1 + \frac{s}{8}\right)}{s^2 \left(1 + s\right) \left(1 + \frac{s}{3}\right) \left(1 + \frac{s}{5}\right) \left(1 + \frac{s}{7}\right)}$$

Questo tipo di rappresentazione è la più conveniente, quando si vogliono rappresentare sistemi mediante diagrammi di Bode.

2.10.1 Guadagno stazionario

Il guadagno stazionario è qualcosa che si userà molto frequentemente nel corso della trattazione; per questo motivo, è fondamentale chiarire alcuni concetti che lo riguardano.

Si consideri una specie di “esperimento”: data la funzione di trasferimento in forma “costanti di tempo”, il sistema deve essere BIBO-stabile, ossia non avere poli non negativi). L’esperimento è: dato $H(s)$ esternamente stabile, valutare $Y(s)$ con una $U(s)$ specifica: a gradino. Dato dunque un ingresso limitato come questo, anche l’uscita dovrà essere limitata. Dato nella fattispecie un gradino di ampiezza \bar{u} , tolto un transitorio iniziale, l’uscita dovrà essere limitata su un certo valore \bar{y} . Si intende calcolare questo valore. Considerando dunque:

$$Y(s) = H(s)U(s)$$

Ci si chiede: è possibile utilizzare una proprietà delle trasformate di Laplace, ossia il teorema del valore finale? Beh, essendo il sistema esternamente stabile, tenendo conto che il teorema del valore finale afferma che:

$$\bar{y} = \lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sY(s)$$

Risultato utilizzabile solo se il limite esiste, cosa verificata se il sistema è esternamente stabile, allora si può dire che:

$$\lim_{s \rightarrow 0} sH(s)U(s) = \lim_{s \rightarrow 0} s \frac{\bar{u}}{s} H(s) = \bar{u}H(s) = K_{st} \cdot \bar{u}$$

Questo risultato dice che, in regime permanente (steady state):

$$K_{st} = \frac{\bar{y}}{\bar{u}}$$

Finora è stato tuttavia considerato un caso abbastanza standard, “fortunato”; un caso meno fortunato è quando sono presenti poli nell’origine del dominio di Laplace. In tal caso, è necessario definire un guadagno stazionario generalizzato, dove si dovrà attribuire un significato fisico particolare, differente. Si considera r il numero di poli nell’origine per quanto riguarda il sistema; se $r = 0$, il guadagno stazionario per motivi storici viene chiamato K_p (dal momento che questa teoria veniva solitamente applicata su sistemi meccanici, dunque riguarda nella fattispecie un guadagno di “posizione”). In questo caso:

$$K_p = \lim_{s \rightarrow 0} H(s)$$

Se $r = 1$, si può definire il guadagno stazionario di velocità, K_v , come:

$$K_v = \lim_{s \rightarrow 0} sH(s)$$

Per $r = 2$, si può definire il guadagno stazionario di accelerazione, K_a :

$$K_a = \lim_{s \rightarrow 0} s^2 H(s)$$

2.10.2 Puntualizzazione sull'uso dei teoremi della \mathcal{L} -trasformata

Quando si utilizzano degli strumenti, buona cosa è comprendere quali sono le condizioni nelle quali è effettivamente possibile utilizzarli. Dato un sistema con un determinato ingresso, si vuole calcolarne l'uscita in regime permanente:

Un esempio pratico potrebbe essere:

$$H(s) = \frac{1}{1-s}$$

Dove l'ingresso è un gradino unitario:

$$U(s) = \frac{1}{s}$$

Ha senso calcolare l'uscita a regime permanente, \bar{y} ? Se sì quanto vale?

Beh, il limite, matematicamente, è assolutamente calcolabile; il fatto che abbia significato è un altro conto: il sistema presenta un polo nel semipiano destro, dunque è instabile, dunque per quanto matematicamente calcolabile, esso non ha significato fisico. In questo caso, il limite, di fatto, non esiste.

Capitolo 3

La risposta in frequenza di sistemi LTI

In questo capitolo della trattazione ci si occuperà della risposta in frequenza per sistemi lineari tempo-invarianti. Come si procederà? Beh, si studierà il comportamento di un sistema LTI in presenza di una particolare classe di ingressi: gli ingressi sinusoidali. Il fatto di utilizzare segnali di questo tipo non è limitante in alcun modo: dal momento che il sistema è lineare, vale il principio di sovrapposizione degli effetti; poichè mediante la combinazione lineare di sinusoidi è possibile rappresentare un'enorme classe di segnali, sovrapponendo gli effetti si possono ottenere grandi risultati. Questo significa che, dato un segnale interpretabile come somma di vari segnali, è possibile dare in ingresso la somma di diversi segnali, in uscita avere la somma dei contributi dei segnali di ingresso.

3.1 Esistenza di una uscita sinusoidale

Si incomincia la prima sezione, con un teorema fondamentale.

Teorema

Si supponga che il sistema (A, B, C, D) con funzione di trasferimento $H(s)$ non abbia autovalori in $\pm j\omega_0$, e si applichi l'ingresso $\tilde{u}(t)$:

$$\tilde{u}(t) = U \sin(\omega_0 t + \varphi), \quad t \geq 0$$

Allora esiste uno stato iniziale per cui l'uscita è sinusoidale, e vale:

$$\tilde{y}(t) = Y \cdot \sin(\omega_0 t + \psi), \quad t \geq 0$$

Dove:

$$Y = |H(j\omega_0)| \cdot U$$

$$\psi = \varphi + \arg \{H(j\omega_0)\}$$

Inoltre, se il sistema è asintoticamente stabile, per qualunque stato iniziale risulta:

$$\lim_{t \rightarrow \infty} [y(t) - \tilde{y}(t)] = 0$$

Osservazioni:

1. Si sta eccitando un sistema con un ingresso sinusoidale. Esiste uno stato iniziale per cui l'uscita è immediatamente sinusoidale, ossia è sinusoidale per $t \geq 0$;
2. La pulsazione della sinusoide in uscita dal sistema è la stessa del segnale eccitante.

Queste osservazioni, questi risultati, sono assolutamente fondamentali per le applicazioni: indipendentemente dalla stabilità iniziale, questi risultati sono comunque validi.

3.2 Definizione di risposta in frequenza

Si incomincia subito con una definizione.

Definizione

La funzione complessa $H(j\omega)$ definita come:

$$H(j\omega) = C(j\omega I - A)^{-1}B + D \quad \omega \geq 0$$

Ossia definita per valori non negativi di ω , tali per cui $j\omega$ non sia un polo di $H(s)$, viene chiamata “risposta in frequenza” del sistema (A, B, C, D) .

Formalmente:

$$H(j\omega) = H(s)|_{s=j\omega}$$

Si noti che vale la seguente osservazione:

$$H(-j\omega) = H^*(j\omega)$$

Questa osservazione potrà essere molto utile quando si forniranno le principali rappresentazioni grafiche per la risposta in frequenza.

3.3 Esempio pratico

Si consideri il sistema caratterizzato mediante la seguente funzione di trasferimento:

$$H(s) = \frac{2}{1 + \frac{s}{10}}$$

Eccitato dall'ingresso:

$$u(t) = 5 \sin(\omega t)$$

Si vuole trovare la risposta, l'uscita relativa a questo segnale per $\omega = 1, 10, 100$ rad/s.

Per $\omega = 1$, si ha:

$$Y_1 = 5 \cdot \frac{2}{\sqrt{1 + \frac{1}{100}}} \simeq 10$$

$$\arg \{H(s)\}_1 = -\arctan 0,1 = -0,0997$$

Per $\omega = 10$, si ha:

$$Y_{10} = 5 \cdot \frac{2}{\sqrt{1 + \frac{100}{100}}} = 5 \cdot \frac{2}{\sqrt{2}} = 5\sqrt{2}$$

$$\arg \{H(s)\}_{10} = \arctan 1 = -\frac{\pi}{4}$$

Per $\omega = 100$, si ha:

$$Y_{100} \simeq 0,2$$

$$\arg \{H(s)\}_{100} \simeq -$$

3.4 Rappresentazioni grafiche della risposta in frequenza

Dove risiede l'importanza, il fatto per cui è necessario dedicare parte della trattazione al fine della comprensione della risposta in frequenza? Beh, la risposta in frequenza è fondamentale non solo ai fini di un'analisi del sistema, ma anche e soprattutto per il progetto di controlli per il sistema: le tecniche nel dominio del tempo per il progetto di sistemi esistono, sono molto intuitive poichè operano “direttamente” sul sistema, e hanno un significato fisico molto concreto, ma sono molto complicate da attuare. Oltre un certo punto non si riesce a comprendere, a partire da una certa scelta di progetto, quale sia la conseguenza nel dominio del tempo. Le tecniche di analisi e progetto in frequenza sono un po' più complicate da comprendere rispetto a quelle nel dominio del tempo, poichè il passaggio di dominio di fatto rende più “astratte” le idee nascoste dietro di esse; da attuare tuttavia queste tecniche sono decisamente più semplici rispetto alle altre, quindi nella letteratura della controllistica si ha solitamente a che fare con un parco di metodi molto più ampio.

3.4.1 Diagrammi di Bode

Alcune nozioni riguardanti i diagrammi di Bode sono date per scontate dai corsi di Elettrotecnica / Elettronica; ciò che si introdurrà in questa trattazione sono un breve ripasso e alcuni “trucchi” per la rappresentazione dei diagrammi di modulo e fase.

Prima cosa fondamentale da ricordare: data una funzione di trasferimento $H(s)$, imponendo che $s = j\omega$ si considera, di tutta la funzione di trasferimento, solo la risposta in frequenza. Partendo dunque dalla forma in costanti di tempo:

$$H(j\omega) = K_{st} \cdot \frac{\prod_i \left(1 - \frac{j\omega}{z_i}\right) \prod_j \left(1 + 2\xi_j \frac{j\omega}{\omega_{n,j}} + \frac{-\omega^2}{\omega_{n,j}^2}\right)}{\prod_k \left(1 - \frac{j\omega}{p_k}\right) \prod_l \left(1 + 2\xi_l \frac{j\omega}{\omega_{n,l}} + \frac{-\omega^2}{\omega_{n,l}^2}\right)}$$

Questa è una funzione di variabile complessa, ed è una funzione che assume valori complessi al variare della frequenza ω considerata¹. In altre parole, si tratta di un numero complesso che varia. Come ciascun numero complesso, esso può essere descritto mediante la parte reale e immaginaria (uti-

¹Si ricordi sempre che si confondono il termine “frequenza” con quello più idoneo “pulsazione”; si considera, nella trattazione, esclusivamente la pulsazione ω ; $[\omega] = \text{rad/s}$

lizzando una rappresentazione cartesiana) o mediante una rappresentazione in modulo e fase (rappresentazione polare).

Una generica funzione di risposta in frequenza, valutata in decibel (dB), avrà una forma di questo tipo:

$$|H(j\omega)|_{\text{dB}} = 20 \log_{10}(K_{st}) + 20 \sum_i \log_{10} \left| 1 - \frac{j\omega}{z_i} \right| + 20 \sum_j \log_{10} \left| 1 + 2\xi_j \frac{j\omega}{\omega_{n,j}} - \frac{\omega^2}{\omega_{n,j}^2} \right| - \left(r 20 \log_{10}(j\omega) \right)$$

Ragionamento simile per quanto riguarda la fase:

$$\angle H(j\omega)_{\text{dB}} = 20 \log_{10}(K_{st}) + 20 \sum_i \log_{10} \angle \left(1 - \frac{j\omega}{z_i} \right) + 20 \sum_j \log_{10} \angle \left(1 + 2\xi_j \frac{j\omega}{\omega_{n,j}} - \frac{\omega^2}{\omega_{n,j}^2} \right) - \left(r 20 \log_{10}(j\omega) \right)$$

Si scompone tutto in diversi contributi, sfruttando la proprietà del logaritmo, quindi si considerano singolarmente.

Cosa significa studiare la risposta in frequenza? Beh, come è noto dallo studio della trasformata di Laplace, la variabile s è complessa, e ha una forma del tipo:

$$s = \sigma + j\omega$$

Il dominio della funzione di trasferimento, dunque, è il piano complesso. Al fine di visualizzare correttamente la funzione di trasferimento sarebbe necessaria una rappresentazione tridimensionale: il piano complesso come dominio, la quota come valore della funzione di trasferimento. Si noti che, essendo la funzione di trasferimento una funzione che assume valori complessi, sarà necessario usare separatamente due grafici: uno per rappresentare il modulo, l'altro per rappresentare la fase.

Precedentemente si è detto che in prossimità di un polo la funzione di trasferimento tende a divergere, ossia ad assumere valori molto elevati.

In prossimità dei poli, vi sono dei coni che tendono a valori infiniti. Dualmente, in presenza di zeri, la funzione va in una sorta di "avvallamento", tende a crollare, come in una buca.

Questa è una rappresentazione del modulo della funzione di trasferimento; a partire da essa, è possibile comprendere cosa sia la risposta in frequenza: osservando, di tutti i valori della funzione di trasferimento, solo quelli che essa acquisisce sull'asse immaginario, ossia $j\omega$ (imponendo $\sigma = 0$), quella che si osserva è la risposta in frequenza.

Esempio pratico

In presenza di poli immaginari puri, la risposta in frequenza assume un comportamento particolare; data ad esempio la seguente funzione di trasferimento:

$$H(s) = \frac{1}{1 + s^2}$$

Determinarne la risposta in frequenza.

Si può vedere facilmente che l'espressione della risposta in frequenza sia:

$$H(j\omega) = \frac{1}{1 - \omega^2}$$

In questo caso, dunque, $\omega_n = 1$, e $\xi = 0$. Si può vedere facilmente che per $\omega \rightarrow 0$, il modulo tende a 1; quando $\omega \rightarrow \infty$, il modulo tende a zero; quando $\omega \rightarrow 1^\pm$, la funzione di risposta in frequenza tende a divergere.

Questo comportamento è interessante: solo nel caso in cui si hanno poli sull'asse immaginario, la risposta in frequenza tende a divergere. Questo fatto è abbastanza intuibile: la risposta in frequenza, come già detto, rappresenta di fatto l'osservazione dei valori che la funzione acquisisce sull'asse immaginario; dal momento che il polo deforma, fa divergere solo i punti in cui il polo è prossimo, di fatto sull'asse immaginario non si potrà osservare altro che una deformazione derivante dal polo, ma non certo valori tendenti a infinito. Se però i poli sono immaginari, il punto divergente sarà proprio sull'asse immaginario, dunque la risposta in frequenza, ossia l'insieme dei valori della funzione di trasferimento assunti sull'asse immaginario, tenderà a infinito.

Si provi a questo punto a interpretare il discorso fatto: dato uno zero nel semipiano sinistro del dominio di Laplace, esso, come già detto, comporta la nascita di una sorta di "buco" nella funzione di trasferimento. Osservando un piano complesso, si può vedere che, se la buca è a "sinistra" dell'asse immaginario, da quel valore di pulsazione in poi, "allontanandosi" dallo zero, si avrà una crescita della funzione di trasferimento. Discorso duale per quanto riguarda il polo: dal momento che il polo fa divergere verso $+\infty$ la funzione di trasferimento, se il polo è "a sinistra" dell'asse immaginario, esso provocherà una deformazione dello spazio tale per cui, allontanandosi da esso, il valore della funzione di trasferimento decrescerà. Se i poli o gli zeri sono a destra dell'asse immaginario, il ragionamento è opposto: un polo fa sempre divergere la funzione a $+\infty$, ma quello che capita ora è che, "allontanandosi" dal polo, la funzione crescerà. Questo detta il fatto che la funzione di trasferimento caratterizzi un sistema instabile. Dualmente, lo

zero a destra avrà un effetto opposto. Questo discorso è valido soprattutto in termini di fase (si tornerà sull'argomento).

Una piccola osservazione: la pulsazione alla quale si ha la “rottura” coincide sempre con il modulo di uno zero; si consideri ad esempio la seguente funzione di trasferimento:

$$H(s) = s + 1$$

La risposta in frequenza sarà:

$$H(j\omega) = 1 + j\omega$$

La pulsazione dello zero, nella risposta in frequenza, coincide con il modulo dello zero:

$$\omega_z = |z| = 1$$

Quello che bisogna in pratica capire è: dato uno zero in un certo punto, tagliando con l'asse $j\omega$ la funzione di trasferimento e considerando solo i valori su di esso, si ha un effetto di questo genere. Se lo zero è a parte reale positiva, il sistema viene detto a “fase non minima”; discorso simile se sono presenti poli sul semipiano destro della trasformata di Laplace. Si può verificare, come già accennato, che zeri e poli per la fase invertono i propri contributi: uno zero a destra farà decrescere la fase, un polo a destra la farà crescere. Si noti che questa regola si può usare se si è avuta l'accortezza di usare la forma a costanti di tempo, altrimenti si rischia di commettere errori. Per questo motivo, essa è preferibile. Per quanto riguarda i poli complessi coniugati, il modulo dipende dallo smorzamento.

Se i poli si avvicinano all'asse immaginario, lo smorzamento diminuirebbe fino a diventare nullo. Il modulo, nel frattempo, continuerebbe ad aumentare, fino a divergere a $+\infty$, sulla risposta in frequenza. Discorso analogo si può fare riguardo la fase: se $\xi \rightarrow 0$, la crescita/decrecita di fase è sempre più rapida. Quando $\xi = 0$, ossia si han poli immaginari puri, la decrecita di fase è sostanzialmente “a gradino”, immediata.

Alcuni accorgimenti per il disegno dei diagrammi di Bode

Si considerano a questo punto alcuni casi particolari da conoscere quando si intende rappresentare, mediante diagrammi di Bode, la risposta in frequenza associata a una funzione di trasferimento. Spesso (se non sempre) un diagramma di Bode va rappresentato mediante strumenti informatici quali

MATLab, tuttavia è necessario conoscere alcune idee di base, in modo da poter individuare eventuali errori del software.

1. Dati poli complessi coniugati, del tipo:

$$s^2 + 2\xi\omega_n s + \omega_n^2$$

Si interpreta il polinomio come un generico:

$$s^2 + as + b$$

Si devono identificare i due parametri del polinomio, ossia ω_n e ξ . Per farlo, è suggeribile utilizzare le seguenti osservazioni:

$$\omega_n = \sqrt{b}$$

Da qui, poi:

$$\xi = \frac{a}{2\omega_n}$$

Questi dati sono estremamente utili: al fine di identificare la pulsazione di picco, si possono passare queste coordinate in coordinate cartesiane, ricavando:

$$\omega_{pk} = \omega_n \sqrt{1 - \xi^2}$$

Il valore del picco, M_{pk} , è calcolabile come:

$$M_p = \frac{1}{2\xi\sqrt{1 - \xi^2}} \quad \text{Se } \xi \leq 0,707$$

Volendo disegnare la funzione di trasferimento, si parte dal guadagno stazionario K_{st} , quindi si scende, a partire da ω_n , a -40 dB/dec ; per disegnare anche il picco, è sufficiente riportare nel grafico ω_{pk} e M_{pk} , quindi usare queste rappresentazioni, più dettagliate.

2. Data una funzione di trasferimento $H'(S)$ senza poli nell'origine e guadagno unitario, si abbia:

$$H(s) = K_p H'(s)$$

Il diagramma di Bode a bassa frequenza sarà una costante: dal momento che $H'(s)$ non ha poli nell'origine, fino al primo polo il guadagno sarà costante. Poi si useranno i criteri abituali per disegnare il resto del diagramma.

3. Se si ha un polo nell'origine, ossia, data la solita $H'(s)$, un'espressione del tipo:

$$H(s) = \frac{K_v}{s} H'(s)$$

Essendoci un polo nell'origine si parte con una pendenza pari a - 20 dB/dec . Si può a questo punto individuare il valore della pulsazione per cui si incrocia l'asse 0 dB, come:

$$\omega = |K_v|$$

4. Data una funzione di trasferimento con due poli nell'origine, ossia del tipo:

$$H(s) = \frac{K_a}{s^2} H'(s)$$

Il discorso è del tutto analogo al precedente, con due varianti: la prima, il fatto che la pendenza iniziale sarà di - 40 dB/dec (a causa del doppio polo), la seconda il fatto che questa volta l'intercetta con l'asse 0 dB avrà un'espressione del tipo:

$$\omega = \sqrt{|K_a|}$$

5. Se è presente uno zero nell'origine, ossia si ha un'espressione del tipo:

$$H(s) = sK' H'(s)$$

Il parametro K' si può calcolare come:

$$K' = \lim_{s \rightarrow 0} \frac{1}{s} H(s)$$

La posizione dell'intercetta della retta (di pendenza + 20 dB/dec) con l'asse 0 dB sarà:

$$\omega = \frac{1}{|K'|}$$

6. Nel caso ci siano due zeri nell'origine, ossia con una funzione di trasferimento del tipo:

$$H(s) = s^2 K'' H'(s)$$

Il parametro K'' si può calcolare come:

$$K'' = \lim_{s \rightarrow 0} \frac{1}{s^2} H(s)$$

A partire da ciò, si può verificare il fatto che la pulsazione di intercetta con l'asse 0 dB valga:

$$\omega = \frac{1}{\sqrt{|K''|}}$$

Alcune osservazioni aggiuntive, alcuni “trucchi aggiuntivi” verranno ora presentati. Un primo trucco può riguardare la “taratura” del diagramma del modulo: data la fattorizzazione zeri-poli, ossia basata sul K_∞ , si può dire che il comportamento ad alta frequenza dipende dal numero di zeri e poli presenti nel sistema. Si può facilmente vedere che:

$$\lim_{\omega \rightarrow \infty} |H(j\omega)| = \frac{|K_\infty|}{\omega^{n-m}}$$

Ad alta frequenza il modulo della risposta in frequenza si presenterà come una retta di pendenza $20 \cdot (n - m)$ dB/dec. L'intersezione con l'asse a 0 dB, nella fattispecie, sarà:

$$\omega = \sqrt[n-m]{|K_\infty|}$$

Per quanto riguarda la fase, il segno del guadagno stazionario è fondamentale: il diagramma della fase si può abbozzare ma risulta essere molto approssimativo; fondamentale è avere quantomeno un'idea del suo andamento, al fine di rilevare eventuali errori nel calcolatore. Un'idea per ricavare un andamento asintotico della fase potrebbe essere quella di approssimare la variazione di fase con una retta che va da un quinto della pulsazione di rottura a cinque volte la pulsazione di rottura; questa tecnica è più precisa rispetto a quella classica, basata sul considerare variazioni prima e dopo una decade dalla pulsazione di rottura.

3.4.2 Diagrammi polari e diagrammi di Nyquist

Al fine di presentare l'idea di diagramma polare, quindi poi di diagramma di Nyquist, si sceglie di presentare un esempio teorico/pratico, in grado di evidenziare la necessità di una rappresentazione di questo tipo.

Si consideri nella fattispecie il sistema caratterizzato dalla seguente funzione di trasferimento:

$$H(s) = \frac{1}{1+s}$$

Considerandola limitata al solo asse immaginario, ossia $s = j\omega$, si può ricavare banalmente l'espressione della risposta in frequenza:

$$H(j\omega) = \frac{1}{1+j\omega}$$

I diagrammi di Bode di modulo e fase di questa risposta in frequenza hanno un andamento di questo genere:

Si supponga a questo punto di fare un'operazione di questo genere: si consideri un piano di Gauss, ossia le cui ascisse rappresentano la parte reale di un numero complesso, le ordinate la parte immaginaria del medesimo; il numero complesso in questione sarà l'andamento della risposta in frequenza del sistema. Si suppone a questo punto di collezionare dei valori di modulo e fase su questo "piano dei valori della funzione", uno per ciascun valore della pulsazione ω ; per ciascuno di essi si utilizza una rappresentazione di tipo vettoriale, ossia si identifica ciascun punto mediante un vettore. Data ad esempio $\omega = 0$, la parte immaginaria si annulla, il vettore risultante ha modulo unitario, dunque il primo vettore identificabile sul piano sarà $(0, 1)$. Aumentando la pulsazione, la fase del sistema tenderà a diminuire, come d'altra parte il modulo: al crescere di ω , la componente immaginaria al denominatore tende ad aumentare il proprio "peso", provocando sostanzialmente due effetti: una variazione di fase tendente al seguente valore

$$\lim_{\omega \rightarrow \infty} \arg \{H(j\omega)\} = \angle \frac{1}{j\omega} = \frac{-\pi}{2}$$

e il fatto che, crescendo ω , il modulo tende a ridursi:

$$\lim_{\omega \rightarrow \infty} |H(j\omega)| = 0$$

La fase si stabilizzerà dunque a -90° , con modulo nullo; involupando le punte di ciascun vettore, si otterrà un diagramma di questo tipo:

Ossia il diagramma polare della risposta in frequenza, per quanto concerne le pulsazioni ω positive.

Data la risposta in frequenza positiva, la risposta in frequenza concernente le frequenze negative è del tutto analoga: è sufficiente simmetrizzare la funzione rispetto all'asse reale, e “chiudere” il grafico². Il diagramma risultante da questo processo di simmetrizzazione è comunemente noto come “diagramma di Nyquist”.

3.4.3 Esempi teorici/pratici

Si considerano a questo punto due esempi “pratici”; il primo sarà in grado di mostrare un ulteriore esempio di rappresentazione dei diagrammi, mentre il secondo proporrà un esempio “patologico”, finora non presentato, riguardante una casistica particolare, per quanto a volte presente in sistemi di cui è necessaria una rappresentazione mediante diagramma di Nyquist.

Esempio pratico

Data la funzione di trasferimento, ricavarne un diagramma di Nyquist qualitativo:

$$H(s) = \frac{1}{(1 + s\tau_1)(1 + s\tau_2)}$$

Innanzitutto, si può vedere immediatamente che la risposta in frequenza avrà un andamento di questo tipo:

Per quanto riguarda il diagramma di Nyquist, le osservazioni da fare sono analoghe a quelle precedentemente affrontate: il primo polo fa abbassare la fase di $\frac{\pi}{2}$, il secondo di altri $\frac{\pi}{2}$, dunque per valori molto elevati della pulsazione ω ci si può aspettare che la fase valga $-\pi$. Analogamente a prima, il modulo tenderà asintoticamente a 0, dal momento che i due poli, non compensati da alcuno zero, faranno decrescere il modulo fino a renderlo infinitesimo. Questo diagramma viene dunque “simmetrizzato”, ottenendo il seguente diagramma di Nyquist:

Il diagramma di Nyquist di una funzione di questo tipo ha un andamento di questo genere, riconducibile a quello di una sorta di cardioide.

Esempio teorico/pratico

Si considera a questo punto un esempio un po' più “particolare”, al fine di introdurre un tassello mancante alla teoria, e un metodo di risoluzione per casi

²Si noti che ciò che è stato appena affermato non è del tutto vero, ma presto verrà puntualizzato in modo da includere alcune casistiche particolari.

di questo tipo. La casistica “patologica”, per quanto concerne la rappresentazione che si sta trattando, è quella delle funzioni di trasferimento contenenti poli sull’asse immaginario. Si analizza, nella fattispecie, il seguente esempio pratico:

$$H(s) = \frac{1}{s}$$

Si propone a questo punto la rappresentazione mediante diagrammi di Bode:

E soprattutto quella mediante diagrammi di Nyquist: di fatto, la risposta in frequenza di questa funzione di trasferimento esiste (come si può vedere sostituendo a s la limitazione sul piano immaginario, $j\omega$); si può immaginare dunque che il solo asse immaginario, considerato da $-\infty$ a $+\infty$, sia la rappresentazione in diagramma di Nyquist della funzione. Si può nella fattispecie osservare che per pulsazioni prossime a 0, il modulo tende a crescere enormemente, andando dunque verso $-\infty$. Per pulsazioni molto grosse il modulo tende ad annichilirsi, verso un intorno negativo di 0; si può vedere meglio ciò esprimendo la funzione in termini della risposta in frequenza:

$$H(s) \longrightarrow H(j\omega) = \frac{1}{j\omega} = \frac{-j}{\omega}$$

Per quanto riguarda il diagramma di Nyquist, è sufficiente “simmetrizzare”, ottenendo un diagramma di questo tipo:

Si aggiunge a questo punto un’osservazione teorica finora non affrontata: in presenza di poli sull’asse immaginario, di fatto si ha una discontinuità della funzione di trasferimento su di esso; la caratteristica che deve avere il diagramma di Nyquist, è l’essere una curva rigorosamente chiusa; in questo esempio, i problemi sono sostanzialmente due, riguardo questa caratteristica:

- Il fatto che vi è questa discontinuità sull’asse immaginario, causata dal fatto che si ha una divergenza verso ∞ dei valori della funzione di risposta in frequenza;
- Il fatto che il diagramma non è chiuso, poichè è su di una retta, da $-\infty$ a $+\infty$.

Il secondo problema è il più semplice da trattare: in realtà, da $-\infty$ a $+\infty$ (ossia dai valore di pulsazione $\omega = 0^-$ a quello di valore $\omega = 0^+$), è presente la cosiddetta “circonferenza canonica di raggio infinito”: una retta, di fatto, si può pensare come una degenerazione di una circonferenza con un raggio molto grosso, matematicamente modellizzabile come infinito.

Il secondo errore è più complicato, per quanto comunque risolvibile con un piccolo stratagemma: il fatto che vi sia una singolarità può essere risolto, “aggirando” il polo mediante una semicirconferenza di raggio piccolo a piacere, ossia con una curva del tipo:

$$s = re^{j\vartheta} \quad \vartheta \in \left[-\frac{\pi}{2}; +\frac{\pi}{2}\right]$$

Si noti un ulteriore fatto: in questo caso, si ha un polo singolo; se il polo sull’asse immaginario (nella fattispecie in questo caso sull’origine) fosse stato doppio, si sarebbe avuto qualcosa del tipo:

$$H(s) \frac{1}{s^2} \longrightarrow \frac{1}{(re^{j\vartheta})} = \frac{1}{r^2 e^{j2\vartheta}} \quad \vartheta \in \left[-\frac{\pi}{2}; +\frac{\pi}{2}\right]$$

In altre parole, il metodo di risoluzione è il seguente: per ogni i -esimo polo sull’asse immaginario, introdurre r_i semicirconferenze che lo “aggirano”, dove r_i è la molteplicità del polo in questione.

Si ribadisce un fatto fondamentale: le due chiusure da introdurre, in casi di questo tipo, sono le semicirconferenze appena descritte, e la “curva di chiusura” (in questo caso la circonferenza di raggio infinito); fondamentale è anche l’orientamento di questa curva: il diagramma di Nyquist va considerato orientato in senso orario; al fine di mantenere la coerenza con il resto del diagramma, la chiusura finale, in questo caso e in ogni caso, va fatta collegando a partire dal punto a pulsazione $\omega = 0^-$ a quello con $\omega = 0^+$, tassativamente e sempre.

3.5 Il criterio di Nyquist per lo studio della stabilità di sistemi retroazionati

Il “percorso” sull’asse immaginario poc’anzi indicato, in presenza di poli immaginari puri, è detto “percorso di Nyquist”. Questo percorso, completato secondo le modalità descritte, abbraccia tutto il semipiano destro esclusi poli immaginari, che se inclusi non permetterebbero di ricavare i risultati proposti.

A partire da questa rappresentazione, si proporrà un criterio fondamentale per la determinazione della stabilità di sistemi retroazionati; prima di ciò, tuttavia, è necessario introdurre alcune premesse fondamentali.

Un generico sistema di controllo a catena chiusa si baserà sempre su di una forma di questo tipo:

La cui funzione di trasferimento risultante è:

$$G_{ry}(d) = \frac{G(s)}{1 + G(s)H(s)}$$

Questa funzione si può ri-scrivere mediante il seguente artificio algebrico:

$$G_{ry}(s) = \frac{1}{H(s)} \cdot \frac{G(s)H(s)}{1 + G(s)H(s)}$$

Ciò potrebbe suggerire una visione grafica un po' alternativa rispetto a quella finora analizzata, basata sull'uso di un sistema di controllo basato su retroazione unitaria:

Lo studio della stabilità esterna su sistemi di questo tipo si effettua studiando i poli della funzione di trasferimento, globale, ossia gli zeri di $1 + G(s)H(s)$. La singola $H(s)$ non conta: dal momento che ci si è riportati in questa forma mediante l'artificio algebrico, mediante una cancellazione zeri-poli si può riportare il sistema i cui poli dipendono esclusivamente dal termine prima proposto. Ciò che si vedrà, inoltre, è che $H(s)$ molto frequentemente è una costante, dunque non introdurrebbe radici a prescindere dalle cancellazioni.

Proponiamo a questo punto una serie di “ingredienti” per il criterio di Nyquist:

- $n_{p,a}$: numero di poli della funzione di anello $G_a(s)$, definita come:

$$G_a(s) = G(s)H(s)$$

Di tutti i poli, si considerano solo quelli con parte reale positiva;

- $n_{p,r}$: numero di poli della funzione di trasferimento globale, $G_{ry}(s)$, con parte reale positiva;
- Il percorso di Nyquist precedentemente proposto;
- Si definisce il “punto critico” C_P con riferimento alla seguente struttura di controllo:

Il valore:

$$C_P = \mp \frac{1}{K}$$

Si noti che K è un coefficiente; quello che spesso si vorrà fare, è lo studio di stabilità del sistema al variare del parametro K , ossia ci si chiede se esista un qualche controllore in grado di stabilizzare il sistema al variare del parametro K .

- N indica il “numero di rotazioni” compiute dal diagramma di Nyquist intorno al punto critico. N è un numero con segno: si considera $N > 0$ se le rotazioni sono in senso orario, $N < 0$ se le rotazioni sono in senso antiorario, in accordo con il sistema di riferimento e verso di percorrenza finora definito;
- Il diagramma di Nyquist della funzione di anello $G_a(s)$.

Si noti che se il diagramma di Nyquist passa esattamente per il punto critico, si suol dire che N “non è ben definito”. Ciò significa che vi è un passaggio da una situazione di stabilità esterna a una di instabilità esterna. Questo capita se la funzione di trasferimento presenta dei poli sull’asse immaginario, ossia dei punti sulla “borderline”, sulla linea di confine tra la stabilità e l’instabilità.

Teorema (Principio della fase di Cauchy)

Dato N ben definito, il numero di poli di G_{ry} con parte reale positiva è dato da:

$$n_{p,r} = n_{p,a} + N$$

Teorema (Criterio di Nyquist)

A partire dal teorema di Cauchy appena proposto, Nyquist ha fatto alcune osservazioni, ottenendo un risultato molto interessante in ambito di sistemi lineari tempo-invarianti.

Dato N ben definito, il sistema retroazionato (riferendosi alla figura precedentemente proposta) è esternamente stabile se e solo se $n_{p,r} = 0$, ossia se $N = -n_{p,a}$.

Esempio pratico

Dato un sistema caratterizzato dalla seguente funzione di trasferimento:

$$G(s) = \frac{10}{s + 1}$$

Determinarne la stabilità.

In questo caso, $G_a(s) = G(s)$, dunque la funzione di anello non presenta poli instabili. Si può da ciò evincere che $n_{p,a} = 0$. In questo caso si ha $K = 1$, dunque:

$$C_P = \mp \frac{1}{K}$$

Dal momento che la retroazione è negativa, si considera il valore negativo: $C_P = -1$.

Si deve a questo punto disegnare il diagramma di Nyquist:

Si può osservare che il punto critico non è mai circondato dal diagramma di Nyquist, dunque $N = 0$; dal momento che $N = n_{p,a}$, il sistema è esternamente stabile.

Si può a questo punto considerare una piccola variante: introducendo un K generico, è possibile ottenere uno studio parametrico delle caratteristiche del sistema. Sostanzialmente, sono presenti tre casistiche:

1. Se si verifica la condizione:

$$-\frac{1}{K} < 0$$

Il sistema è esternamente stabile;

2. Se si ha come condizione:

$$0 < -\frac{1}{K} < 10$$

Si ha $N = +1$, dunque $n_{p,r} = 1$, e il sistema non è asintoticamente stabile;

3. Se si verifica la condizione:

$$-\frac{1}{K} > 10$$

Allora, come nel primo caso, $N = 0$, $n_{p,r} = 0$, dunque il sistema è stabile.

Capitolo 4

Le caratteristiche dei sistemi di controllo retroazionati con un ingresso e una uscita (SISO)

Finora sono stati introdotti alcuni richiami concernenti alcuni strumenti utili, quindi è stato analizzato un criterio fondamentale in grado di determinare le caratteristiche di stabilità di un sistema: il criterio di Nyquist.

Questo capitolo entra nel cuore della trattazione, e sarà la base per lo studio del problema del progetto di un controllore.

4.1 Una struttura di un sistema di controllo

La struttura generale analizzata all'inizio della trattazione aveva una forma di questo tipo:

Raccolte le informazioni disponibili per ciascun blocco e raccolte le informazioni sull'uscita desiderata, si decide quale deve essere la legge del controllo, in maniera automatica. Di solito, tecnologicamente, il sistema viene realizzato mediante blocchi che gestiscono segnali a bassa potenza, dunque si utilizzano blocchi di attuazione/amplificazione quali \mathcal{A} . La struttura appena mostrata è la più generale; nel corso della trattazione ci si occuperà prevalentemente di una struttura particolare, di questo tipo:

Si analizzano a questo punto i principali elementi contenuti in questa struttura di controllo: $G_p(s)$ indica la funzione di trasferimento dell'impianto, del sistema da controllare, normalmente dato; $G_d(s)$ tiene conto del legame tra disturbi esterni additivi (d_2) e l'uscita; d_1 è un disturbo additivo introdotto dall'attuatore \mathcal{A} (un offset ad esempio; G_c , G_y , G_r sono blocchi da progettare. G_c nella fattispecie è il cuore del controllo, nonchè il pezzo

più importante da progettare. G_t è di fatto un trasduttore, o meglio la sua funzione di trasferimento, e di solito non va progettato; esso deve avere una dinamica ampia, in modo da poterlo considerare costante nel range di applicazioni in cui viene utilizzato; potrà avere zeri e poli ad alte frequenze, ma non a basse, in modo da avere un guadagno costante. Stesso discorso per quanto concerne l'attuatore: non si imparerà a realizzare, ma esso sarà scelto in modo da essere considerabile costante, ossia in modo da non introdurre zeri o poli nella banda passante del sistema retroazionato complessivo.

In questa trattazione non verranno approfonditi aspetti tecnologici per quanto riguarda la realizzazione di sistemi di controllo, bensì esclusivamente aspetti metodologici: si imparerà a scegliere, a ottenere specifiche sui vari blocchi, in modo da poter introdurre le basi necessarie per la realizzazione tecnologica di un controllo.

G_r e G_y servono per gestire il rapporto di scala in regime permanente, o in altre parole il guadagno stazionario per il sistema di controllo. Normalmente si imposterà esclusivamente G_y ; G_r sarà tuttavia presente fisicamente su alcuni sistemi; in tal caso, esso sarà chiamato "trasduttore del riferimento".

Come detto precedentemente, l'uscita y va assolutamente misurata; il riferimento r è "l'uscita desiderata", ossia il segnale che il sistema di controllo deve cercare di inseguire, di riprodurre. I disturbi possono essere di due categorie, come già accennato: d_1 e d_2 sono disturbi additivi, d_t è anch'esso additivo ma di origine diversa, come si evidenzierà maggiormente in seguito, poichè è posizionato sul ramo di retroazione. Un disturbo può essere di tipo polinomiale (come si vedrà meglio in seguito), ma anche di tipo sinusoidale, o più generalmente parametrico.

Al fine di fissare le idee fondamentali necessarie per proseguire nella trattazione, si consideri il seguente problema: il posizionamento di un'antenna parabolica. Il sistema di controllo per il posizionamento sarà costituito da sistemi di ingranaggi con motori di tipo elettrico o pneumatico (pneumatici, utili per antenne particolarmente grosse). Supponendo di usare motori elettrici, ben noti, l'impianto in questione, il sistema da controllare, è costituito dall'antenna da posizionare; il motore di fatto è l'attuatore del sistema, l'antenna è G_p ; tra gli ingranaggi vi è un potenziometro, che si usa come trasduttore di posizione angolare. Al fine di ordinare al sistema di posizionare l'antenna secondo una certa angolazione, si usa come "console" di controllo un trasduttore di riferimento, G_r ; G_c e G_y sono parametri ancora da progettare. Volendo modellizzare l'intero sistema, si potrà dire che l'antenna avrà una certa inerzia, un certo attrito viscoso al movimento, e altri fenomeni. Il motore muove l'antenna, l'informazione viene "mandata indietro" mediante il K , verrà sottratta dall'informazione del movimento, costruendo il segnale di errore. Questo segnale viene utilizzato per correggere la posizione rispetto

a quella attuale, muovendo il motore. Modellizzando tutto ciò mediante un diagramma a blocchi di tipo simile a quello precedentemente introdotto, è possibile ottenere una rappresentazione ulteriore del sistema.

Questo esempio, molto qualitativo, può essere usato per identificare alcuni dei parametri, dei requisiti fondamentali per dimensionare un sistema di controllo. Questi parametri potrebbero essere nella fattispecie i seguenti:

- Stabilità: il sistema risultante deve assolutamente essere stabile;
- Fedeltà di risposta: un sistema deve essere in grado di riprodurre l'uscita desiderata nella maniera più fedele possibile sia in regime permanente sia nel transitorio;
- Il sistema deve avere una buona reiezione dei disturbi additivi: d_1 , d_2 , d_t devono essere eliminati;
- Esiste un'ulteriore classe di disturbi che il sistema deve essere in grado di attenuare: i cosiddetti disturbi "parametrici", ossia disturbi che provocano variazioni del sistema di controllo.

Si considera brevemente, a questo punto, un esempio pratico.

Esempio pratico

A partire dalle nozioni appena introdotte, si supponga che $\mathcal{A} = 1$, $G_r = 1$, $G_d = 1$, $G_c = K_C$ (un generico guadagno parametrico), $d_1 = 0$, $d_t = 0$, $d_2 \neq 0$; infine:

$$G_t \cdot G_y = \frac{1}{K_d}$$

Dove K_d è anche chiamato "coefficiente di proporzionalità tra guadagno e uscita". Si vuole esprimere il contributo dell'uscita quando agiscono i due ingressi del sistema: riferimento r e disturbo d_2 ; ovviamente, essi dovranno essere trattati in maniera assolutamente diversa: r è un ingresso da riprodurre, d_2 un ingresso da attenuare.

Dal momento che il sistema è lineare, è possibile applicare il principio di sovrapposizione degli effetti, ottenendo:

$$Y(s) = G_{ry}(s) \cdot R(s) + G_{d_2y}(s) \cdot D_2(s)$$

I contributi sull'uscita saranno ovviamente due: il contributo del riferimento, e quello del disturbo.

Applicando le regole fondamentali dell'algebra degli schemi a blocchi, ricordando lo schema fondamentale di lavoro con la retroazione, si può scrivere immediatamente che:

$$G_{ry}(s) = \frac{K_c G_p(s)}{1 + K_c G_p(s) \frac{1}{K_d}}$$

$$G_{d2y}(s) = \frac{1}{1 - (-1) \frac{1}{K_d} K_c G_p(s)}$$

Dunque:

$$Y(s) = \frac{K_c G_p(s)}{1 + K_c G_p(s) \frac{1}{K_d}} R(s) \frac{1}{1 - (-1) \frac{1}{K_d} K_c G_p(s)} D_2(s)$$

Cosa si può fare a questo punto? Beh, si può considerare K_c e provare a considerarlo molto elevato, tendente a valori infiniti:

$$\lim_{K_c \rightarrow \infty} Y(s) = K_d \cdot R(s) + 0$$

Se si fa crescere il guadagno, l'uscita è proporzionale al solo riferimento! Il sistema è sostanzialmente ideale, dal momento che $R(s)$ è esattamente riprodotta, a meno di una costante moltiplicativa K_d , che permetterà di regolarne un'eventuale amplificazione. Con un K_c molto elevato, dunque, si ottiene un'eccellente fedeltà e una perfetta reiezione del rumore.

Abbiamo scoperto la gallina dalle uova d'oro? No: attribuire valori troppo elevati a K_c comporta da un lato gli effetti proposti, dall'altro, si può vedere cosa capiti sotto il punto di vista della stabilità, analizzando il diagramma di Nyquist. Si supponga che il sistema in studio sia un generico sistema elettromeccanico, con una funzione di trasferimento del tipo:

$$G(s) = \frac{K_v}{s \left(1 + \frac{s}{p_m}\right) \left(1 + \frac{s}{p_e}\right)}$$

Generalmente un sistema elettromeccanico ha due poli: uno causato dai limiti meccanici del sistema, p_m , uno causato dai limiti elettrici di banda passante, p_e . Il diagramma di Nyquist di una funzione di trasferimento di questo tipo è qualcosa del genere:

Il punto critico, essendo la reazione negativa, sarà di questo tipo:

$$C_P = -\frac{1}{K_c}$$

Al variare del punto critico cambiano le condizioni del sistema; se si considera $K_c \rightarrow \infty$, il punto critico tende a 0, dunque n_a da un lato è sempre

nullo, ma N può variare da 0 a 2, rendendo di fatto il sistema instabile. Il sistema è sicuramente fedele, ma instabile, a causa del criterio di Nyquist. Come dice il proverbio, non si può avere la botte piena e la moglie ubriaca: troppo guadagno migliora le caratteristiche del sistema, ma lo rende anche instabile. In genere, raddoppiando il guadagno si dimezza l'errore; tenendo conto di questa considerazione, si può stabilire una soglia massima consentita per l'errore, e regolarsi di conseguenza.

4.2 La stabilità nei sistemi di controllo con retroazione

Finora si è parlato di quattro fondamentali caratteristiche dei sistemi di controllo; in questo paragrafo, nella fattispecie, si considererà un aspetto particolare: non è sufficiente parlare di stabilità, ossia chiedere che il sistema sia stabile, bensì è necessario introdurre una sorta di “grado di stabilità” per il sistema, ossia dire “quanto” il sistema è stabile, introducendo concetti di “stabilità relativa”. Si tratta in sostanza dello studio di problemi di robustezza del sistema: si cambiano determinati parametri, in modo che il sistema sia ben lontano dall'essere instabile. Si introducono nella fattispecie due “margini”, atti a quantificare la robustezza del sistema di controllo.

Margine di guadagno

Il primo dei due margini riguarda il guadagno: il margine di guadagno è l'estremo superiore dei fattori moltiplicativi della funzione di trasferimento $G(s)$ (quella “con il K in cascata”, nel sistema rappresentato mediante schemi a blocchi), che il sistema retroazionato può tollerare senza perdere la proprietà di stabilità esterna.

Si supponga ad esempio di avere un diagramma polare di questo tipo:

Se il punto critico è inizialmente posizionato su -1, cambiando il guadagno, accade che il punto critico cambia. Si può tuttavia ragionare in maniera diversa: supponendo che il punto critico sia fisso, non modificabile, si può pensare che cambiando il guadagno della funzione di trasferimento cambi la possibilità di “inglobare” il punto critico, facendo variare le proprietà di stabilità del sistema, come si sa dal criterio di Nyquist. La domanda è: qual è il più grande dei fattori moltiplicativi che si possono introdurre, prima che il sistema retroazionato divenga instabile? Chiamando questo punto m_G , si deve avere che:

$$m_G \cdot C_P = -1 \implies m_G = \left| \frac{1}{\overline{OC}} \right|$$

Ossia, tanto è più grande il segmento \overline{OC} , il segmento che congiunge l'origine con il punto critico C_P , tanto più stabile sarà il sistema; più si sta lontani, meglio è.

Margine di fase

Si consideri a questo punto la seguente situazione:

Il margine di fase si definisce come il massimo ritardo di fase della funzione $G(s)$ che il sistema retroazionato può tollerare senza perdere la proprietà di stabilità esterna.

Prima, introducendo un termine moltiplicativo m_G , si introduceva di fatto una variazione del guadagno della curva; ciò implica una variazione delle “dimensioni” della curva, il cui guadagno troppo elevato avrebbe rischiato di farle “abbracciare” il punto critico. Questa volta la trasformazione è differente: al variare del ritardo di fase φ , il diagramma di Nyquist del sistema retroazionato di fatto “ruoterà”, non “variando” le proprie ampiezze, ma rischiando comunque di “inglobare” il punto critico.

Qual è il massimo ritardo di fase tollerabile prima di perdere la proprietà di stabilità del sistema? Beh, si provi a intuire dalla seguente illustrazione:

Dal momento che la fase introduce esclusivamente una rotazione del diagramma di Nyquist, l'idea per calcolare il margine di fase potrebbe essere quella di intersecare il diagramma di Nyquist con la circonferenza di raggio unitario: dal momento che il punto critico è stato fissato su -1, introducendo un ritardo di fase si potrebbe indurre una variazione tale da inglobare; calcolando l'angolo mediante la circonferenza goniometrica, si può determinare m_φ , ossia il margine di fase della funzione di trasferimento.

Si noti che, mediante il software MATLAB, è possibile calcolare i margini di fase e modulo mediante il comando

```
margin()
```

I margini di modulo e fase, oltre sul diagramma di Nyquist, trovano un'interpretazione importante anche sui diagrammi di Bode di modulo e fase:

- Il margine di guadagno è, preso il punto in cui la fase vale -180° (riferendosi al fatto che -1 ha fase pari a π), la distanza dal punto in cui il modulo inizia a valere 0 dB, ossia 1.
- Il margine di fase si guarda per quella funzione in cui il modulo vale 1 (in altre parole, l'insieme dei punti appartenenti alla circonferenza goniometrica, ossia la circonferenza di raggio unitario), la distanza da essi al punto in cui la fase vale -180° .

Il senso è sempre questo: il punto “limite” è il punto critico, ossia -1 : il punto con modulo unitario e fase pari a π ; i margini, per quanto concerne modulo e fase, sono quei valori atti ad “allontanarsi” dal punto critico fissato, in modo da evitare l'instabilità del sistema.

4.3 Sistemi di controllo: la risposta in regime permanente a ingressi polinomiali

4.3.1 Concetti preliminari

L'argomento di questa sezione sarà la fedeltà della risposta in regime permanente, in presenza di ingressi di tipo particolare. Si ricorda che, per regime permanente, si intende la risposta del sistema una volta terminati i transitori, ossia dopo un tempo considerevolmente maggiore rispetto alle costanti di tempo del sistema. I particolari ingressi in questione, saranno quelli di tipo polinomiale.

Che senso ha studiare ingressi esclusivamente di questo tipo, ossia esclusivamente appartenenti alla classe dei polinomi? Beh, gli ingressi polinomiali sono molto, molto interessanti, dal momento che da soli riescono già a caratterizzare in maniera molto approfondita un sistema.

La generica espressione di un ingresso polinomiale è la seguente:

$$r(t) = R_0 \frac{t^K}{K!}$$

Questa è un'espressione di un generico segnale polinomiale; per $K = 0, 1, 2$, si avranno rispettivamente segnali a gradino, a rampa, a parabola, dotati delle rispettive trasformate di Laplace:

$$r(t) = R_0 \implies R(s) = \frac{R_0}{s}$$

$$r(t) = R_0 t \implies R(s) = \frac{R_0}{s^2}$$

$$r(t) = \frac{R_0}{2}t^2 \implies R(s) = \frac{R_0}{s^3}$$

Per quanto riguarda la trattazione in questione, questi sono i primi ingredienti; l'altro ingrediente fondamentale per parlare di fedeltà di risposta è il cosiddetto “sistema errore”: volendo misurare la fedeltà di risposta sarà necessario quantificare la bontà di riproduzione del riferimento, quindi quantificare un errore. Dato un sistema di controllo, supponendo di provare con uno dei segnali appena introdotti, si effettua sostanzialmente un confronto con il sistema ideale, ossia con il sistema che ripropone immediatamente l'uscita, senza alcuna manipolazione e senza alcun errore. L'idea fondamentale, dunque, è quella di introdurre un errore $e(t)$ definibile come:

$$e(t) = y_d(t) - y(t)$$

Si noti che gli attori potranno essere scambiati: il nostro interesse è quello di quantificare la distanza tra i due segnali, non quello di considerarne il segno. Nel caso del sistema ideale, l'ingresso dovrà essere completamente coincidente con l'uscita, a meno di una certa amplificazione, di una certa moltiplicazione delle ampiezze: K_d :

$$Y_d(s) = K_d R(s)$$

Il sistema errore $E(s)$, nel dominio di Laplace, verrà così definito:

$$E(s) = Y_d(s) - Y(s) = K_d \cdot R(s) - G_{ry}(s) \cdot R(s)$$

Si può raccogliere, ottenendo:

$$= [K_d - G_{ry}(s)] R(s)$$

Questo è l'errore che si commette; esso è di fatto modellizzabile mediante una funzione di trasferimento fittizia, $G_{re}(s)$, definibile come:

$$G_{re}(s) = K_d - G_{ry}(s)$$

Definizione: un sistema di controllo è di tipo “K” (si noti che non si usa la parola “ordine”, parola che in questo ambito verrà esclusivamente associata a segnali, non a sistemi) se l'errore in regime permanente corrispondente a un ingresso di ordine K è pari a una costante non nulla.

Cosa significa ciò? Beh, semplice: esaurito il transitorio, quando l'ingresso è di ordine 0, l'errore deve essere finito e non nullo, dunque l'uscita deve essere una costante, che differisce di un certo valore dall'ingresso; lo stesso discorso si può pensare per una retta, o per una parabola:

A seconda del tipo, si vedrà che i sistemi di controllo possono essere più fedeli a ingressi di diverso tipo.

Per ora, la definizione di “tipo” del sistema dipende esclusivamente dal segnale: un sistema può essere classificato a seconda del “tipo” solo per merito della risposta che propone a un dato ingresso di riferimento.

Un altro ingrediente per la trattazione è il seguente teorema.

Teorema: un sistema di controllo è di tipo “K” se e solo se la funzione di trasferimento di errore ($G_{re}(s)$) presenta nel punto $s = 0$ uno zero di molteplicità K .

Questo teorema non deve assolutamente stupirci: di fatto, osservando le precedenti espressioni delle trasformate di Laplace degli ingressi, si vede che ciascuna di esse propone almeno un polo nell’origine; a seconda dell’ingresso proposto, si avrà un certo numero di zeri nell’origine, atti a “cancellare” la presenza dei poli introdotti dagli ingressi, di fatto rendendo esistente il limite per le basse frequenze del valore della funzione di trasferimento, verificando le ipotesi del teorema del valore finale della trasformata di Laplace.

Data la funzione di errore, si sostituirà la $G_{ry}(s)$ dopo aver assunto di usare una struttura di controllo di questo genere:

Si può scrivere, come ben noto, $G_{ry}(s)$ come:

$$G_{ry}(s) = \frac{G(s)}{1 + \frac{1}{K_d}G(s)}$$

Si consideri a questo punto $G(s)$ come rapporto di una certa funzione al denominatore e una al denominatore:

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

Si può riscrivere $G_{ry}(s)$ come:

$$G_{ry}(s) = \frac{\frac{N(s)}{D(s)}}{1 + \frac{1}{K_d} \frac{N(s)}{D(s)}}$$

Da qui:

$$\Rightarrow = \frac{N(s)}{D(s) + \frac{N(s)}{K_d}}$$

Si recupera a questo punto la definizione di funzione di trasferimento di errore, e si introduce ciò che è stato appena scritto per ottenere un risultato molto importante:

$$G_{re}(s) = K_d - G_{ry}(s) = K_d - \frac{G(s)K_d}{K_d + G} = \frac{K_d^2}{K_d + G(s)}$$

Sostituendo:

$$= \frac{K_d^2}{K_d + \frac{N(s)}{D(s)}} = \frac{K_d^2 D(s)}{D(s)K_d + N(s)}$$

Questa è la funzione di errore; dal momento che il teorema si riferisce agli zeri di questa funzione si può osservare che essi siano strettamente imparentati con i poli della funzione di trasferimento $G(s)$. Da questa osservazione, si può estrarre un teorema molto importante, alla base di tutta la teoria del controllo che verrà ora introdotta.

Teorema: un sistema di controllo è di tipo K se e solo se la funzione di trasferimento $G(s)$ presenta nel punto $s = 0$ un polo di molteplicità K .

A questo punto si hanno tutti gli ingredienti per vedere, in uno di questi tre tipi di sistemi di controllo, come esso si comporti con ingressi test polinomiali di ordine 0, 1, 2, e verificare quale sia la fedeltà di risposta a segnali di questo tipo.

4.3.2 Caso dei sistemi retroazionati

Dato ad esempio un filtro, passa basso, si sa che esso ha una risposta in frequenza di questo tipo:

Dove la sua funzione di trasferimento avrà un andamento del tipo:

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

In questo caso, il regime permanente non è particolarmente interessante: nessuno mette in evidenza l'errore in regime permanente, dal momento che un sistema di questo tipo è fatto per attenuare segnali dotati di determinate caratteristiche, o solo alcune "parti" di alcuni segnali.

Nel caso di sistemi retroazionati, in generale si avranno configurazioni del tipo:

Nel corso della trattazione si considererà il progetto di retroazioni $H(s)$ statiche, ossia indipendenti dal valore di s . Partendo da questa ipotesi, si considerano alcune osservazioni particolari; come ben noto:

$$G_{ry}(s) = \frac{G(s)}{1 + G(s) \cdot H}$$

Si definisce a questo punto il guadagno stazionario del sistema retroazionato, identificandolo mediante K_d , come:

$$K_d = \lim_{s \rightarrow 0} G_{ry}(s)$$

Si noti che questo guadagno non è generalizzato: non si moltiplica per alcun coefficiente s , nella definizione appena introdotta; si suppone implicitamente dunque che la funzione di trasferimento del sistema retroazionato, del sistema controllato, non abbia poli nell'origine. Si consideri un altro fatto interessante: recuperando la definizione di funzione di trasferimento del sistema retroazionato, si può riscrivere il limite come:

$$K_d = \lim_{s \rightarrow 0} \frac{G(s)}{1 + G(s) \cdot H}$$

Se sono presenti poli nell'origine, per quanto riguarda la funzione $G(s)$, la funzione tenderà a divergere; il limite per $s \rightarrow 0$ della funzione di trasferimento, dunque, tenderà a:

$$K_d = \frac{1}{H}$$

Questo risultato è veramente molto interessante: considerando la struttura specifica di sistema di controllo che verrà trattata nel testo, si può dire che generalmente il guadagno stazionario (o rapporto di scala) della funzione di trasferimento del sistema retroazionato sia:

$$K_d = \frac{1}{H} = \frac{1}{G_t \cdot G_y}$$

Dal momento che G_t è un dato del problema (poichè non ci si preoccuperà di progettare il trasduttore), si può progettare G_y , data una specifica su K_d :

$$G_y = \frac{1}{G_t \cdot K_d}$$

Questo è il primo passo per il progetto. Determinare il tipo del sistema è semplice: basandosi sul teorema precedentemente presentato, si è detto che se $G(s)$ presenta poli nell'origine, allora il sistema ha tipo pari alla molteplicità del suddetto polo nell'origine.

Un esempio pratico di sistema di tipo "0" può essere un amplificatore operativo: si sa che esso presenta un polo a bassa frequenza (nell'intorno dei 10 Hz), uno a frequenze alte (intorno dei MHz); non avendo poli in zero a catena aperta, il sistema sarà di tipo 0.

Fino alla fine della sezione, si effettuerà un'operazione precisa: il calcolo dell'errore sui vari sistemi di controllo, al variare dei segnali di riferimento. Si noti fin da subito, come poi si farà notare in seguito, che i risultati ora ottenuti sono validi solo per uno specifico caso: il calcolo degli errori di rappresentazione del riferimento. Questi risultati non devono assolutamente essere confusi con quelli riguardanti i disturbi additivi: per ora si sta esclusivamente trattando la fedeltà di rappresentazione, senza considerare in alcun modo la reiezione dei disturbi, che verrà introdotta solo in seguito.

Quello che si intende quantificare è l'errore in regime permanente; al fine di fare ciò, si sfrutta la validità del teorema del valore finale, lavorando dunque mediante le trasformate di Laplace rappresentanti i sistemi in gioco:

$$E(s) = G_{re}(s)R(s) = \frac{K_d^2}{K_d G(s)} R(s)$$

Si può dire che:

$$e_\infty = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} R(s)$$

Si analizzano, a partire da questa idea, le varie risposte sui vari sistemi.

Sistema tipo "0"

Per quanto riguarda un sistema di tipo "0", si sa che in $G(s)$ non si hanno poli nell'origine. Si ha dunque:

$$K_{st} = K_p = \lim_{s \rightarrow 0} G(s)$$

- Utilizzando un ingresso di ordine "0" (ossia un gradino), si ha una cosa del tipo:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} R(s) = \frac{R_0}{s} \cdot s \cdot \frac{K_d^2}{K_s + G(0)} = \frac{R_0 K_d^2}{K_d + K_p}$$

L'errore sarà dunque, come già annunciato, finito, costante, non nullo;

- Utilizzando un ingresso di ordine "1", ossia "a rampa", si ottiene:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s^2}$$

Il sistema è di tipo "0", dunque ha una $G(s)$ priva di poli nell'origine; per questo motivo, il limite tende a ∞ .

- Introducendo un riferimento parabolico, il risultato sarà del tutto analogo a quello appena esposto: il limite divergerà a ∞ .

Una piccola nota: si è detto che il limite tende a ∞ , ma ciò non è stato giustificato: quello che si sta utilizzando per le dimostrazioni, è il teorema del valore finale, teorema che per essere usato ha bisogno di aver soddisfatte alcune ipotesi: il limite deve esistere ed essere finito. In questo caso il limite calcolato nel dominio di Laplace non è finito, cosa che rende di fatto il teorema non applicabile: non è sufficiente dire che il limite diverga solo perchè esso diverge nel dominio di Laplace: è necessario sviluppare l'espressione nel dominio del tempo, al fine di ricercare risultati. Si sappia che è comunque possibile farlo, per quanto i calcoli siano estremamente lunghi.

Sistema tipo “1”

Nel caso di sistemi di tipo “1”, il guadagno stazionario della funzione di trasferimento sarà un K_v , ossia avrà il significato fisico di “guadagno di velocità”:

$$K_{st} = K_v = \lim_{s \rightarrow 0} sG(s)$$

Dunque, per quanto concerne l'errore di rappresentazione di una funzione di trasferimento retroazionata, in un sistema di tipo “1”, si avranno i soliti tre casi, riferiti ai diversi tipi di ingressi di riferimento:

- Utilizzando un ingresso a gradino:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s}$$

Si moltiplica il denominatore per s , ottenendo:

$$\implies \lim_{s \rightarrow 0} \frac{sK_d^2}{sK_d + sG(s)} R_0 = \lim_{s \rightarrow 0} \frac{sK_d^2}{sK_d + K_v} R_0 = \frac{0}{K_v} = 0$$

Cosa ci dice questo risultato? Introducendo un riferimento a gradino in un sistema di tipo “1”, l'errore risultante è nullo.

- Utilizzando un ingresso a rampa:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s^2} = \frac{K_d^2}{K_v} R_0$$

Questo perchè i due poli introdotti dall'ingresso a rampa cancellano la variabile s al numeratore, ma comunque permettono di ottenere K_v al denominatore.

- Utilizzando un ingresso a parabola:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s^3}$$

Questo limite, come nel caso precedente, diverge, come si dovrebbe dimostrare nel dominio del tempo (come fatto nei casi precedenti, per il sistema di tipo "0").

Sistema tipo "2"

Si completa la carrellata per quanto riguarda i sistemi di tipo "2", ossia la cui $G(s)$ presenta un polo doppio nell'origine del dominio di Laplace. In questo caso, il guadagno stazionario, avrà un'espressione del tipo:

$$K_{st} = K_a = \lim_{s \rightarrow 0} s^2 G(s)$$

Si ragiona dunque rapidamente con i tre ingressi, al fine di completare il discorso:

- Nel caso di ingresso a gradino:

$$e_\infty = \lim_{s \rightarrow 0} s \cdot \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s} = 0$$

- Stesso ragionamento, per quanto riguarda un ingresso a rampa;
- Per quanto riguarda un ingresso a parabola, si verifica ciò:

$$e_\infty = \lim_{s \rightarrow 0} s \frac{K_d^2}{K_d + G(s)} \frac{R_0}{s^3} = \frac{K_d^2}{K_d + K_a}$$

Si noti la seguente osservazione fondamentale: man mano che aumenta l'indice del tipo del sistema di controllo, questo diventa sempre più "preciso", più "fedele" nella rappresentazione di ingressi di ordine inferiore al tipo. Ciò ci dà anche un'idea del fatto che la parabola, dei tre segnali, è il più "difficile" da rappresentare, da inseguire per il sistema, a regime permanente. Qualcuno potrebbe a questo punto proporre di progettare sistemi di controllo sempre di tipo 3 o 4, in modo da avere la certezza di rappresentare con fedeltà massima

il segnale $r(t)$, annullando di fatto l'errore. Questo non è possibile, poichè maggiore è il tipo del sistema, più difficile è da realizzare sia in termini di progetto, sia in termini di costi; si deve cercare, nei limiti del possibile, di utilizzare sempre il sistema di tipo inferiore possibile, per quanto le specifiche richiedano, in modo da avere progetti semplici da realizzare, e poco costosi.

4.4 Sistemi di controllo: la risposta in regime permanente a disturbi additivi

4.4.1 Disturbi polinomiali

Si è parlato delle caratteristiche di un sistema di controllo, tra cui la possibilità di attenuare disturbi di vario genere. Una categoria molto frequente di disturbi che si rilevano in sistemi è quella dei cosiddetti “disturbi polinomiali”, quali gradini, rampe, parabole; si sappia da subito che essi, generalmente, verranno considerati esclusivamente sul ramo diretto dello schema a blocchi, dunque molto raramente sulla retroazione.

Come si sa dai capitoli introduttivi alla trattazione, i disturbi sono ingressi; l'obiettivo della presente sezione è quello di quantificare il contributo sull'uscita di ciascun rumore, ossia studiare l'effetto che hanno sull'uscita del sistema.

Si consideri un esempio teorico/pratico: se si ha un sistema di tipo “1”, dato un riferimento a gradino, esso dovrebbe essere rappresentato con un errore nullo, dal momento che il tipo del sistema è superiore all'ordine del segnale di riferimento da utilizzare. Se vi fossero disturbi, tuttavia, la rappresentazione non sarebbe con errore nullo neanche al termine del transitorio, dal momento che vi sarebbero eventi aggiuntivi in grado di rovinare la rappresentazione del riferimento.

I sistemi che verranno trattati dispongono tuttavia di una fondamentale proprietà: la linearità! Essendo il sistema lineare, è possibile applicare il principio di sovrapposizione degli effetti, dunque separare i contributi sull'uscita dei segnali di riferimento (già considerati nella precedente sezione) e dei disturbi. Si noti assolutamente che i risultati precedenti non possono essere utilizzati per quanto riguarda i disturbi: di fatto, tutto ciò che è stato precedentemente calcolato e presentato vale solo per la funzione di trasferimento complessiva del sistema retroazionato, a partire dal riferimento verso l'uscita ($G_{ry}(s)$). Se per i riferimenti sono stati proposti risultati generali e immediatamente applicabili per qualsiasi sistema il cui schema a blocchi presenti una topologia collegabile a quella “specificata” utilizzata, per i disturbi si utilizzerà un approccio differente, più “didattico”: ogni volta che sarà

necessario calcolare il contributo di un disturbo sull'uscita, si dovrà calcolare l'espressione nel dominio di Laplace, quindi utilizzare il teorema del valore finale.

Esiste solo un risultato generale, per quanto fondamentale, che ora verrà proposto e semplificherà molti dei calcoli da effettuare per ogni sistema: essendo l'obiettivo della sezione quello di insegnare a quantificare l'errore sull'uscita, sarà necessario proporre, come precedentemente fatto, l'errore tra la funzione di uscita reale $Y(s)$ e quella desiderata, $Y_d(s)$ (per comodità si considera il tutto nel dominio di Laplace). Applicando il principio di sovrapposizione degli effetti, si può ottenere qualcosa di questo tipo:

$$E^{r,d_1}(s) = Y(s) - Y_d(s) = G_{ry}(s)R(s) + G_{d_1y}(s)D_1(s) - K_d R(s)$$

Si sommano dunque gli effetti del riferimento, del disturbo $d_1(t)$, e si sottrae l'uscita del sistema ideale, quello desiderato. Si ricorda a questo punto, applicando il teorema del valore finale, che l'errore in regime permanente, quando agiscono $r(t)$ e $d_1(t)$, è:

$$e^{r,d_1} = \lim_{s \rightarrow 0} s E^{e,d_1}(s) = \lim_{s \rightarrow 0} \{(G_{ry}(s) - K_d) R(s) + G_{d_1,y}(s) D_1(s)\}$$

Ciò che si sa, tuttavia, è che:

$$\lim_{s \rightarrow 0} G_{ry}(s) = K_d$$

Dunque:

$$e^{r,d_1} = \lim_{s \rightarrow 0} s E^{e,d_1}(s) = \lim_{s \rightarrow 0} \{(K_d - K_d) R(s) + G_{d_1,y}(s) D_1(s)\}$$

Ma quindi la parentesi interna va a 0, dunque:

$$e^{r,d_1} = \lim_{s \rightarrow 0} s G_{d_1y}(s) D_1(s)$$

Questo risultato è fondamentale, e si può interpretare nella seguente maniera: l'errore sull'uscita, quando agisce un determinato disturbo, coincide con il contributo sull'uscita di quel disturbo! Il contributo, dunque, è calcolabile semplicemente "spegnendo" tutti i generatori, gli ingressi del sistema, a meno dello specifico disturbo. In altre parole, si può dire che:

$$E^{d_1}(s) = Y^{d_1}(s)$$

Esempio pratico

Si supponga che agisca un disturbo polinomiale $d_1(t)$; si vuole quantificare l'errore sull'uscita quando agisce il suddetto.

Dal teorema precedente, si può dire che:

$$E^{d_1}(s) = Y^{d_1}(s)$$

Ma l'uscita è pari a:

$$Y^{d_1}(s) = G_{d_1y}(s)D_1(s) = \frac{G_p(s)}{1 + G_p(s)G_tG_yG_c(s)A}D_1(s)$$

Ogni volta, dunque, sarà sufficiente calcolare un limite di questo tipo, risolverlo, e quantificare l'errore.

4.4.2 Disturbi sinusoidali

Dopo la trattazione dei disturbi polinomiali, si vuole proporre una trattazione riguardo i disturbi sinusoidali. Al fine di poterli presentare, tuttavia, è necessario riprendere alcuni concetti riguardo la risposta in frequenza di un tipico sistema di controllo:

Qual è la risposta in frequenza del sistema? Beh, si supponga che la funzione di anello, $G_a(s) = G(s) \cdot H$ abbia un andamento di questo tipo (cosa che capiterà molto frequentemente):

La $G_{ry}(s)$ avrà una forma di questo tipo, come ben noto:

$$G_{ry}(s) = \frac{G(s)}{1 + H \cdot G(s)} = \frac{1}{H} \frac{G(s) \cdot H}{1 + G(s) \cdot H}$$

La seconda formula, come ben noto, rappresenta la funzione di trasferimento relativa allo schema a blocchi con retroazione unitaria. Questa espressione dunque è pari a:

$$G_{ry}(s) = \frac{1}{H} \frac{G_a(s)}{1 + G_a(s)} = \frac{1}{H} T(s)$$

La funzione $T(s)$ è di solito chiamata “funzione di sensibilità complementare”; come esiste la funzione di sensibilità complementare $T(s)$, esiste la funzione di sensibilità, $S(s)$, definibile come:

$$S(s) = \frac{1}{1 + G_a(s)}$$

Il termine “complementare” è usato perchè tra le due espressioni esiste un legame molto interessante:

$$T(s) + S(s) = 1 \quad \forall s \in \mathbb{C}$$

Disegnare gli andamenti di queste due funzioni è piuttosto semplice, a partire dalla conoscenza dell'andamento della funzione di anello, $G_a(s)$, quantomeno asintoticamente; si osservi che:

$$\text{Se } |G_a(s)| \gg 1 \implies T(j\omega) \simeq 1$$

$$\text{Se } |G_a(s)| \ll 1 \implies T(j\omega) \simeq |G_a(s)|$$

Si rapporta a “1” (o 0 dB) il valore del modulo della funzione di anello, $|G_a(s)|$, in modo da poter osservare quale sia l'andamento della risposta in frequenza della funzione di sensibilità complementare rispetto a quello di $|G_a(s)|$. Rapportare a “1” serve a considerare un valore asintotico, in modo da vedere se al denominatore prevale “1” o “ $|G_a(s)|$ ”. Quando dunque il modulo della funzione di anello assumerà valori bassi, considerevolmente minori di 1, la funzione di sensibilità complementare “seguirà” la funzione di anello in modulo; dualmente, quando il modulo della funzione di anello assumerà valori maggiori di 1, la funzione $T(j\omega)$ assumerà valori prossimi a 1.

Si faccia lo stesso ragionamento per quanto riguarda $S(s)$ e la sua risposta in frequenza $S(j\omega)$:

$$\text{Se } |G_a(s)| \gg 1 \implies S(j\omega) \simeq \left| \frac{1}{G_a} \right|$$

$$\text{Se } |G_a(s)| \ll 1 \implies S(j\omega) \simeq 1$$

Questa funzione ha un comportamento, come suggerisce il nome, abbastanza “complementare” rispetto a quello di $T(j\omega)$: per valori bassi del modulo della funzione di anello la funzione di sensibilità assumerà valori prossimi all'unità (0 dB); per valori di $|G_a(s)|$ elevati, la funzione $S(j\omega)$ assumerà valori prossimi al reciproco del modulo della funzione di anello; in scale logaritmiche, quali i decibel (dB), ciò significa semplicemente “ribaltare” il grafico rispetto all'asse delle ascisse, ottenendo dunque, per valori elevati di $|G_a(s)|$, una funzione di fatto simmetrica a quella del guadagno di anello in modulo, rispetto all'asse delle ascisse, delle pulsazioni.

Si parla di alta frequenza e di bassa frequenza; fondamentale è l'identificazione di ω_c , ossia della pulsazione in cui la funzione di anello vale 1, 0 dB. Fatte queste premesse, è ora possibile trattare disturbi di tipo sinusoidale, ossia del tipo:

$$d_2(t) = a_2 \sin(\omega_2 t) \quad \forall \underline{\omega_2} \leq \omega_2 \leq \overline{\omega_2}$$

Dove $d_2(t)$ è un disturbo additivo introdotto sul ramo diretto dello schema a blocchi del sistema; si noti che, fino a quando non si specificherà il contrario, tutti i discorsi introdotti di qui in poi riguarderanno esclusivamente il ramo diretto del sistema, non quello della retroazione, che avrà bisogno di una trattazione notevolmente diversa.

Si noti che, dal momento che spesso si trattano sistemi reali, definire precisamente la pulsazione del disturbo, ω_2 , è molto improbabile; quello che si definisce è un range di frequenze al quale ω_2 potrebbe appartenere, i cui bound inferiore e superiore sono rispettivamente $\underline{\omega_2}$ e $\overline{\omega_2}$. La larghezza di banda si può calcolare in diverse maniere; una classicamente utilizzata può riguardare la banda a - 3 dB. Date queste puntualizzazioni, è ora possibile trattare i contributi relativi a disturbi di tipo sinusoidale.

Dato un disturbo di tipo sinusoidale, non è possibile conoscere molte informazioni riguardo esso; potrebbe essere possibile conoscere a_2 , la larghezza dei bound, o altro; si vuole quantificare il contributo massimo di $d_2(t)$, in modo da progettare di conseguenza il sistema; l'errore, dunque, dovrà essere limitato:

$$|e_\infty^{d_2}| \leq \rho_2$$

Il contributo sull'uscita, quando agisce questo disturbo, è pari a:

$$\begin{aligned} E^{d_2}(s) &= Y^{d_2}(s) = G_{d_2}(s)D_2(s) = D_2(s) \cdot \frac{G_d(s)}{1 + G_t G_y G_c(s) A G_p(s)} = \\ &= G_d(s) \frac{1}{1 + G_a} D_2(s) = G_d(s) S(s) D_2(s) \end{aligned}$$

Questo fatto è piuttosto interessante; si supponga, per ora, per semplicità (cosa che in seguito nella trattazione non si ripeterà spesso) il fatto che $G_d(s) = 1$. Si può dire che:

$$E^{d_2}(s) = S(s) D_2(s)$$

A questo punto si riprende un risultato ben noto: la risposta a un segnale sinusoidale:

$$e_\infty^{d_2} = y^{d_2}(t) = a_2 \cdot |S(j\omega_2)| \sin(\omega_2 t + \psi)$$

Si deve avere che:

$$|e_\infty^{d_2}| < \rho_2 \implies a_2 |S(j\omega)| < \rho_2$$

Da qui, si può ricavare la seguente condizione:

$$|S(j\omega_2)| < \frac{\rho_2}{a_2} \quad \forall \omega_2 \in [\underline{\omega}_2; \overline{\omega}_2]$$

Se la frequenza del disturbo è in questo intervallo, la condizione deve valere su tutto l'intervallo. Dallo studio di $S(j\omega)$, si sa che, volendo che essa assuma valori bassi, bisogna considerare valori elevati del modulo, in modo che essa si possa costruire come il “simmetrico” del modulo su scala logaritmica, quindi permettere di assumere valori bassi. Perchè il modulo sia grande, è necessario che la pulsazione di passaggio del sistema, ω_c , sia molto avanzata: se la banda passante del sistema (strettamente imparentata con ω_c è molto elevata, “prima di ω_c ” ci sarà lo “spazio” per ottenere un guadagno molto elevato, di conseguenza un simmetrico “molto piccolo”, proprio come si può desiderare. Al fine di aumentare dunque la reiezione dei disturbi sinusoidali, il cui valore è fortemente legato a quello della funzione di sensibilità $S(s)$, è necessario avere $S(s)$ molto basse, e per far ciò avere una banda passante del sistema molto larga.

Si considera da adesso fino alla fine della sezione l'altro caso, che comunque per la trattazione e per l'uso delle nozioni che verrà fatto sarà meno utilizzato: il caso di disturbi sul ramo di retroazione. Osservando lo schema generale, i disturbi di questo tipo sono quelli identificati mediante la variabile $d_t(t)$; si considera dunque qualcosa dalla forma:

$$d_t(t) = a_t \sin(\omega_t t) \quad \forall \omega_t \leq \omega_t$$

Si noti che, per disturbi di questo tipo, non si considera un upper bound: essi sono, come vedremo, disturbi importanti soprattutto a frequenze elevate, dunque non si limitano le frequenze “in alto” dal momento che la trattazione mostrerà che un upper bound non è fondamentale. Osserviamo cosa capita, in presenza del solo $d_t(t)$, sfruttando i risultati finora proposti:

$$e_\infty^{d_t} = y^{d_t}(t)$$

Passando nel dominio di Laplace:

$$\begin{aligned} E^{d_t}(s) &= Y^{d_t}(s) = G_{d_t y}(s) D_t(s) = \frac{-G_y G_c(s) A G_p(s)}{1 + G_a(s)} D_t(s) = \\ &= \frac{1}{G_t} \cdot \frac{-G_a(s)}{1 + G_a(s)} D_t(s) = -\frac{1}{G_t} T(s) D_t(s) \end{aligned}$$

Questa volta, come si può notare, il problema è stato ricondotto all'altra funzione di sensibilità, quella complementare. Ciò comporterà sicuramente osservazioni finali differenti, per quanto concerne la reiezione del disturbo sinusoidale. Si osservi che l'errore, dunque, assumerà valori di questo genere:

$$e_{\infty}^{dt} = Y^{dt}(t) = a_t \cdot \frac{1}{G_t} |T(j\omega_t)| \sin(\omega_t t + \psi_t)$$

In modulo, l'errore deve essere inferiore a un certo valore:

$$|e_{\infty}^{dt}| \leq \rho_t \implies a_t \frac{1}{G_t} |T(j\omega_t)| \leq \rho_t$$

La condizione da soddisfare, questa volta, sarà:

$$|T(j\omega_t)| \leq \frac{\rho_t}{a_t} G_t \quad \forall \omega_t \geq \underline{\omega}_t$$

Cosa significa tutto ciò? Beh, semplice: la reiezione del disturbo sinusoidale sul ramo di retroazione si può ottenere imponendo valori molto bassi di $T(j\omega)$, nell'intorno della frequenza ω_t ; conoscendo l'andamento di $T(j\omega)$, è necessario che, da ω_t in poi, la funzione soddisfi la condizione. Dal momento che $T(j\omega)$ ha un andamento duale rispetto a $S(j\omega)$, è necessario che, da ω_t in poi, essa assuma valori al di sotto di uno ben prefissato. Questo si ottiene, come prima, lavorando sulla pulsazione ω_c di passaggio per l'asse 0 dB, ma in maniera opposta: è necessario che ω_c sia bassa, in modo che, dopo essa, la funzione di trasferimento assuma valori minori, in modo da poter attenuare sufficientemente $T(j\omega)$ (che segue l'andamento di $|G_a(s)|$) in modo da attenuare conseguentemente il segnale.

Si noti che sul ramo di reazione e sul ramo diretto si hanno condizioni assolutamente opposte: se da un lato per la reiezione del rumore sul ramo diretto è necessario allargare la banda del sistema, per la reiezione del rumore sul ramo di retroazione è necessario ridurre la banda del sistema. Da qua si può capire perchè non sono presenti upper bound su ω_t : non servono, dal momento che il sistema dovrà attenuare disturbi al di sopra di una certa frequenza, non appartenenti a una certa banda limitata in un intervallo. Il fatto che si debba esclusivamente limitare la banda in questo senso, ottenendo dunque condizioni di fatto opposte a quelle precedenti, implica il dover effettuare delle scelte in termini di blocchi utilizzati: tecnologicamente, i disturbi $d_t(t)$ derivano dall'uso di cattivi trasduttori sul ramo di retroazione; per quello che si vedrà, si supporrà molto spesso di aver a disposizione buoni sistemi di trasduzione, ogni qual volta non si vogliano considerare disturbi sul ramo di retroazione, in modo da introdurre solo condizioni sui disturbi sul ramo diretto.

4.5 La sensibilità alle variazioni parametriche

Una volta trattati i disturbi polinomiali e sinusoidali, ci si interesserà di un'altra categoria di disturbi, o meglio di "variazioni"; ci si pone sostanzialmente il seguente quesito:

Si supponga che vi siano disturbi di tipo parametrico, ossia disturbi riguardo la definizione dei blocchi $G(s)$ e $H(s)$ (si sappia fin da subito che essi verranno considerati separatamente), ossia la presenza di indeterminazione sulle caratteristiche dei blocchi, δG e δH . Sulle funzioni di trasferimento possono esserci variazioni di parametri che modificano le loro caratteristiche (in seguito verrà mostrato un esempio che permetterà di comprendere questo fatto); la domanda è: come cambia il comportamento del sistema di controllo, e la sua funzione di trasferimento?

Al fine di rispondere a questo quesito si definiscono le funzioni di sensibilità, differenziando due casi specifici: variazioni della sola G , e variazioni della sola H :

Variazioni della sola G

Si definisce la funzione di sensibilità $S_G^{G_{ry}}$ come la funzione di sensibilità della funzione di trasferimento del sistema retroazionato G_{ry} rispetto alla funzione di trasferimento del sistema da retroazionare ad anello aperto, G , come le variazioni di tipo relativo di G_{ry} rispetto a un valore "nominale", rispetto a loro volta le variazioni di G rispetto a un valore nominale:

$$S_G^{G_{ry}} = \frac{\frac{\partial G_{ry}}{G_{ry,nom}}}{\frac{\partial G}{G_{nom}}} = \frac{\partial G_{ry}}{\partial G} \cdot \frac{G_{nom}}{G_{ry,nom}}$$

Nel caso della struttura di controllo che viene utilizzata nella trattazione, facendo i conti con le derivate parziali, si può ricavare che:

$$S_G^{G_{ry}} = \frac{1}{1 + G_a}$$

Variazioni della sola H

Si definisce analogamente a prima la funzione di sensibilità $S_H^{G_{ry}}$ come la funzione di sensibilità della funzione di trasferimento del sistema retroazionato G_{ry} rispetto al blocco di retroazione H , come le variazioni di tipo relativo di G_{ry} rispetto a un valore "nominale", rispetto a loro volta le variazioni di H rispetto a un valore nominale:

$$S_H^{G_{ry}} = \frac{\frac{\partial G_{ry}}{G_{ry,nom}}}{\frac{\partial H}{H_{nom}}} = \frac{\partial G_{ry}}{\partial H} \cdot \frac{H_{nom}}{G_{ry,nom}}$$

Nel caso della struttura di controllo ormai abituale, essa vale:

$$S_H^{G_{ry}} = -\frac{G_a}{1 + G_a}$$

Che rappresentazioni hanno, in frequenza, queste due espressioni? Beh, il primo caso ha semplicemente la funzione di sensibilità, il secondo caso un'espressione simile alla funzione di sensibilità complementare. Ciò non è del tutto positivo, dal momento che i due risultati sono uno in contrasto con l'altro: per avere piccole variazioni su G , S deve essere piccola, cosa che si verifica a bassa frequenza; aumentando la frequenza S tende ad assumere valori elevati, cosa negativa. Ciò si ripercuote sulla funzione di trasferimento del sistema retroazionato:

$$\frac{\partial G_{ry}}{G_{ry,nom}} = S_G^{G_{ry}} \frac{\partial G}{G}$$

Una variazione di G rispetto al valore nominale, al valore per cui è stato progettato il sistema di controllo, comporta una variazione della funzione di trasferimento del sistema retroazionato, rispetto a quella nominale, dipendente da $S_G^{G_{ry}}$.

Volendo ridurre i disturbi su G , non si potranno ridurre quelli su H , dal momento che G e H variano con pesi determinati dalle funzioni di sensibilità tra loro complementari, che dunque avranno un comportamento complementare tra loro. Dal momento che fondamentale è considerare, nella trattazione, variazioni di G , considerano un H "ben fatto", sarà necessario "spendere" su H , dal momento che servirà un oggetto in grado di non introdurre disturbi di alcun tipo.

Esempio pratico

Si consideri il seguente circuito basato sull'uso di amplificatore operazionale:

Precedentemente è stato proposto un buon modello per l'amplificatore operazionale ad anello aperto:

$$A(s) = \frac{A_0}{\left(1 + \frac{s}{\omega_1}\right) \left(1 + \frac{s}{\omega_2}\right)}$$

Si consideri un valore nominale di A_0 pari a 10^5 ; si consideri il fatto (molto, molto irrealistico, nel senso di molto pessimistico) che A_0 possa variare dal valore nominale fino a diventare 10^7 :

$$A_0 \in [10^5; 10^7]$$

Dati:

$$\omega_1 \simeq 10 \text{ rad/s}; \quad \omega_2 \simeq 10^6 \text{ rad/s}; \quad R_1 = 1 \text{ k}\Omega; \quad R_2 = 10 \text{ k}\Omega$$

Ci si chiede: nel caso vi sia una grossa perturbazione sul guadagno, fino a raggiungere il valore di 10^7 , causata ad esempio dalla sostituzione dell'amplificatore operazionale in uno più moderno, di quanto cambia la funzione di trasferimento ad anello chiuso, $G_{ry}(s)$?

Si calcola la variazione di A_0 rispetto al valore nominale:

$$\frac{\delta A_0}{A_{0,nom}} = \frac{10^7 - 10^5}{10^5} \simeq 10^2 = 10000\%$$

Si ha dunque una variazione relativa, ad anello chiuso, pari al 10000 %, dunque molto, molto elevata. Considerando la funzione di trasferimento unitaria, si avrà:

$$H = \frac{R_1}{R_1 + R_2} \simeq 0,1$$

Dunque, la funzione di anello, $A \cdot H$, sarà semplicemente A normalizzata per un fattore 10; si costruisce a questo punto la funzione di sensibilità, a partire dalla funzione di anello, dove la funzione di anello è coincidente alla funzione del sistema, normalizzata di 10 (rispetto al valore nominale, partendo dunque da 10^4 anzichè 10^5):

Ciò che si vuole mettere in evidenza è: come varia G_{ry} rispetto al valore nominale quando si usa la funzione di sensibilità rispetto ad A , ossia quella che prevede variazioni del 10000 %? Si supponga di utilizzare questo circuito a bassa frequenza; in questo caso, dal momento che $|S(j\omega)|$ si costruisce come la simmetrica della funzione di anello, essa assumerà valori pari al reciproco di 10^4 , ossia 10^{-4} :

$$S_A^{G_{ry}} = 10^{-4}$$

La variazione della funzione di trasferimento del sistema retroazionato sarà:

$$\frac{\partial G_{ry}}{G_{ry,nom}} = S_A^{G_{ry}} \cdot \frac{\delta A_0}{A_{0,nom}} = 10000\% \cdot 10^{-4} = 1\%$$

Cosa significa questo risultato? Nonostante il caso estremamente patologico (che non si verifica, in pratica, sostanzialmente mai), la retroazione ha un effetto tale da attenuare notevolmente la variazione parametrica della funzione di trasferimento ad anello aperto, riducendo di fatto la sensibilità ai disturbi di tipo parametrico.

Poichè progettando sistemi di controllo con guadagno alto su bassa frequenza si è interessati a trattare la funzione $S(j\omega)$ per basse frequenze, dal momento che si tratta esclusivamente “il ramo diretto”, si suppone che il trasduttore su H sia molto valido, ossia che i resistori siano di qualità eccellente, non subiscano variazioni importanti durante il funzionamento del sistema.

4.6 Introduzione ai problemi di laboratorio

In questa sezione verranno analizzati alcuni aspetti di progetto, che poi potranno essere completati mediante simulazioni in sede di laboratorio. Si considererà in sostanza un esempio pratico di come operare, e verranno proposti gli esempi di calcolo per trattare le specifiche finora studiate.

Si richiede la progettazione di un sistema di controllo, considerando (come si farà di solito) $G_R = 1$; si considerano d_1 e d_2 ingressi polinomiali, con questo andamento:

$$d_1(t) = \gamma_1$$

$$d_2(t) = \gamma_2 \cdot t$$

d_1 è un ingresso a gradino, d_2 a rampa. Per rapporto di scala si intende sostanzialmente il guadagno stazionario del sistema, ossia il K_d . Lo scopo dell’esercizio è sostanzialmente ricavare G_c e G_y . Verranno ora affrontate le quattro richieste del problema, in modo da soddisfarle tutte contemporaneamente.

Specifica 1

Si ha un rapporto di scala unitario; si ricava l’espressione di G_y a partire dalle altre, come:

$$K_d = 1 \implies H = \frac{1}{K_d} = G_t \cdot G_y$$

Dato che il guadagno del trasduttore è noto, si può ricavare G_y come:

$$G_y = \frac{1}{K_d \cdot G_t} = 1$$

Specifica 2

Si chiede di quantificare e limitare l'errore in regime permanente causato dalla fedeltà di risposta, dunque causato dall'errore sul riferimento r . Nel caso dell'esercitazione, r è un segnale a rampa, dunque l'errore sarà sicuramente finito e non nullo. Come in ogni ambito dell'ingegneria, le specifiche di progetto devono essere rispettate e in maniera assolutamente "stretta": è richiesto che l'errore in presenza di rampa unitaria sia finito e non nullo, dunque il sistema deve essere di tipo "1".

Ci si può fare la seguente domanda: conoscendo il tipo di sistema, nel controllore G_c , è necessario introdurre poli nell'origine? Beh, la risposta è semplice: il sistema è di tipo 1; si ha tuttavia che:

$$G = G_c \cdot A \cdot G_p$$

G ha bisogno di un polo nell'origine, ma esso è già presente in G_p , come si può leggere dalle specifiche, dunque nel controllore progettato, G_c , non sarà necessario introdurre alcun polo. Come si sa, infatti, nella sua forma più generale $G_c(s)$ ha un'espressione del tipo:

$$G_c(s) = \frac{K_c}{s^r}$$

In questo caso, $r = 0$, per i motivi sopra citati.

Dalla tabella delle specifiche per i riferimenti, si ha che:

$$|e_\infty^r| = \left| \frac{K_d^2}{K_v} R_0 \right| \leq 0,15 \longrightarrow \frac{1}{K_v} \leq 0,15$$

Si ricordi che:

$$K_v = \lim_{s \rightarrow 0} s \cdot G(s)$$

Da qui:

$$K_v = \lim_{s \rightarrow 0} s \cdot G_c(s) \cdot A \cdot G_p(s) = AK_c \frac{25}{2}$$

Dunque:

$$|e_\infty^r| = \left| \frac{1}{AK_c \frac{25}{2}} \right| \leq 0,15 \implies |K_c| \geq 5,614$$

Specifica 3

Per specifiche sui disturbi, è fondamentale un accorgimento: non usare mai assolutamente e per nessun motivo la tabella dei riferimenti: essa è valida solo e soltanto per quanto riguarda i riferimenti.

Si ha che:

$$d_1(t) = \gamma_1, \quad |\gamma_1| < 5,5 \cdot 10^{-3}$$

Si scrive a questo punto l'espressione dell'errore e se ne calcola il contributo:

$$E_\infty^{d_1} = Y^{d_1}(s) = \frac{G_p(s)}{1 + G_t \cdot G_y \cdot G_c \cdot A \cdot G_p} D_1(s)$$

L'errore nel tempo si può calcolare mediante il teorema del valore finale:

$$|e_\infty^{d_1}| = \lim_{s \rightarrow 0} s E^{d_1}(s) = \lim_{s \rightarrow 0} s \frac{\gamma_1}{s} \frac{G_p(s)}{1 + G_t \cdot G_y \cdot G_c \cdot A \cdot G_p} D_1(s)$$

A questo punto, una piccola variante sul tema: si supponga di non aver precedentemente dedotto il tipo del controllo (cosa fatta per quanto riguarda la specifica sul riferimento); si potrà dire che il valore di r tale per cui il limite è finito e non nullo sarà 0, anche da qua. Facendo infatti i conti:

$$\implies \frac{G_p(0)}{1 + G_t \cdot G_y \cdot G_c \cdot A \cdot G_p(0) K_c} \frac{\max\{\gamma_1\}}{s} \leq 0,015$$

Da qui si può ricavare che:

$$K_c \simeq 3,86$$

Si noti un fatto: le ipotesi e le specifiche sono trattate rigorosamente in maniera separata; il valore di K_c da scegliere, al termine del progetto, sarà quello più elevato, ossia quello in grado di soddisfare contemporaneamente tutte le specifiche.

Specifica 4

Si considera a questo punto l'ultima specifica, quella che richiede la trattazione dell'errore concernente il disturbo a rampa, d_2 :

$$E^{d_2}(s) = Y^{d_2}(s) = D_2(s) \cdot \frac{G_d(s)}{1 + G_t G_y G_c A G_p}$$

Da qua:

$$|e_{\infty}^{d_2}| = \lim_{s \rightarrow 0} s \cdot Y^{d_2}(s) = \frac{3,5 \cdot 10^{-3}}{s^2} \cdot \frac{1}{\frac{25}{2} K_c \cdot 0,095} \leq 7,5 \cdot 10^{-4}$$

Si risolve questa disequazione e si riesce dunque a completare l'ultima specifica, ricavando K_c .

4.7 La risposta transitoria di sistemi di controllo con retroazione

Una volta affrontato il problema dello studio della fedeltà di risposta nel caso di particolari ingressi, quali regime permanente con riferimenti polinomiali, si può passare allo studio del transitorio.

Un'analisi formale della risposta in transitorio del sistema è estremamente complicata, rispetto al caso del regime permanente; verranno dunque introdotte e considerate scorciatoie, ipotesi semplificative che permetteranno di districarci in maniera molto più facile in questi percorsi. Fondamentalmente, le ipotesi semplificative che verranno sempre considerate sono due:

- La risposta al transitorio verrà considerata sempre per quanto riguarda un ingresso a gradino: esso è di fatto il segnale di ingresso più “esigente” all'origine, per quanto riguarda il transitorio, dal momento che la discontinuità all'origine comporta una notevole “fatica”, un notevole ostacolo da saltare, per il sistema;
- Si considereranno i sistemi progettati come assimilabili a un sistema “prototipo”, introducendo come prototipo una dinamica dominante del secondo ordine; ciò equivale a dire che ciò che regola il transitorio è una coppia di poli complessi coniugati; vengono trascurati tutti i contributi di altri eventuali poli presenti nel sistema.

Esempio pratico 1

Si consideri un sistema, composto da elementi ideali, di questo tipo:

Come è ben noto, l'equazione che lega ingresso e uscita è qualcosa di questo genere:

$$y = \frac{R_2}{R_1 + R_2} u$$

Il legame tra ingresso e uscita è statico; ciò significa che, introducendo un ingresso a gradino, R_0 , la reazione del sistema, l'uscita, sarà immediata e

non avrà fenomeni transitori di alcun tipo; l'uscita sarà identica all'ingresso, a meno di un fattore di scala dettato dal partitore.

Esempio pratico 2

Si consideri un nuovo esempio:

Si potrebbe vedere, con il modello più approssimativo, che:

$$\frac{y}{u} \simeq 1 + \frac{R_3}{R_4}$$

Osservando tuttavia su di un oscilloscopio digitale l'andamento dell'uscita, essa presenterebbe un transitorio:

La risposta ha un andamento di questo tipo, ossia non segue prontamente l'ingresso; questo perchè, in un modello di questo tipo, non ci considerano effetti transitori che di fatto non sono trascurabili, in questo genere di casistiche. Il legame statico è insufficiente a rappresentare il transitorio di un sistema; volendo considerarlo, è necessario introdurre una modellizzazione dinamica del sistema, in grado di considerare anche fenomeni di questo tipo.

4.7.1 Caratterizzazione generale nel tempo

Considerando il problema della risposta a gradino nel transitorio, ci si aspetta che il sistema risponda in un certo modo, dato l'ingresso per ipotesi a gradino.

Ciò che si può notare è che l'uscita potrà presentare andamenti di diversi tipi; per questo motivo, fondamentale è l'introduzione di un certo insieme di parametri in grado di caratterizzare la risposta nel tempo, in modo da poter effettuare operazioni sia di analisi sia di progetto. Verranno ora presentati alcuni parametri.

- Tempo di salita: esistono due fondamentali definizioni di questo parametro: una, che non utilizzeremo, calcola o misura il tempo impiegato per passare dal 10% al 50% del valore a regime. Quella che verrà utilizzata è la seguente: “il tempo di salita è il tempo impiegato a raggiungere per la prima volta il valore di regime permanente”.
- Sovraelongazione: si definisce come il valore massimo meno il valore in regime permanente, normalizzato per quest'ultimo (si parla di “sovraelongazione relativa”):

$$\hat{s} = \frac{y_{max} - y_{\infty}}{y_{\infty}}$$

La sovraelongazione fornisce un'idea di quanto i poli complessi coniugati siano vicini all'asse immaginario, $j\omega$: più lo smorzamento ξ è piccolo, più il sistema è instabile, dunque la sovraelongazione grossa;

- Tempo di assestamento: tempo impiegato per la risposta a rientrare in una fascia a $\pm\varepsilon$; di solito si utilizzano valori del 1%, 2%, 3%, 4%, 5%, rispetto al valore a regime.

4.7.2 Parametri della risposta al gradino di sistemi prototipo del secondo ordine

Si consideri un sistema di questo genere:

In questo caso, si può calcolare rapidamente il fatto che:

$$G_{ry}(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

Questo è il sistema chiamato “sistema prototipo del II ordine”.

Si considerino dunque sistemi di questo tipo: introducendo un gradino unitario, si avrà qualcosa del tipo:

$$Y(s) = G_{ry}(s) \cdot U(s), \quad U(s) = \frac{1}{s}$$

Si può dimostrare, mediante lunghi conti, che antitrasformando, si ottenga un'espressione di questo genere, nel dominio del tempo:

$$y(t) = 1 - \frac{1}{\sqrt{1-\xi^2}} e^{-j\omega_n t} \sin \left\{ \omega_n t \sqrt{1-\xi^2} + \arctan \frac{\sqrt{1-\xi^2}}{\xi} \right\}$$

Si può dimostrare, derivando quest'espressione e cercandone il massimo, che:

$$\hat{s} = e^{-\frac{\pi\xi}{\sqrt{1-\xi^2}}} = f_1(\xi)$$

Questa formula si può utilizzare sia in fase di analisi sia in fase di progetto: date specifiche, di fatto, utilizzando questa formula si può progettare la posizione dei poli complessi coniugati, invertendo la formula e trovando:

$$\xi = \sqrt{\frac{\left(\frac{\ln(\hat{s})}{\pi}\right)^2}{1 + \left(\frac{\ln(\hat{s})}{\pi}\right)^2}}$$

Si può osservare che, dato un certo valore \hat{s}_0 , dire che \hat{s} è inferiore a questo valore, significa dire che lo smorzamento dei poli complessi coniugati, ξ , sia superiore a un certo valore ξ_0 . Un criterio per posizionare i poli/autovalori è proprio quello di posizionarli in un settore di questo tipo:

Si può trovare un collegamento tra smorzamento e tempo di salita:

$$t_s = \frac{1}{\omega_n \sqrt{1 - \xi^2}} \left\{ \pi - \arctan \frac{\sqrt{1 - \xi^2}}{\xi} \right\}$$

Si può dunque dire che:

$$t_s \cdot \omega_n = f_2(\xi)$$

Per quanto riguarda il tempo di assestamento, invece:

$$t_a \simeq -\frac{\ln(\varepsilon)}{\omega_n \xi} \longrightarrow \omega_n t_a = f_3(\varphi, \varepsilon)$$

Si noti che spesso ε è fornito in termini di percentuale; il valore da utilizzarsi in questo ambito è sempre il valore relativo, ma numerico; si normalizza per 100 il valore percentuale, ottenendolo così.

4.8 Risposta in frequenza di un sistema di controllo

La risposta in frequenza finora è stata studiata solo per quanto riguarda il regime permanente; essa però contiene informazioni che possono tornare utili anche per il solo transitorio. Studiamo alcune nozioni riguardo a ciò, in modo da poter introdurre una caratterizzazione nel dominio della frequenza dei sistemi di controllo, come appena fatto per quanto riguarda il dominio del tempo.

4.8.1 Caratterizzazione generale in frequenza

Si consideri il seguente andamento per quanto riguarda la risposta in frequenza della funzione di sensibilità complementare, $T(j\omega)$:

I parametri fondamentali che identificano questa funzione sono T_p , ossia l'altezza del massimo della funzione, ω_c , ossia la frequenza di passaggio per l'asse 0 dB, ω_p , ossia la pulsazione del picco di ampiezza T_p , ω_B , ossia la frequenza di taglio a - 3 dB (sia per la funzione di anello G_a sia per T , tendenzialmente).

Gli stessi parametri possono essere considerati anche per quanto riguarda la funzione di sensibilità e la sua risposta in frequenza, $S(j\omega)$:

In questo caso il massimo della funzione si identifica con S_r ; la pulsazione nella quale esso è localizzato, si indica mediante ω_r .

Come si vedrà tra breve questi parametri, per ora solo introdotti, saranno molto importanti sia in sede di analisi sia in sede di progetto; si introdurrà inoltre tra breve un sistema per stimarli e utilizzarli.

4.8.2 Parametri della risposta in frequenza di sistemi prototipo del secondo ordine

Come già fatto per quanto riguarda il dominio del tempo, esistono relazioni in grado di determinare alcuni parametri a partire da altri; si propongono a questo punto alcune espressioni in grado di quantificare alcuni parametri delle risposte in frequenza, espressioni che di fatto non verranno mai utilizzate, dal momento che si utilizzeranno metodi grafici presto presentati. Si può dimostrare che:

$$\omega_p = \omega_n \cdot \sqrt{1 - 2\xi^2} = \omega_n f_4(\xi)$$

$$T_p = \frac{1}{2\xi\sqrt{1 - \xi^2}} = f_5(\xi)$$

$$\omega_B = \omega_n \sqrt{1 - 2\xi^2 + \sqrt{2 - 4\xi^2(1 - \xi^2)}} = \omega_n f_6(\xi)$$

Si può dimostrare che, generalmente, le altre funzioni han espressioni del tipo:

$$\omega_c = \omega_n \cdot f_7(\xi)$$

$$\omega_r = \omega_n f_8(\xi)$$

$$S_r = f_9(\xi)$$

Non si riportano alcune delle formule, anche tra le più importanti; questo per evidenziare il fatto che esse non sono molto utili, per l'approccio che si intende utilizzare in questa trattazione. Quello che si vuole mettere per ora in evidenza è il seguente concetto: nei sistemi prototipi del secondo ordine, una volta individuati i parametri necessari per la caratterizzazione del sistema, essi possono essere "convertiti" nel dominio della frequenza; le specifiche vengono sovente introdotte nel dominio del tempo, poichè è

possibile riscontrare più rapidamente effetti pratici di determinate scelte; dal momento che è tuttavia più facile lavorare nel dominio della frequenza, sarà necessaria questa operazione di “conversione”, tra breve esposta. Quasi ogni volta le prestazioni, le specifiche, dunque, saranno date nel dominio del tempo; nel resto della trattazione si imparerà a progettare controllori nel dominio della frequenza, a partire dalla traduzione di queste specifiche.

4.9 Relazioni tra la risposta al gradino e la risposta in frequenza nei sistemi prototipo del secondo ordine

Nella sezione precedente sono già stati evidenziati alcuni legami tra alcuni parametri nel dominio del tempo e nel dominio della frequenza. Si vuole esporre a questo punto, in maniera maggiormente dettagliata, come conviene operare per effettuare la conversione; in seguito si introdurrà un esempio pratico che chiarificherà ulteriormente.

Sono stati introdotti parametri in grado di caratterizzare la risposta nel dominio del tempo del sistema all'eccitazione di un segnale a gradino, e parametri in grado di caratterizzare la risposta in frequenza. Si possono fare, a partire dalle introduzioni finora proposte, alcune considerazioni, quale ad esempio la seguente: si consideri la seguente espressione:

$$\omega_B t_s = f_2(\xi) \cdot f_6(\xi)$$

Esistono legami tra dominio del tempo e dominio della frequenza, quali ad esempio quello appena proposto: si sa che, per avere un tempo di salita elevato, ossia una durata breve del transitorio, è necessaria una banda molto elevata per il sistema: banda elevata significa grossa ampiezza spettrale, dunque presenza di sinusoidi molto “veloci”, in grado di permettere rapide variazioni delle ampiezze, nel dominio del tempo.

Ai fini di caratterizzare un sistema questa può essere un'idea valida, ma non sicuramente sufficiente e soddisfacente. Esistono, come si può osservare dalle altre espressioni, altre osservazioni effettuabili: ω_c e ω_B per esempio sono distanziate da un coefficiente prossimo a 0,67.

Si potrebbero estrarre molte altre idee, che però non sarebbero sufficienti al fine di introdurre un metodo di analisi e di progetto completo. Ciò che serve realmente, a questo punto, è qualcosa in grado di collegare univocamente un parametro nel tempo a uno nella frequenza.

Vengono incontro, a tal fine, alcuni grafici, uno su tutti il seguente:

Il punto di partenza per il progetto, per la conversione da parametri dal dominio del tempo a dominio della frequenza, sarà tendenzialmente sempre questo grafico: un'associazione biunivoca tra valori dello smorzamento di poli di un sistema a dinamica dominante del secondo ordine e valore della massima sovraelongazione. Data dunque la sovraelongazione nel dominio del tempo, è possibile, mediante questo grafico, ricavare il valore minimo di ξ necessario per soddisfare la specifica.

Si tratta di un grosso punto di partenza; si osservi il seguente grafico:

A partire dall'informazione ricavata sullo smorzamento è possibile ricavare molte altre informazioni, quali ad esempio i valori di T_p , S_r , a partire dai quali è possibile procedere con il progetto. Conoscendo altre specifiche nel dominio del tempo è possibile ricavare, a partire da esse, dallo smorzamento ξ e da altri grafici, quali:

Tutte le altre informazioni in grado di caratterizzare la risposta in frequenza del sistema.

Esempio teorico/pratico

A cosa serve tutto ciò che è stato finora detto? Beh, si propone un primo esempio teorico/pratico, in grado quantomeno di dare meglio l'idea di ciò che è stato finora detto. Un sistema fisiologico che può essere molto interessante studiare è l'occhio umano, e nella fattispecie i suoi movimenti. Tra i vari movimenti che l'occhio può fare, esiste il cosiddetto movimento "saccodico" dell'occhio: quando si legge una riga e si arriva al termine di essa, istintivamente, l'occhio si muove velocemente in modo per arrivare all'inizio della riga successiva. La domanda che ci si pone è: quanto tempo ci impiega l'occhio per effettuare un movimento di questo tipo?

Si introduce la seguente modellizzazione:

Si ha a che fare con un'equazione di questo tipo, considerando il momento di inerzia J dell'occhio, l'attrito viscoso β e l'elasticità dei muscoli K , si ha un'equazione del tipo:

$$\tau(t) = J\ddot{\vartheta} + \beta\dot{\vartheta} + K\vartheta$$

Questo è un modello in grado di descrivere, in prima approssimazione, il movimento dell'occhio. I parametri J , β e K possono essere stimati, mediante metodi di vario genere, in un paziente; nel dominio di Laplace si ha un'espressione di questo tipo:

$$\frac{\Theta(s)}{T(s)} = \frac{1}{s^2 J + s\beta + K} = \frac{1}{K} \frac{\frac{K}{J}}{s^2 + \frac{\beta}{J}s + \frac{K}{J}}$$

Questa equazione è nella forma classica del sistema a dinamica dominante del secondo ordine:

$$G(s) = K \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

I procedimenti di stima portano a trovare, a partire dai valori nel tempo esposti, valori in frequenza di questo genere:

$$\xi = 0,7 \quad \omega_n = 120 \text{ rad/s}$$

La domanda a questo punto è: quanto impiega il “sistema di controllo” dell’occhio a muoverlo, a fine riga? Beh, è un classico problema di studio del tempo di salita: di fatto il movimento rapido è assimilabile a un gradino; lo studio del tempo coincide col tempo di salita. Si può ricavare, a partire dai grafici (come si può provare a fare), che:

$$t_s \cdot \omega_n = 3,3 \text{ rad}$$

Dunque:

$$\frac{3,3 \text{ rad}}{120 \text{ rad/s}} = 27 \text{ ms}$$

4.10 Curve di T e S a modulo costante

4.10.1 Curve di T a modulo costante

Si presenta a questo punto uno strumento che una volta veniva utilizzato molto spesso, ora è caduto un po’ in disuso. Molte tecniche di controllo, per quanto riguarda i sistemi elettronici, sono nate per quanto riguarda gli amplificatori: spesso, a certe condizioni, gli amplificatori diventavano oscillatori, ossia uscivano da uno stato di stabilità. Al fine di studiare la stabilità Nyquist formulò il criterio precedentemente esposto, Bode altri, e così via. Ora si introdurrà un ulteriore criterio, atto a verificare in altri modi ancora la stabilità dei sistemi.

Si consideri il solito sistema di controllo a reazione unitaria:

Con la funzione di anello e il suo diagramma di Nyquist è possibile trarre conclusioni riguardo la stabilità della funzione di trasferimento del sistema retroazionato. Si vuol vedere, ora, come da un grafico della funzione di anello sia possibile desumere un grafico della funzione del sistema retroazionato, G_{ry} ; data $G_a(s)$ e dunque $G_a(j\omega)$, si può dire che essa sia un generico numero complesso, dunque sia rappresentabile in forma cartesiana:

$$G_a = A + jB$$

Come si sa, la funzione di anello e la funzione di trasferimento sono legate dalla seguente relazione:

$$G_{ry} = \frac{G_a}{1 + G_a} = \frac{A + jB}{1 + A + jB} = T$$

Si consideri un particolare valore della funzione T , T_{p0} , e nella fattispecie di esso il suo modulo quadro, $|T_{p0}|^2$:

$$|T_{p0}|^2 = \frac{A^2 + B^2}{(1 + A)^2 + B^2}$$

Quali sono i particolari valori di A e B tali per cui si ottenga $|T_{p0}|^2$? Beh, risolvendo l'equazione, si può ottenere:

$$\left(A + \frac{|T_{p0}|^2}{|T_{p0}|^2 - 1} \right)^2 + B^2 = \left(\frac{|T_{p0}|}{|T_{p0}| - 1} \right)^2$$

Come si può banalmente osservare, questa è una circonferenza di raggio pari a r :

$$r = \frac{|T_{p0}|}{|T_{p0}| - 1}$$

E centrata nel punto:

$$C = \left(-\frac{|T_{p0}|^2}{|T_{p0}|^2 - 1}, 0 \right)$$

Vengono spesso proposti grafici di questo tipo:

Accanto a ogni circonferenza è scritto un certo valore di $|T_{p0}|$: ciò caratterizza le circonferenze che danno luogo a $|T|$ costante. Il dominio nel quale è rappresentata questa circonferenza, per quanto sia cartesiano e dunque riconducibile a un banale piano di Nyquist (o di Gauss), è noto in questo ambito come piano di Hall. La costruzione dell'analisi in questo ambito si effettua nel seguente modo: sul grafico si appuntano modulo e fase della funzione di anello, e la ω_n cui è associato il punto; in ingresso si userà G_a in coordinate polari dunque, mentre in uscita si avrà l'andamento di $|T|$: sfruttando la vicinanza tra la curva di G_a e le circonferenze per $|T|$ costanti, si riesce a rappresentare il diagramma di Bode semplicemente rilevando i valori. Quando si trova una circonferenza tangente alla curva di G_a , quello sarà il punto di massimo del diagramma.

4.10.2 Curve di S a modulo costante

Un ragionamento analogo a quello appena introdotto è effettuabile anche per quanto riguarda la funzione $|S|$:

$$S = \frac{1}{1 + G_a} \longrightarrow |S|^2 = \frac{1}{(1 + A)^2 + B^2}$$

Dato un particolare valore, S_{r0}^2 del modulo di S , si può dire che:

$$S_{r0}^2 = \frac{1}{(1 + A)^2 + B^2}$$

Si trova dunque il seguente luogo geometrico:

$$(A + 1)^2 + B^2 = \frac{1}{S_{r0}^2}$$

Ossia una circonferenza di raggio pari al reciproco di S_{r0} , centrata nel punto $(-1, 0)$.

Si noti un fatto estremamente interessante riguardo ciò, che va a considerare un concetto precedentemente introdotto: il punto in cui è sempre centrata qualsiasi circonferenza rappresentante un valore a modulo costante sul piano di Hall di $|S|$, è esattamente coincidente con il punto critico considerato per la definizione dei margini di guadagno e di fase per quanto concerne il diagramma di Nyquist. Questo diagramma funziona in maniera analoga rispetto al precedente: data come ingresso G_a , si può ottenere come uscita il diagramma di $|S(j\omega)|$, ossia la risposta in frequenza della funzione di sensibilità.

Questi due sottocapitoli sono serviti per introdurre e spiegare almeno in maniera superficiale le curve di $|S|$ e $|T|$ a modulo costante; esse in pratica vengono utilizzate nella seguente maniera: scritto il diagramma polare sulla carta, per ogni punto che si introduce sul diagramma si appunta la pulsazione; si riporta dunque il valore dei vari punti su di un diagramma di Bode, tracciando di fatto quelle curve. Presto queste nozioni verranno riprese ed estese, in ambito di una nuova rappresentazione, molto più utile di questa.

4.11 Indicatori di margini di stabilità

In precedenza, come già accennato, si è parlato dell'importanza dello studio dei margini di guadagno, confrontando con il punto “-1” qualsiasi valore.

Queste tecniche sono relativamente vecchie, e, rispetto a quelle basate sull'uso di calcolatori, potrebbero sembrare anche sorpassate. Ciò non è così:

negli anni '80 / '85 si è iniziato a parlare di problemi di robustezza per quanto riguarda il controllo; i recenti cambiamenti in ambito di controllistica sono partiti proprio su osservazioni sul diagramma di Hall e sul piano di Nyquist:

Si veda questa figura: se la funzione di trasferimento presentata sul piano di Nyquist avesse una “gobba”, di fatto non si avrebbe alterazione sui margini di stabilità rispetto al caso in cui la funzione fosse più “liscia”, ma la condizione di stabilità sarebbe meno veritiera. Il caso di questo tipo di curvatura non è per niente raro: se il sistema presenta una risonanza, cosa frequente in molti sistemi meccanici dove si hanno poli complessi poco smorzati, potrebbero vedersi andamenti di questo genere.

Margini di fase e di guadagno non sono parametri sufficienti per una caratterizzazione completa dello studio della stabilità di un sistema: garantire il fatto che siano soddisfatti non garantisce il fatto che il sistema abbia “una certa stabilità”.

Si può introdurre un buon criterio: dal grafico della funzione di sensibilità, si può vedere che il picco deve essere il più piccolo possibile: un indicatore dei margini di stabilità, nella fattispecie, è proprio il picco massimo della funzione di sensibilità.

Cerchiamo di capire meglio questa affermazione, mediante le seguenti osservazioni:

Si consideri x il vettore distanza tra il punto critico, -1 , e il punto in questione della funzione di anello, G_a (quello sulla circonferenza a raggio unitario); si può dire, facendo la somma vettoriale, che:

$$-1 + x = G_a \longrightarrow x = 1 + G_a$$

Da qui, si può vedere banalmente che:

$$\frac{1}{x} = \frac{1}{1 + G_a} = S$$

Cosa ci dice tutto ciò? Beh, se x è grande, il vettore x tende a “spingere” in là la “gobba”. S sarà conseguentemente piccola, dunque il picco sarà ridotto, e i margini di guadagno e fase saranno significativi.

Per questo motivo, in qualche maniera, tra le specifiche se ne troverà (in maniera indiretta) una del tipo:

$$|S_r| \leq S_{r0}$$

Come visto prima: si cercherà di limitare il valore del modulo del picco a un certo valore; ciò permetterà di aumentare la stabilità del sistema retroazionato.

4.12 Esempio (traduzione di $\hat{s} \leq \hat{s}_0$)

Si considera a questo punto un esempio pratico di traduzione di specifica dal dominio del tempo a dominio della frequenza, per poi introdurre delle formule interessanti atte a calcolare i margini di guadagno e fase a partire da altre.

Si progetti un sistema in cui la sovralongazione relativa abbia un valore inferiore o uguale al 10% del valore a regime.

Di questo sistema non si sa niente, dunque si suppone che esso abbia una dinamica dominante del secondo ordine; da qui, utilizzando il primo grafico, si può imporre una sovralongazione pari a 0,1, e ricavare (graficamente) il fatto che:

$$\xi \geq 0,59$$

Si supponga di utilizzare il valore limite; si possono ricavare, mediante gli altri grafici, informazioni sui moduli dei valori di picco di T e S :

$$|T_{p0}| = 1,05$$

$$|S_{r0}| = 1,35$$

A questo punto si hanno le informazioni necessarie per calcolare margini di guadagno e margini di fase; nella fattispecie, sarà necessario risolvere le seguenti disequazioni:

$$m_G \geq \frac{1}{T_p} + 1$$

$$m_\varphi = 2 \arcsin \left(\frac{1}{2T_p} \right)$$

Oppure:

$$m_G \geq \frac{1}{S_r - 1} + 1$$

$$m_\varphi = 2 \arcsin \left(\frac{1}{2S_r} \right)$$

Dei valori ricavati si considerano solo quelli più stringenti, i “più larghi”: essi saranno, per il sistema, i margini di fase minimi necessari al fine di garantire il fatto che le specifiche relative alla sovralongazione massima siano consentite. Si noti che questi margini non sono assolutamente relativi allo

studio della stabilità del sistema, bensì ad una specifica sul transitorio. La stabilità è una naturale conseguenza della soddisfazione di questi margini, tuttavia questa è una condizione sufficiente, ma non necessaria, al fine dello studio della stabilità del sistema.

4.13 La carta di Nichols

Finora è stata introdotta una caratterizzazione delle curve a modulo costante sul piano di Hall per quanto riguarda $|T|$ e $|S|$; esiste tuttavia una variante rispetto a questa rappresentazione, basata sull'uso di alcuni accorgimenti e soprattutto su di una scala semilogaritmica: la carta di Nichols.

Sul piano di Hall si ha purtroppo un grosso difetto, per quanto riguarda la rappresentazione delle curve: esse sono piuttosto limitate, ossia sono in grado di rappresentare range di valori relativamente stretti, limitati: non è possibile rappresentare curve troppo piccole o troppo grosse. A questo tipo di rappresentazione, come già suggerito, viene incontro la rappresentazione logaritmica: si utilizza un piano in cui le ascisse sono rappresentative di una fase, mentre le ordinate di un modulo in decibel (dB). Su questo piano l'ingresso, ciò che va inserito, al solito è la risposta in frequenza della funzione di anello, $G_a(j\omega)$; in uscita, si spiegherà tra breve cosa si avrà.

Si consideri prima di tutto un piccolo esempio pratico, atto a migliorare la comprensione di ciò che è stato finora introdotto. Si consideri il caso di un integratore, ossia di un dispositivo la cui risposta in frequenza ha un andamento di questo genere:

$$G_a(j\omega) = \frac{1}{j\omega}$$

La fase è costante a -90° , mentre il modulo, inizialmente dotato di valori estremamente elevati, tende a scendere, a decrescere all'aumentare della pulsazione ω considerata.

Si consideri un altro esempio pratico: si supponga di avere una funzione di questo tipo:

$$G_a(j\omega) = \frac{1}{j\omega(1 + j\omega\tau_1)}$$

In questo caso disegnare la curva è un poco più complicato. Per farlo, tendenzialmente, si può vedere che per $\omega \rightarrow 0$, la fase vale -90° ; raccogliendo un certo insieme di punti, appuntando i valori delle pulsazioni ω e rappresentando sul diagramma modulo e fase per ciascun punto, quindi congiungendoli in modo da formare l'andamento della curva. Ci si può aspettare che, per

pulsazioni elevate, il modulo della funzione tenda a 0 (dunque a $-\infty$ dB), la fase a -180° , per il contributo dei due poli.

Si noti un fatto: il punto centrale della carta di Nichols ha modulo “1” e fase “ -180° ”. Questo non è un punto casuale, dal momento che è già stato visto più e più volte e con un significato ben preciso: quello di “punto critico”. Ci si può aspettare dunque che, visivamente, dalla carta di Nichols si possano estrarre informazioni molto importanti per quanto riguarda la stabilità dei sistemi retroazionati, a partire dal diagramma della risposta in frequenza della funzione di anello.

Fondamentale nella fattispecie è il seguente fatto: per i sistemi a rotazione di fase minima è possibile indicare, su di questo piano, il fatto che il sistema sia stabile e con quali margini di stabilità.

Il margine di guadagno m_G è dato dalla distanza tra l’intersezione con l’asse verticale e il punto centrale della carta di Nichols; il margine di fase m_φ è dato dalla distanza tra l’origine e l’intersezione con l’asse orizzontale, l’asse a 0 dB. Queste ultime osservazioni non devono stupire: da un lato, il margine di fase si può quantificare come la distanza del solo modulo, per fase pari a $-\pi$, dal punto critico -1; dualmente, considerando l’asse 0 dB, ossia la circonferenza di raggio unitario sul diagramma di Nyquist, la distanza dal punto critico è sostanzialmente l’arco di circonferenza unitaria precedentemente considerato.

Si noti che tutto ciò vale esclusivamente se si considera il secondo quadrante della carta di Nichols: se si avesse un margine di fase calcolato dall’origine verso “sinistra” e/o un margine di guadagno dall’origine verso l’ “alto”, si avrebbe un margine negativo, ossia si sarebbe già “oltrepassato” il punto critico, rendendo di fatto il sistema instabile ancora prima di incominciare un’analisi dettagliata delle specifiche. Si sappia inoltre che si potrebbero avere andamenti patologici, causati ad esempio da risonanze contenute nel sistema, che porterebbero a mettere eventualmente in dubbio la validità dei margini identificati. I margini sono così definiti, e anche se la situazione è “antipatica”, la loro definizione è sempre e comunque valida.

A cosa serve la carta di Nichols, oltre ad ottenere immediatamente una nuova definizione di stabilità? Beh, come si può vedere in MATLAB o in una carta tradizionale, la carta di Nichols è corredata da un insieme di curve a modulo costante, per la funzione T . Queste funzioni servono sostanzialmente per vedere, frequenza per frequenza, qual è la curva più vicina a ciascun punto della funzione di anello. In uscita dalla carta, confrontando la funzione di anello e le varie curve, sarà possibile ottenere, al variare delle pulsazioni ω appuntate per ogni punto, un andamento di $|T(j\omega)|$. Talvolta, per lo stesso motivo, si possono trovare curve tarate in gradi; esse sono le curve a fase costante per T , ossia $\angle T(j\omega)$.

Questa è la forma che di solito si trova per quanto riguarda il piano; noi intendiamo utilizzarlo, per il resto della trattazione, in maniera differente rispetto a questa, in un modo particolare. Si riprende un esempio precedente, in modo da meglio illustrare cosa si intenderà fare per il resto della trattazione.

Si consideri il progetto del sistema di controllo con $\hat{s}_0 = 10\%$, $\hat{s} \leq \hat{s}_0$. Da qui, si ricava con i grafici che:

$$\xi \geq \xi_0 \simeq 0,59$$

Da qui, si può ricavare banalmente che:

$$|T_{p0}| \leq 1,05 \quad |S_{r0}| \leq 1,35$$

Il concetto di stabilità relativa serve a limitare mediante l'uso di questi valori, a partire dai quali si possono ricavare margini di guadagno e margini di fase minimi per quanto riguarda la funzione di anello da progettare (mediante G_c , G_y e quant'altro).

Le informazioni su T_p e S_r dicono che, in frequenza, la funzione T deve avere un picco massimo inferiore a T_{p0} , mentre la funzione S deve avere un picco massimo inferiore a S_{r0} . A volte qualcuno fornisce direttamente questi valori, ma è più utile e didattico ricavarli dai grafici a partire dalle specifiche nel dominio del tempo.

Ciò che si intende fare è mettere in evidenza, nel grafico, una singola circonferenza, ossia quella limite, quella di T_{p0} . Di tutte le circonferenze della carta di Nichols solo una è interessante, ossia quella a $|T_{p0}|$ costante! Dal momento che le specifiche nel dominio del tempo richiedono indirettamente il fatto che si rispettino le specifiche su di essa, si dovrà progettare la funzione di anello in modo che essa non tagli, non oltrepassi in alcun punto, sulla carta di Nichols, la circonferenza a $|T_{p0}|$ costante, ma sia sempre fuori da essa. Si introduce inoltre una seconda circonferenza, ossia quella a $|S_{r0}|$ costante: la carta di Nichols solitamente presenta solo circonferenze per quanto riguarda la funzione T , ma, per gli usi che si intende fare, fondamentale è anche la circonferenza a $|S_{r0}|$ costante. In sostanza, l'introduzione di queste due circonferenze introduce due zone proibite sulla carta di Nichols, tali per cui la funzione di anello non deve assolutamente violarle.

A partire da queste funzioni è possibile anche determinare i margini di guadagno e di fase minimi che la funzione di anello deve soddisfare: analogamente a prima, il minimo per ciascuna circonferenza è dato dall'intersezione con l'asse orizzontale (per quanto riguarda la fase) e con quello verticale (per quanto riguarda il modulo) per ciascuna circonferenza; di

questi margini si sceglierà dunque quello più stringente, quello che soddisfa entrambe le condizioni, in modo da poter soddisfare tutte le specifiche contemporaneamente.

Capitolo 5

Metodi di sintesi basati sull'impiego della risposta in frequenza

Questo capitolo tratterà metodi non tanto di “sintesi” come il suo nome suggerisce, quanto di progetto, dal momento che si tratteranno tecniche di progetto nel dominio della frequenza non fornendo una strada “unica”, basata sull'uso di passi da seguire rigorosamente ai fini di ottenere un progetto di un certo tipo, bensì fornendo un certo numero di gradi di libertà per quanto riguarda il progetto, andando un po' “a tentativi” (anche se ciò sembra brutto da dirsi, in realtà non lo sarà così tanto, come si vedrà).

I metodi che verranno analizzati sono basati sul dominio della frequenza. Esistono metodi efficaci anche nel dominio del tempo, quale ad esempio la stima e retroazione degli stati, nella quale si assegnano gli autovalori in modo da stabilizzarli. Dal momento che si intende introdurre metodi atti a permettere la progettazione non solo in modo da ottenere la stabilità, ma anche in modo da ottenere un certo grado di stabilità, nonché soddisfare altre specifiche, si utilizzeranno metodi basati sul dominio reciproco. Si utilizzeranno approcci comunque “classici”, lavorando su $s = j\omega$, traducendo tutte le specifiche fornite da tempo a frequenza, per poi progettare il controllore.

La struttura del controllore che si intende progettare sarà la seguente:

$$G_c(s) = \frac{K_c}{s^r} R_d(s) R_i(s)$$

Oltre al già citato primo elemento, si introducono due funzioni aggiuntive, $R_d(s)$ e $R_i(s)$, che avranno un guadagno stazionario unitario al fine di permettere di caratterizzare interamente il comportamento stazionario del sistema nel solo K_c .

Si è già visto come lavorare a regime permanente al fine di tradurre le specifiche su di esso in informazioni quali numero di poli nell'origine per quanto riguarda il controllore, valore minimo del modulo di K_c , e altro. Questa parte del progetto, che riguarda esclusivamente il regime stazionario, è detta “progetto statico” del controllore, poichè riguarda esclusivamente il caso di $t \rightarrow \infty$. Si è visto in seguito che le specifiche di transitorio conducono a informazioni su di un valore di ω_c (pulsazione di attraversamento della funzione di anello $G_a(s)$ per l'asse 0 dB), e alla definizione di margini di stabilità (margini di modulo e margini di fase), introducendo dunque poi il “progetto dinamico”. I disturbi sinusoidali sono stati tradotti in vincoli sugli andamenti delle funzioni $S(s)$ e $T(s)$ (le funzioni di sensibilità), vincolandone ω_c , dunque indirettamente la banda passante.

Volendo soddisfare le specifiche a regime permanente, è prima di tutto necessario fissare i parametri che riguardano il regime permanente, e non toccarli più: fissato il comportamento a bassa frequenza quello non deve in alcun modo essere modificato; per questo motivo le reti di correzione, R_d e R_i , devono avere un guadagno unitario per $s \rightarrow 0$: non si deve in alcun modo modificare le caratteristiche stazionarie del sistema dopo il progetto statico.

In sostanza, il progetto del controllore si può riassumere in tre passi:

1. Traduzione delle specifiche (nella loro interezza: tutte le specifiche devono essere tradotte);
2. Progetto statico (coinvolgente sostanzialmente il dimensionamento del guadagno stazionario K_c del controllore e l'introduzione di eventuali poli nell'origine) e loop shaping (dare la “forma” alla funzione di anello, mediante la modifica di R_i e R_d , in grado di introdurre una parte dinamica, ossia variabile con la variabile di Laplace, s);
3. Verifica delle prestazioni ottenute, mediante simulazione del sistema di controllo.

Per le prime due fasi si utilizzeranno modelli semplici, basati spesso sull'uso di grafici che permetteranno comunque la definizione di un buon progetto; per quanto riguarda la simulazione, deve essere possibile effettuarla con precisione arbitraria, dunque saranno necessari modelli estremamente fini del sistema modellizzato.

5.1 Funzioni compensatrici elementari

5.1.1 Controllore ad azione proporzionale

Si consideri a questo punto un esempio semplice: dato come riferimento il solito schema a blocchi, che verrà sostanzialmente utilizzato per tutta la trattazione, in cui si ha:

$$G_p(s) = \frac{10}{s(s+1)(s+10)}$$

Sistema piuttosto comune, quale ad esempio un normale DC-motor comandato in armatura, con un polo nell'origine (provocato dall'integrazione della velocità), un polo meccanico a bassa frequenza e uno elettrico a frequenza più elevata, si considerino:

$$G_r = A = G_y = G_b = G_d = 1$$

$$d_1 = d_t = d_2 = 0$$

Si considerino, come specifiche:

$$|e_\infty^r| = 1; \quad \hat{s} \leq 10\%$$

Il controllore ad azione proporzionale ha espressione del tipo:

$$G_c(s) = \frac{K_c}{s^r}$$

Può capitare di dover introdurre alcuni poli nell'origine, anche se a questo punto il controllore sarebbe ad azione integrale, più che ad azione proporzionale. Al variare di K_c varia la stabilità, ma anche altre prestazioni del sistema quali sovraelongazione.

Si ha a che fare con due specifiche ben distinte tra loro, che dunque andranno trattate separatamente.

Si tratti per ora la specifica sul modulo dell'errore in regime permanente:

$$|e_\infty^r| \leq 1$$

Dal momento che si suppone che il sistema venga eccitato mediante un riferimento a rampa, e che l'errore dovrà essere finito e non nullo, il sistema sarà di tipo 1. La funzione di trasferimento sul ramo diretto dovrà dunque avere un polo nell'origine:

$$G(s) = A \cdot G_c(s) \cdot G_p(s)$$

Dovrà avere un polo nell'origine. Si osservi tuttavia che $G_p(s)$ ha intrinsecamente un polo nell'origine, dunque $G_c(s)$ non avrà bisogno di alcun polo: $r = 0$. Si ha che:

$$|e_\infty^r| = \left| \frac{K_d^2}{K_v} R_0 \right|$$

Ma:

$$K_d = \frac{1}{H} = \frac{1}{G_t G_y} = 1$$

Inoltre, $R_0 = 1$, dal momento che la rampa è unitaria; per tal motivo, si ha che:

$$|e_\infty^r| = \left| \frac{1}{K_v} \right|$$

Ma, per quanto riguarda il guadagno stazionario di velocità, K_v , si può dire che:

$$\begin{aligned} K_v &= \lim_{s \rightarrow 0} sG(s) = \lim_{s \rightarrow 0} sG_c(s)AG_p(s) = A \lim_{s \rightarrow 0} sG_c G_p K_c = \\ &= K_c \cdot \frac{s \cdot 10}{s \cdot 1 \cdot 10} = K_c \end{aligned}$$

Dunque:

$$K_c = K_v$$

Si può dunque dire che:

$$|e_\infty^r| = \left| \frac{1}{K_c} \right| \leq 1 \implies K_c \geq 1$$

In questo modo, è stato progettato il controllore sotto il punto di vista del progetto statico: K_c deve essere almeno pari a 1, affinché sia soddisfatta la specifica sull'errore in regime permanente.

Si analizzi a questo punto la seconda specifica, ossia quella che interessa la sovralongazione; dai grafici, si può evincere semplicemente che:

$$\xi \geq 0,59$$

Dato dunque lo smorzamento ξ , si può ottenere:

$$T_{p0} = 1,05; \quad S_{r0} = 1,35$$

Si possono quindi ricavare, o mediante la carta di Nichols o mediante il calcolo analitico, i seguenti margini di guadagno e fase, da rispettare affinché la specifica sulla sovraelongazione sia a sua volta rispettata:

$$m_\varphi = 58^\circ$$

$$m_G = 11,5 \text{ dB}$$

Il fatto che le specifiche siano rispettate alla pulsazione ω_c è verificabile sul diagramma di Bode di modulo e fase: vedendo che $\omega_c = 0,8 \text{ rad/s}$, la fase è pari a -132° , dunque il sistema è stabile con 48° di margine di fase (sostanzialmente l'unico margine interessante). Ciò non va bene: il margine di fase soddisfa la condizione di stabilità, ma non quella dello smorzamento dei poli necessario per la soddisfazione della condizione concernente la sovraelongazione.

Si noti che anche il diagramma di Nichols sembra fornire informazioni riguardo questo fatto, dal momento che la funzione di anello, $G_a(s)$, attraversa le curve a modulo costante rappresentate. Ciò non vuol dire in realtà niente: il sistema in questione non è infatti un sistema del secondo ordine, quindi la condizione sulle curve a modulo costante non è necessaria e sufficiente, ma solo sufficiente: non avendo attraversamenti delle curve si ha la garanzia che le specifiche siano rispettate, ma avendone non si ha la garanzia del fatto che le specifiche non siano rispettate. In questo caso comunque i diagrammi di Bode forniscono informazioni negative riguardo i margini di fase, come si può verificare anche da una simulazione nel dominio del tempo, più significativa dei criteri finora usati.

Questo è il risultato in seguito alla scelta di un K_c pari a 1: la condizione sul regime permanente è soddisfatta, quella su ξ no.

Si potrebbe agire su K_c e modificarlo, ma non è possibile per alcuni motivi:

- Una volta terminato il progetto statico, sarebbe buona cosa mantenerlo intatto e non modificarlo;
- Volendo aumentare K_c , la specifica sull'errore sarebbe ancora soddisfatta, ma quella riguardo la sovraelongazione sarebbe ancor meno rispettata: aumentare K_c significa traslare "in alto" la curva G_a sul piano di Nichols, poichè si aumenta il guadagno, dunque questa via non si può prendere;
- Volendo diminuire K_c , per ora unitario, si andrebbe contro la specifica riguardante l'errore a regime permanente; si potrebbero migliorare le

prestazioni per quanto riguarda l'altra specifica, ma di sicuro peggiorare la prima, dunque perderci sotto un punto di vista per guadagnarci sotto un altro, scelta infelice.

5.1.2 Controllore ad azione anticipatrice

Si consideri a questo punto una delle due funzioni dinamiche per ora non considerate: $R_d(s)$, o “rete anticipatrice” (il pedice d sta per “derivativa”, poichè, come si vedrà, l'andamento in frequenza della funzione è sostanzialmente quello di un derivatore). La funzione ha una forma di questo tipo:

$$R_d(s) = \frac{1 + \frac{s}{\omega_{z,d}}}{1 + \frac{s}{\omega_{p,d}}}, \quad \omega_{p,d} = m_d \cdot \omega_{z,d}, \quad m_d > 1$$

Questa è la struttura generale delle reti di questo tipo; il comportamento in frequenza, come si può immaginare, sarà il seguente:

Si parte da 0 dB dal momento che il guadagno stazionario è unitario, al fine di non modificare le informazioni di K_c ; al crescere di m_d , si tende ad aumentare la pulsazione del polo, distanziandola da quella dello zero, posto prima di essa. I grafici contengono pulsazioni normalizzate, in modo da poterle adattare a un progetto qualunque. Il fatto che lo zero sia posto prima del polo provoca la presenza di una zona in cui sia il guadagno sia la fase tendono ad aumentare, ottenendo dunque un recupero di fase; scegliendo la pulsazione normalizzata più idonea, si riesce a modificare al variare della frequenza (utilizzando una certa frequenza di riferimento, come si vedrà in seguito in esempi pratici) il comportamento della funzione di anello, effettuando la cosiddetta operazione di “loop shaping”, introducendo dunque un andamento selettivo con la frequenza della funzione di anello.

5.1.3 Controllore ad azione attenuatrice

Una volta considerata $R_d(s)$, si consideri $R_i(s)$, ossia la rete ad azione attenuatrice (dunque duale alla precedente, la cui risposta in frequenza tendeva ad ottenere un aumento del guadagno e un recupero di fase). Come prima, si avrà un guadagno stazionario unitario, un polo a frequenza bassa, e uno zero a frequenza maggiore rispetto a quella del polo:

$$R_i(s) = \frac{1 + \frac{s}{\omega_{z,i}}}{1 + \frac{s}{\omega_{p,i}}}, \quad \omega_{z,i} = m_i \cdot \omega_{p,i}, \quad m_i > 1$$

In questo caso l'andamento sarà quello di un integratore (il pedice i sta per l'appunto per “rete integrativa”):

I grafici sono in realtà del tutto analoghi a quelli studiati per quanto riguarda $R_d(s)$; è sufficiente considerare il fatto che la fase e i guadagni subiscono variazioni in negativo anziché in positivo, dunque, utilizzando gli stessi disegni con però un segno “-” dinnanzi a ogni valore, si hanno sostanzialmente anche i grafici per quanto concerne questo tipo di rete.

5.2 Esempio di progetto di rete integrativa

Si considera a questo punto l'esempio precedentemente introdotto di realizzazione del controllore, utilizzando però una rete attenuatrice in grado di soddisfare entrambe le specifiche. Come suggerisce la specifica sul margine di stabilità, al fine di rendere il sistema stabile e soddisfare le specifiche, è necessario tenere d'occhio, per la pulsazione ω_c (parametro in realtà ancora da progettare), il margine di fase, pari a 58° .

Come deve essere la ω_c finale rispetto a quella ottenuta con il progetto iniziale, con $K_c = 1$? Beh, essa deve essere tale, come la definizione di ω_c suggerisce, da avere modulo unitario della funzione di anello. La fase deve essere inoltre tale da rispettare il minimo margine di fase, in modo da avere la certezza di soddisfare la specifica. Si considererà eventualmente un margine di 60° , in modo da alleggerire la notazione e ottenere un margine sul margine. La fase, alla frequenza ω_c , dovrà dunque essere maggiore o uguale di -120° ; negli altri punti non è particolarmente interessante: la specifica richiede sempre e comunque che il margine sia rispettato sulla pulsazione di taglio dell'asse 0 dB selezionata (e ancora da selezionare, per ora); il resto non importa.

Si vuole usare una rete attenuatrice: essa, al variare della frequenza, fa diminuire il modulo, ossia lo attenua; si introduce dunque un'attenuazione variabile al variare della pulsazione. La fase diminuirà di conseguenza: se da un lato si attenua il modulo, dall'altro si tende a perdere fase, cosa che di fatto ci è un po' antipatica: una rete di questo tipo fa perdere margine di fase, non ne fa recuperare. Ciò può portare a intuire in quale range di frequenze debba essere ω_c : la rete deve attenuare, dunque si può solo scegliere valori di pulsazione tale per cui modulo e fase sian sufficientemente superiori rispetto a quelli dell'attuale ω_c (vi deve essere una zona di guadagno e una fase positiva, dal momento che la rete attenuatrice “abbassa” la funzione di anello). Una possibile candidata per la ω_c potrebbe essere la seguente:

$$\omega_c = 0,1 \text{ rad/s}$$

Infatti, la fase è circa -96° , quindi si può sperare che la rete attenuatrice sia tale da non abbassare troppo la fase. Il modulo è circa 20 dB, dunque, in

0,1 rad/s, sarà necessario avere, in seguito all'introduzione della rete, 0 dB: essa dovrà abbassare, per $\omega = \omega_c$, il modulo della funzione di anello di 20 dB.

Sul grafico si procede a questo punto in questo modo: si sceglie un'ascissa normalizzata a valori elevati, la più elevata possibile, in modo da perdere la minor fase possibile; si tende a scegliere il valore di fondo scala sull'asse orizzontale, ottenendo dunque qualcosa del genere:

$$\left. \frac{\omega}{\omega_{p,i}} \right|_{\omega=\omega_c, des} = 100$$

Si fa in modo che il modulo scenda di 20 dB, dunque si seleziona un valore di m_i pari a 10. Si avrà:

$$\frac{\omega_c}{\omega_{p,i}} = 100 \longrightarrow \omega_{p,i} = 10^{-3} \text{ rad/s}$$

Da ciò si può ricavare la pulsazione dello zero della rete:

$$\omega_{z,i} = 10 \cdot 10^{-3} = 10^{-2} \text{ rad/s}$$

Una volta introdotti questi parametri nel controllore, per una pulsazione pari a 0,1 rad/s la funzione attraversa l'asse 0 dB, guadagnando alcuni gradi rispetto al valore della fase con la precedente ω_c , ottenendo dunque il risultato. Si noti che l'andamento della fase, in seguito, sarà decrescente e raggiungerà livelli molto bassi; ciò non ci interessa in alcun modo, dal momento che l'unico punto importante è il passaggio per l'asse 0 dB, ossia la cosiddetta "circonferenza di raggio unitario"; in tutti gli altri punti tendenzialmente la fase può assumere i valori che preferisce, ma non interesseranno particolarmente la trattazione.

Si può notare, disegnando il diagramma di Nichols, che la forma della funzione di anello è decisamente cambiata; si potrebbe però vedere che la condizione sull'attraversamento delle circonferenze continua a non essere rispettato. Prima di operare, conviene tuttavia osservare, simulare la risposta nel tempo: di fatto come già detto la condizione riguardo le curve a modulo costante è solo sufficiente, non necessaria, affinché la specifica sia rispettata; si può osservare, dalla simulazione nel dominio del tempo, che ambo le specifiche sono soddisfatte, nonostante la condizione non lo sia. Ci si potrebbe a questo punto fermare qui: il controllore di fatto funziona, anche se Nichols non è soddisfatto. Il controllore ora progettato ha una forma di questo tipo:

$$G_c(s) = K_c \cdot R_i(s) = 1 \cdot \frac{1 + \frac{s}{0,01}}{1 + \text{frac}s0,001}$$

Abbiamo dunque soddisfatto al contempo la condizione sull'errore a regime stazionario, dal momento che per frequenze basse la rete integrativa non agisce in alcun modo, sia la condizione sulla sovraelongazione. Si sappia che la rete integrativa comunque non è una soluzione ottimale: come si vedrà, essa introduce inesorabilmente una latenza nel sistema, rallentandolo, come si dirà eventualmente in seguito.

Si può migliorare a questo punto la condizione sul tempo di salita? Beh, per farlo, sarebbe necessario aumentare la pulsazione ω_c , e per fare ciò "alzare" la funzione di anello; in questo modo, i punti associati alle varie pulsazioni si "alzeranno". Aumentando il valore di K_c si può migliorare il progetto, trasladando verso l'alto il diagramma di Nichols, arrivando dunque addirittura a soddisfare la condizione sufficiente e migliorando in questo caso il progetto. Si sappia che di solito è meglio non esagerare sotto questo punto di vista, quindi, ottenuto un progetto che soddisfi tutte le specifiche, si può già essere piuttosto contenti.

5.2.1 Esempio di progetto di rete anticipatrice

Si ripeterà a questo punto la procedura di progetto, utilizzando questa volta una rete di tipo anticipatrice; si faranno meno osservazioni di prima, dal momento che alcuni concetti sono già stati di fatto fissati nella precedente sottosezione.

Il problema considerato è sempre il solito, ma a questo punto si utilizza un terzo approccio, basato su di una rete anticipatrice, dunque derivativa. Beh, a questo punto si faccia la seguente osservazione, duale alla precedente: per come è realizzata, una rete anticipatrice tende a guadagnare e a recuperare fase, dunque bisognerà comportarsi in modo duale a prima. Un buon modo di lavorare (assolutamente non l'unico, ma per ora quello suggerito) è quello di considerare una pulsazione di taglio dell'asse 0 dB di poco maggiore rispetto alla precedente, 0,8 rad/s, per compensarlo con la rete anticipatrice; si sceglie una zona con circa 0,5 dB di recupero del modulo, o comunque valori molto bassi, al massimo pari a 1 dB, in modo da poter guadagnare 12° abbondanti sul margine nella nuova pulsazione di passaggio per l'asse 0 dB.

Si sceglie dunque una $\omega_{c,des} \simeq 0,8$ rad/s; si ottiene:

$$\frac{\omega}{\omega_{z,d}} \Big|_{\omega=\omega_{c,des}} = 0,3$$

Si sceglie $m_d = 5$

Da qui:

$$\omega_{z,d} = 2,6 \text{ rad/s}$$

$$\omega_{p,d} = 5 \cdot 2,6 = 13 \text{ rad/s}$$

La rete $R_d(s)$ avrà dunque una forma del tipo:

$$R_d(s) = \frac{1 + \frac{s}{2,6}}{1 + \frac{s}{13}}$$

ω_c sostanzialmente non subisce variazioni sensibili, e la fase recupera quel tanto che basta per soddisfare la specifica; il progetto si può dire concluso.

Alcune osservazioni che completano il quadro di quelle precedentemente fatte: conviene utilizzare questa strada, ossia quella di procedere seguendo sempre le stesse linee guida, ma senza avere “ricette”: il progettista deve potersi muovere liberamente, al fine di ottenere risultati validi in qualsiasi situazione.

Si noti un altro fatto: finora il loop shaping, ossia la modifica della funzione di anello, è stata fatta solo ed esclusivamente per $\omega = \omega_{c,des}$; in realtà, questo tipo di procedimento si può fare per qualsiasi pulsazione: la modellazione della funzione di anello utilizzando lo stesso metodo si può scegliere considerando qualsiasi riferimento di pulsazione; si è scelta quella di passaggio per l’asse 0 dB poichè è la più comoda e significativa, per quello che finora è stato visto.

5.2.2 Simulazione degli esempi 1, 2, 3

In questa sottosezione si raccoglieranno osservazioni effettuabili mediante il software Simulink, unito ovviamente alla sua base, MATLAB; queste osservazioni troverebbero riscontro in un’eventuale simulazione, che si consiglia di fare, al fine di verificare il successo nel progetto o effettuare eventuali migliorie. Verranno considerati i tre esempi finora utilizzati, basati sui tre tipi di controllori finora introdotti. In un progetto verranno solitamente assegnate le specifiche; le prestazioni ottenute dovranno essere documentate, in modo da verificare e permettere al committente di verificare l’effettivo funzionamento del progetto. Nella fattispecie, fondamentale è documentare le prestazioni che riguardano le specifiche: solo esse sono importanti per il committente, che ha fornito indicazioni riguardo alcuni parametri quali sovraelongazione o errore a regime permanente (come nel caso della simulazione che verrà descritta). Mostrare grafici qualitativi, eventualmente quotati, può essere molto utile nella documentazione.

Esempio 1

Si consideri, per il progetto del primo tipo (con controllore ad azione proporzionale), il seguente parametro:

$$K_c = 0,55$$

Ciò che conviene fare è disegnare i diagrammi di Bode, di Nyquist e di Nichols, in modo da visualizzare almeno la risposta qualitativa. Si può vedere che:

$$\hat{s} = 7,5\%$$

$$t_s = 4,06 \text{ s}$$

$$t_a = 7 \text{ s}$$

La lettura deve essere precisa con un errore pari al 5 %, assolutamente non superiore al 10 %. Al fine di rilevare questi valori, sono state utilizzate le definizioni precedentemente esposte: la formula della sovraelongazione, il primo attraversamento per il valore a regime permanente, e l'ultima volta che si attraversa il "confine" entro cui si rientra nella fascia. Si noti che, per rilevare l'errore in regime permanente rispetto al riferimento, si deve collegare solo e unicamente esso: tutti gli altri collegamenti vanno rimossi, al fine di non sovrapporre gli effetti dei vari ingressi (considerando dunque anche i disturbi e non solo la fedeltà).

Si disegna quindi anche la risposta nel tempo dell'errore al riferimento, $e^r(t)$ (definita come differenza tra il riferimento e l'uscita del sistema reale, evidenziando l'errore a regime e l'errore massimo. Importante è la documentazione del "plant input", ossia del segnale che va in ingresso al controllo: si deve evidenziare il valore del massimo di questo parametro, che risulterà essere fondamentale per varie motivazioni: di fatto, il plant input è il segnale che agisce sul sistema di controllo in modo da regolare l'uscita al valore desiderato; se questo segnale è troppo intenso, cosa che potrebbe capitare, il sistema potrebbe danneggiarsi anche in maniera irrecuperabile. In questo caso, si può vedere che il massimo valore del plant input sarà:

$$u_{max} = 22$$

Esempio 2

In questo caso, il controllo avrà una forma del tipo:

$$G_{c,2}(s) = 4 \frac{1 + \frac{s}{10^{-3}}}{1 + \frac{s}{10^{-2}}}$$

In questo caso, si ha un effetto di tipo diverso, precedentemente non presente (e non presente con nessun altro tipo di rete correttiva): l'effetto coda. Simulando si può vedere che, nonostante il sistema sia di tipo 1, l'errore vada a zero asintoticamente, non immediatamente. Ciò dipende dal fatto che le costanti di tempo dei modi sono estremamente alte, causate dalla presenza di zeri della funzione di trasferimento del sistema. Si può infatti vedere che:

$$G_a = \frac{N_a}{D_a} \implies G_{ry} = \frac{G_a}{1 + G_a} = \frac{N_a}{N_a + D_a}$$

Gli zeri del controllore sono anche zeri della funzione di anello, e quindi radici di G_{ry} , ossia della funzione di trasferimento del sistema retroazionato. Gli zeri del controllore si ritrovano dunque tra riferimento e uscita. Conoscendo tuttavia l'andamento della funzione $T(j\omega)$, strettamente imparentata con $G_{ry}(j\omega)$ (a meno del guadagno stazionario), si vede che, dove dovrebbe esserci uno zero, la funzione è piatta, non guadagna. Questo perché la retroazione fa nascere un polo che cancella lo zero, ma dunque introduce nel sistema un modo a bassa frequenza, che dunque tenderà ad estinguersi per tempi molto elevati, rendendo di fatto fastidioso il tutto.

Si può vedere che questo tipo di sistema ha anche elementi positivi: il massimo comando, in questo ambito (plant input), è:

$$u_{max} = 16$$

Dunque inferiore rispetto al precedente.

Esempio 3

Si considera a questo punto il seguente controllore:

$$G_{c,3}(s) = 1 \cdot \frac{1 + \frac{s}{2,6}}{1 + \frac{s}{13}}$$

I valori meno importanti rilevati verranno poi presentati nella tabella riassuntiva. Si vuole fare un'osservazione: la rete anticipatrice non ha più lo svantaggio dell'effetto coda, che di fatto rende decisamente poco appetibile

una rete di tipo integratrice (a meno di casi particolari per ora non analizzati); la rete integrativa, dunque, non andrà quasi mai utilizzata, a meno di casistiche piuttosto importanti.

Il sistema è molto veloce, ma presenta un grosso problema:

$$u_{max} = 200$$

Ciò è decisamente grave: la rete provoca grosse “botte” al sistema di controllo, quantomeno rispetto ai casi precedentemente proposti, dunque servono sistemi di controllo molto robusti sotto il punto di vista della dinamica dei segnali in ingresso affinché possano supportare una stimolazione di questo genere, o comunque tecniche atte a limitare questo tipo di problemi.

Si tenga comunque conto di ciò: questa rete è la più importante poichè l’effetto coda è estremamente fastidioso; se si può fare a meno di usare una rete attenuatrice, non la si usi.

5.3 Studio del segno di K_c

Precedentemente, parlando di traduzione delle specifiche, sono state introdotte condizioni coinvolgenti esclusivamente il modulo di K_c , ossia nella fattispecie la sua ampiezza; niente è stato finora detto per quanto riguarda il segno.

Scegliere il segno di K_c è fondamentale: non sempre infatti è possibile studiare sistemi con un segno qualsiasi di K_c .

Si supponga che si abbia un diagramma di Nyquist di questo tipo:

In questo caso, si vede che $n_{p,a} = 0$, ma $N = 2$, dunque il sistema è instabile. Mediante il loop shaping è possibile tuttavia modificare la funzione di anello, in modo da stabilizzare il sistema, ottenendo qualcosa di questo tipo:

Ciò che è stato fatto con la rete compensatrice è modificare la funzione in modo da far “girare attorno” al punto critico, allontanandosi da esso. Questo, in questo caso, con $K_c > 0$. Se $K_c < 0$, tuttavia, il diagramma di Nyquist sarà il simmetrico rispetto all’asse verticale di questo, e si avrà tendenzialmente $N = 1$. In questo caso, non esiste modo di togliere il punto dalla circonferenza mediante loop shaping, dal momento che esso è chiuso dalla circonferenza di raggio infinito chiudente le curve.

Si parla in questo ambito di “stabilizzabilità” mediante reti dinamiche: K_c è vincolato da alcune specifiche dunque non si deve toccare, e si deve operare esclusivamente mediante reti dinamiche (le due categorie di reti finora viste più un’eventuale terza tra breve presentata); è dunque necessario ricavare il segno di K_c mediante diagramma di Nyquist, in modo da determinare la stabilizzabilità del sistema e progettare dunque le reti nel modo giusto.

5.4 Attività del “comando” in funzione del riferimento

5.4.1 Analisi

Si vuole a questo punto introdurre qualche strumento utile al fine di trattare analiticamente il problema dell'attività del comando (plant input), già precedentemente toccato. Ciò che si dirà sarà valido, a patto che siano rispettate le seguenti ipotesi:

1. Segnale di riferimento a gradino (caso comunque buono, dal momento che, sotto il punto di vista del transitorio, è il segnale più esigente);
2. Le reti anticipatrici e attenuatrici progettate a bassa frequenza rispetto a ω_c ;
3. Non vi siano poli di chiusura nel controllore, ossia poli non accompagnati da zeri.

Siamo a questo punto interessati all'attività del comando in presenza di un riferimento; data l'espressione di $U(s)$ dove per U si intende la trasformata di Laplace dell'attività del comando, si ha che:

$$U(s) = \frac{A \cdot G_c(s)}{1 + G_a(s)} \cdot R(s)$$

In sostanza questa è la funzione tra riferimento e comando:

$$= G_c(s)S(s)R(s)$$

Si può dimostrare che, sotto le ipotesi, il valore massimo del comando è nell'origine, e si può calcolare mediante il teorema del valore iniziale:

$$\begin{aligned} u_{max} &= \lim_{t \rightarrow 0} u(t) = \lim_{s \rightarrow \infty} sU(s) = \lim_{s \rightarrow \infty} sG_c(s)AS(s)R(s) = \\ &= \lim_{s \rightarrow \infty} sG_c(s)AS(s) \frac{R_0}{s} = \\ &= A \lim_{s \rightarrow \infty} K_c \frac{1 + \frac{s}{\omega_{z,d}}}{1 + \frac{s}{m_d \omega_{z,d}}} \cdot \frac{1 + \frac{s}{m_i \omega_{p,i}}}{1 + \frac{s}{\omega_{p,i}}} \cdot S(s) = \\ &= AK_c \frac{m_d}{m_i} R_0 = u_{max} \end{aligned}$$

Dove, date più reti, m_d e m_i sono le produttorie di tutti i singoli $m_{d,j}$ e $m_{i,k}$ di ogni singola j -esima e k -esima rete.

Volendo applicare ciò agli esempi, si può vedere (considerando $R_0 = 40$ per esempio) che:

1. Per il primo esempio:

$$u_{max} = 1 \cdot 0,55 \cdot \frac{1}{1} \cdot 40 = 22$$

2. Per il secondo esempio:

$$u_{max} = 1 \cdot 4 \cdot \frac{1}{10} \cdot 40 = 16$$

3. Per il terzo esempio:

$$u_{max} = 1 \cdot 1 \cdot \frac{5}{1} \cdot 40 = 200$$

Tutti i valori ricavati mediante la simulazione sono assolutamente verificati da questo procedimento algebrico.

5.4.2 Progetto

Ci si pone a questo punto il problema del progetto: si intende introdurre la possibilità di trattare specifiche riguardanti l'attività del comando. In altre parole, quello che si potrebbe richiedere, è il fatto che l'attività di comando sia inferiore ad un certo valore: dato un certo \bar{u} , si chiede che tutti i valori del comando $u(t)$ siano inferiori ad esso:

$$|u(t)| \leq \bar{u}$$

Ciò che si può a questo punto fare è, date le ipotesi precedentemente formulate (in fase di analisi), cercare alcuni risultati. Si sa, da prima, che:

$$u_{max} = \left| A \cdot K_c \cdot \frac{m_d}{m_i} \cdot R_0 \right| \leq \bar{u}$$

Dunque:

$$\left| \frac{m_d}{m_i} \right| \leq \frac{\bar{u}}{AK_c R_0}$$

Si ha dunque un vincolo sulla scelta di m_d e m_i !

Si supponga a questo punto ciò: una volta fatto il progetto, dati A , K_c , m_d , m_i , si supponga di voler verificare il fatto che il sistema sia in zona di saturazione, ossia che il controllo abbia valori di ingresso troppo elevati. Si può invertire la formula, e ricavare il massimo valore di ampiezza della rampa accettabile:

$$R_0 \leq \frac{\bar{u}m_i}{AK_c m_d}$$

Giocando sull'inversione di questa formula, è possibile dunque ricavare vincoli su alcuni parametri, imponendo però la presenza degli altri.

5.5 Attività del comando in funzione di un ingresso sinusoidale

5.5.1 Analisi

Una volta trattato il caso di segnale a gradino, si tratta ora il caso di disturbo sul trasduttore, disturbo tipicamente sinusoidale. Si supponga dunque di aver a che fare con segnali del tipo:

$$d_t(t) = a_t \sin(\omega_t t)$$

Si consideri a questo punto l'espressione di $U(s)$ (con significato sostanzialmente analogo a prima):

$$U(s) = G_{d_t,u} \cdot D_t(s)$$

Utilizzando i risultati noti dalla teoria della risposta in frequenza di sistemi LTI, si può dire, maggiorando il seno a 1, che:

$$u_{max} = a_t \cdot |G_{d_t,u}(j\omega_t)|$$

Dove:

$$G_{d_t,u}(s) = \frac{G_y \cdot (-1) \cdot G_c(s) \cdot A}{1 + G_a(s)}$$

Calcolando mediante una calcolatrice o MATLAB questi parametri (la funzione di risposta in frequenza nella ω_t del disturbo), è possibile quantificare dunque il picco dell'attività del comando.

5.5.2 Progetto

In ambito di progetto, si lavora come tra poco esposto; si consideri:

$$G_{d_t,u}(s) = -G_c(s)G_yAS(s)$$

Per $s = j\omega_t$, l'espressione diventa:

$$G_{d_t,u}(s) = -G_c(j\omega_t)G_yAS(j\omega_t)$$

Si introduce una prima approssimazione: supponendo che il disturbo sia ad alta frequenza, $\omega_t \gg \omega_c$, dunque $S(j\omega_t) \simeq 1$. Si introduce dunque una seconda approssimazione, in realtà molto più grossolana della prima: per $\omega = \omega_t$, si dice che:

$$G_c(s) \sim K_c \frac{m_d}{m_i}$$

Questa è un'approssimazione molto più grossolana; si noti inoltre un fatto: se sono presenti poli nell'origine, l'espressione si riporta alla seguente correzione:

$$G_c(s) \sim K_c \frac{m_d}{m_i} \frac{1}{(j\omega_t)^r}$$

Dunque, usando questa espressione, si può approssimare il massimo dell'attività di comando come:

$$u_{max} \sim AG_y K_c \frac{m_d}{m_i} \left| \frac{1}{(j\omega_t)^r} \right| a_t$$

Questa formula si può utilizzare solo e solamente in fase di progetto: è una formula totalmente inadatta in sede di analisi, ma può tornare utile per il progetto se si conoscono pochi dati.

5.6 Riduzione della complessità delle reti correttrici

Precedentemente è stato fatto cenno ad una possibile terza rete compensatrice, ma non si è più approfondito l'argomento.

Il discorso è: quando si progetta un controllore, è necessario evitare di esagerare con i gradi di libertà: le reti correttrici di fatto costano, dunque al massimo se ne usino tre o quattro, nei progetti relativi alla trattazione.

Aspetto fondamentale è il seguente: un controllore, affinché esso sia fisicamente realizzabile, deve avere una funzione di trasferimento propria, ossia in cui il grado del numeratore sia minore o uguale di quello del denominatore: il numero di zeri deve essere al più pari a quello di poli presenti nel sistema. Si può tuttavia osservare ciò: come ben noto, il sistema di controllo ha un'espressione del tipo:

$$G_c(s) = \frac{K_c}{s^r} R_d(s) R_i(s)$$

Si supponga che nel problema il controllore debba introdurre K_c ma anche un polo nell'origine; come detto, il numero di zeri presenti nel sistema deve essere al più pari a quello di poli, ma se il controllore ha un polo nell'origine, perchè non sfruttarlo per introdurre e posizionare esclusivamente uno zero, anzichè una coppia zero-polo come suggerirebbero di fare le reti finora proposte? L'idea, nella fattispecie, potrebbe essere quella di usare una funzione di controllo di questo tipo:

$$G_c(s) = \frac{K_c}{s} \left(1 + \frac{s}{\omega_z} \right)$$

Questa rete è detta "Rete P.I." (o Proporzionale Integrale); questo nome deriva dal fatto che questa rete si può sviluppare, evidenziando due contributi:

$$G_c(s) = \frac{K_c}{s} + \frac{K_c}{\omega_t}$$

Ossia un contributo integrativo, come si può osservare, e uno puramente proporzionale. Progettare reti di questo tipo è meno costoso, dunque, se è presente un polo nell'origine, è preferibile progettare reti di questo tipo anzichè di altri. Basta posizionare, mediante i grafici che forniscono le informazioni sugli zeri, la posizione dello zero, e quindi scegliere quanta fase è necessario recuperare, ricavando poi l'ascissa da quella normalizzata come:

$$\frac{\omega}{\omega_z} \Big|_{\omega=\omega_c} = \omega_{norm}$$

5.7 Realizzazione di controllori analogici mediante reti RC-attive

Si consideri una rete attiva di questo tipo:

Distinguendo le impedenze Z_1 e Z_2 , è noto che la funzione di trasferimento ha un'espressione di questo tipo:

$$\frac{V_u}{V_e} = -\frac{Z_2}{Z_1}$$

Effettuati i passaggi, si può ottenere:

$$\frac{V_u}{V_e} = -\frac{R_4}{R_1 + R_2} \cdot \frac{1 + sR_2C_2}{1 + sC_2(R_2 + R_4)} \cdot \frac{1 + sR_1C_1}{1 + sC_1\frac{R_1R_3}{R_1+R_3}}$$

Si possono individuare sostanzialmente tre contributi: uno puramente statico, due dinamici. Si osservi il primo contributo: si tratta di funzioni fattorizzate nella forma “costanti di tempo”; ciò che si può osservare, è che il termine al numeratore avrà una costante di tempo sempre inferiore rispetto a quella al denominatore, dal momento che al denominatore si ha lo stesso termine, più un contributo dovuto a R_4 ; si può dire che questa sia una rete integrativa, dal momento che la costante di tempo del denominatore è sempre maggiore di quella del numeratore, dunque la frequenza del polo al denominatore sempre minore di quella dello zero.

Ragionamento del tutto duale si può fare per la seconda funzione: la costante di tempo al numeratore sarà sempre maggiore di quella al denominatore, dal momento che al denominatore si ha la somma armonica di R_1 , presente anche al numeratore, e di R_3 , dunque il risultato sarà sempre inferiore a R_1 e R_3 ; questa rete sarà di tipo derivativo.

Mediante questo tipo di rete RC è possibile progettare il controllo, utilizzando le tecniche teoriche finora introdotte. Ciò presenta sostanzialmente un vantaggio e uno svantaggio: se da un lato una struttura del genere permette di progettare ciò che si vuole come si vuole, questa è anche difficile da utilizzare: variando il valore di un parametro di fatto si variano molti parametri, dunque progettare celle di questo tipo è difficile. Ciò che conviene fare è usare un operazionale per ciascuna rete, in modo da disaccoppiare mediante impedenze ciascuna rete, e ottenere dunque reti indipendenti tra loro.

Capitolo 6

Introduzione al controllo digitale

6.1 Struttura di sistema di controllo digitale

Un controllo digitale è un sistema di controllo che al suo interno ha segnali sia a tempo continuo sia a tempo discreto. Uno schema a blocchi per il sistema di controllo, che verrà sempre considerato per il resto della trattazione è il seguente:

Come si può vedere dallo schema, a seconda del blocco considerato si possono avere segnali di tipo numerico o analogico; si suol dire che un sistema di questo tipo, ossia che tratta entrambi i tipi di domini, sia “misto”.

Rispetto allo schema precedente si possono trovare diverse analogie: come prima è possibile trovare l'impianto, ossia il sistema da controllare (plant), di solito a tempo continuo; si può identificare, come si farà tra breve, un controllore; si notano tuttavia molti blocchi “nuovi”, che verranno sempre tra breve presentati più nel dettaglio. In uno schema a blocchi di questo tipo si ha a che fare con un confronto tra segnali analogici; una variante sul tema potrebbe ad esempio uno schema a blocchi in cui si consideri un confronto tra segnali numerici.

Il controllore analogico, G_c , trova la sua realizzazione in tre blocchi: il convertitore A/D, il “computer”, il convertitore D/A. Il cuore del controllo è di fatto il computer: al suo interno sono implementate le istruzioni in grado di realizzare la legge di controllo. Gli altri due blocchi di fatto sono interfacce: il computer lavora di fatto nel mondo numerico, ma il sistema da controllare è analogico; il fatto che sia necessario introdurre un certo numero di interfacce è dunque evidente. La legge di controllo, ossia la funzione di trasferimento del controllore, verrà implementata dunque nel computer, mediante un software

che assiste il progettista.

Da un punto di vista puramente controllistico, di fatto l'interfaccia non è molto interessante; si vedrà che il blocco fondamentale sotto il punto di vista controllistico è il circuito di mantenimento, a causa di effetti che introduce sulla funzione di anello, sulla retroazione. Questo tipo di progetto introduce alcuni vantaggi e alcuni svantaggi:

- Maggiore capacità di elaborazione e di precisione: sicuramente l'implementazione della legge di controllo sarà effettuata in maniera migliore, con maggior precisione;
- Maggior flessibilità, anche in fase di sviluppo: anziché modificare le impostazioni di dispositivi elettronici quali ad esempio resistenze differenti, è solo necessario modificare il codice implementante la funzione di controllo;
- Maggior leggibilità: dal momento che si tratta comunque di programmare, sicuramente è più semplice comprendere il progetto;
- La progettazione è più difficile rispetto a quella classica, analogica;
- La stabilizzabilità è problematica da ottenere: le condizioni di stabilizzabilità sono rese complicate dal circuito di mantenimento (Hold Circuit);
- Serve energia elettrica per alimentare il controllo; spesso una cosa di questo genere non è complicata dal momento che anche i dispositivi di solito vengono alimentati in corrente; se il sistema da alimentare è ad esempio di tipo meccanico, pneumatico, sistemi in cui non sono presenti forme di alimentazione in corrente, l'alimentazione del controllore digitale diviene problematica.

6.2 Modello matematico sul campionamento

Si vogliono a questo punto introdurre alcune nozioni basilari per quanto riguarda il campionamento, procedimento fondamentale al fine di discretizzare il dominio analogico (specialmente per quanto concerne i segnali):

Si consideri un segnale continuo nel tempo t , $f(t)$, e un treno di impulsi $\delta_T(t)$ periodico di un certo tempo T :

$$\delta_T(t) = \sum_{k=0}^{+\infty} \delta(t - kT)$$

In sostanza si moltiplicano il treno di impulsi e il segnale continuo nel dominio del tempo con un moltiplicatore, ottenendo qualcosa di questo genere: Si ha che:

$$f^*(t) = f(t) \cdot \delta_T(t) = \sum_{k=0}^{+\infty} f(t)\delta(t - kT)$$

Ciò si può vedere anche nel seguente modo:

$$f^*(t) = f(0)\delta(t) + f(T)\delta(t - T) + f(2T)\delta(t - 2T) + \dots$$

L'apice * indicherà il fatto che un segnale e, in seguito, anche una generica funzione, sono discretizzati (nel caso di segnali, campionati).

Si è imparato il fatto che i segnali a tempo continuo sono comodi per essere studiati mediante l'uso della trasformata \mathcal{Z} ; per ora, tuttavia, prima di introdurre una trattazione nel dominio della trasformata \mathcal{Z} , si passa per la trasformata di Laplace, mediante la quale si introdurranno alcune nozioni preliminari. Si consideri per ora dunque la trasformata di Laplace del segnale campionato, $f^*(t)$:

$$F^*(s) = f(0) + f(T)e^{-sT} + f(2T)e^{-2sT} + \dots = \sum_{k=0}^{+\infty} f(kT)e^{-skT}$$

Questa è la trasformata di Laplace del segnale considerato.

6.3 Spettro del segnale campionato

Ai fini di studiare lo spettro del segnale campionato, si introduce un metodo alternativo di studiare il treno di impulsi, in modo da ottenere una diversa espressione di $F^*(s)$. Si consideri dunque la seguente espressione:

$$\delta_T(t) = \sum_{n=-\infty}^{+\infty} c_n e^{-jn2st}, \quad \omega_s = \frac{2\pi}{T}$$

Dove:

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \delta_T(t) e^{-jn\omega_s t} dt = \frac{1}{T}$$

Quindi:

$$\delta_T(t) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} e^{jn\omega_s t}$$

Si può dunque dire che:

$$f^*(t) = f(t)\delta_T(t) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} f(t)e^{jn\omega_s t}$$

Da qui:

$$\begin{aligned} F^*(s) &= \int_0^{\infty} \frac{1}{T} \sum_{n=-\infty}^{+\infty} f(t)e^{jn\omega_s t} e^{-st} dt = \frac{1}{T} \sum_{n=-\infty}^{+\infty} \int_0^{\infty} e^{-(st-jn\omega_s)t} dt = \\ &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} F(s - jn\omega_s) \end{aligned}$$

Si può osservare che nel dominio di Laplace, ossia al variare della variabile complessa s , la funzione F^* è una funzione periodica, che si ripete con periodo (nel dominio reciproco) pari a ω_s . Dal momento che si intende ragionare dunque sullo spettro, dato $f(t)$ si può supporre di conoscere il suo spettro di ampiezza, $F(s)$, ma dunque anche il modulo della risposta in frequenza, ossia $|F(j\omega)|$; esso potrà avere ad esempio un andamento passa-basso, di questo tipo:

Come si vede nella seconda figura, dati i conti di prima, si può dire che $F^*(s)$ (o meglio $|F^*(j\omega)|$) sia semplicemente questo spettro, traslato e normalizzato per un fattore T .

Lo spettro del segnale campionato subisce una periodicizzazione rispetto a quello del segnale di partenza, il che porta ad una ripetizione dello stesso spettro per un numero infinito di volte.

6.4 Ricostruzione del segnale di partenza mediante filtro ideale e teorema del campionamento

In molte applicazioni si pone il problema di ricostruire un segnale campionato. Campionato dunque $f(t)$, dal momento che le sue caratteristiche in frequenza sono state descritte nel precedente capitolo, ci si potrebbe chiedere: “come si può tornare indietro?”. La domanda in pratica propone il seguente

dubbio: dato un segnale campionato, a partire da esso è possibile ricostruire in maniera quantomeno soddisfacente il segnale di partenza, prima del campionamento?

Questa operazione si può teoricamente fare, e con una notevole facilità: un modo di ricostruire il segnale è basato sull'uso di un filtro ideale impostato su di una frequenza di taglio pari a $\frac{\omega_s}{2}$.

Filtrando il segnale campionato, di fatto, si riottenrebbe lo spettro "singolo", ossia lo spettro del segnale precedente al campionamento. Da qui appare chiaro ciò: data la pulsazione ω_t , determinante il termine della banda, il teorema del campionamento (teorema di Nyquist o teorema di Shannon) afferma che ω_s , ossia la pulsazione di campionamento, deve essere almeno pari al doppio di ω_t :

$$\omega_s \geq 2\omega_t$$

Ovviamente tutta la teoria finora presentata può essere applicata nel caso il segnale sia limitato in banda, e si consideri come pulsazione di campionamento quella almeno doppia del limite, come il teorema desidera; solo a tali ipotesi ha senso parlare di ricostruzione del segnale senza avere effetti sgradevoli.

6.5 Ricostruzione mediante ZOH

Dato il filtro ideale con la risposta in frequenza precedentemente presentata, si deve immaginare immediatamente una cosa: questo filtro esiste ma solo sulla carta: da un lato la risposta all'impulso dovrebbe avere una durata infinita, dall'altro esso dovrebbe essere non causale: il filtro dovrebbe essere in grado di restituire un segnale ancor prima che esso venga introdotto nel sistema (cosa assolutamente priva di senso).

Una buona tecnica per lavorare con il segnale è basata sul cosiddetto ZOH (Zero Order Hold): si tratta di un sistema che, dato un segnale campionato, ne propone uno a tempo continuo.

L'idea fondamentale è la seguente:

Sostanzialmente l'idea è quella di considerare un campione e mantenere costante la sua ampiezza finché non si abbia una variazione anche nel tempo discreto, "prolungando per continuità" il tempo discreto a tempo continuo. "Ordine zero" significa "derivata nulla", pendenza zero che si usa per l'approssimazione in questione.

Acquisito il campione, è necessario mantenerlo; per far ciò, sostanzialmente, si considera una somma e differenza di gradini, in modo da rendere

il sistema semplicemente studiabile mediante trasformate di Laplace. Considerando un singolo contributo, si può immaginare che si abbia qualcosa del tipo:

$$g_h(t) = \varepsilon(t) - \varepsilon(t - T)$$

Considerando due contributi traslati di un tempo T . Nel dominio di Laplace, ciò si può tradurre in questo modo:

$$G_h(s) = \mathcal{L}\{g_h(t)\} = \frac{1}{s} - \frac{1}{s}e^{-sT} = \frac{1 - e^{-sT}}{s}$$

Questo è il comportamento in frequenza di uno ZOH: in banda passante il guadagno purtroppo non è costante. Ciò non è tuttavia drammatico, dal momento che vi è un range di frequenze piuttosto ampio dove si può considerare almeno in prima approssimazione costante, dal momento che non subisce attenuazioni eccessive. Fondamentale è tuttavia la seguente osservazione: più alta è ω_s , ossia la pulsazione di campionamento, che detta l'ampiezza T del passo di campionamento mediante la relazione:

$$\omega_s = \frac{2\pi}{T}$$

E maggiore sarà la zona in cui il guadagno si può considerare approssimativamente costante. Il filtro ZOH è fondamentale per la ricostruzione, e viene sostanzialmente introdotto nell' "Hold Circuit" dello schema a blocchi. Sotto il punto di vista del controllo, questo circuito è inoltre abbastanza fastidioso: come si può notare dall'espressione della sua funzione di trasferimento, esso in sostanza fa perdere fase nella funzione di anello, causando dunque un ulteriore abbassamento rispetto a quelli già interni ad essa. Dal momento che è tuttavia molto semplice da realizzare, per quanto non abbia guadagno costante in banda e perda fase, esso viene sostanzialmente utilizzato sempre.

6.6 Legame tra $F^*(s)$ e $F(z)$

Si studia a questo punto un legame tra i due domini fondamentali: quello di Laplace, ossia la variabile complessa s , e quello della trasformata zeta, ossia la variabile complessa z . Finora è stata esclusivamente proposta una formulazione del segnale campionato nel dominio di Laplace, di questo tipo:

$$F^*(s) = \sum_{k=0}^{\infty} f(kT)e^{-skT}$$

Per segnali a tempo discreto, tuttavia, la descrizione più idonea è senza dubbio quella mediante trasformata zeta, \mathcal{Z} ; da $f^*(t)$ si sa che:

$$\mathcal{Z}\{f^*(t)\} = F(z) = \mathcal{Z}\left\{\sum_{k=0}^{\infty} f(kT)e^{-skT}\right\} = \sum_{k=0}^{\infty} f(kT)z^{-k}$$

Si possono dunque mappare il dominio di Laplace e il dominio della trasformata zeta secondo le seguenti trasformazioni:

$$F^*(s) = F(z)|_{z=e^{sT}} \quad F(z) = F^*(s)|_{s=\frac{1}{T}\ln(z)}$$

6.7 Schemi a blocchi di sistemi con campionatori

In questo capitolo si sta parlando di sistemi misti, ossia al cui interno si lavora sia con segnali a tempo continuo sia con segnali a tempo discreto. Dato a disposizione un sistema $G(s)$ con un campionatore sull'ingresso di riferimento $R(s)$ (considerato direttamente nel dominio di Laplace), si ha qualcosa di questo tipo:

Dopo $R(s)$ e il campionatore si avrà $R^*(s)$; si ha dunque qualcosa di questo genere:

$$Y(s) = G(s)R^*(s)$$

Quella che potrebbe interessare a questo punto è la funzione di uscita dal blocco campionata nel dominio di Laplace, $Y^*(s)$: in teoria, $G(s)$ è un blocco che potrebbe restituire un segnale analogico; usando un ulteriore campionatore sull'uscita è possibile ottenere nel tempo $y^*(t)$, dunque la sua trasformata di Laplace:

$$Y^*(s) = \mathcal{L}\{y^*(t)\}$$

Essa avrà una forma di questo tipo, come già noto dalla teoria precedentemente studiata:

$$\begin{aligned} Y^*(s) &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} Y(s - jn\omega_s), \quad \omega_s = \frac{2\pi}{T} = \\ &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} G(s - jkn\omega_s)R^*(s - jkn\omega_s) \end{aligned}$$

Si noti la seguente osservazione: $R^*(s)$ è una funzione periodica in ω_s ; scrivendola in questo modo, semplicemente, si impone la periodicità a una funzione già periodica, introducendo una condizione di fatto ridondante. Ciò che si può fare è dunque estrarre dalla sommatoria $R^*(s)$, eliminando da un lato questa ridondanza e dall'altro semplificando il sistema:

$$= R^*(s) \frac{1}{T} \sum_{n=-\infty}^{+\infty} G(s - jkn\omega_s) = R^*(s) \cdot G^*(s)$$

Quella che è stata appena trovata è la seguente relazione:

$$Y^*(s) = R^*(s) \cdot G^*(s)$$

Quello che si ha è un legame tra segnali campionati e “sistemi campionati”, anche se più propriamente bisognerebbe dire “sistemi discreti” (dal momento che solo i segnali vengono di fatto campionati; i sistemi sono “scatole” che reagiscono, in seguito a una certa eccitazione, con una certa uscita). Ciò che si potrebbe fare, in altre parole, è passare alla trasformata zeta dell'espressione, ottenendo qualcosa di questo tipo:

$$Y(z) = G(z) \cdot R(z)$$

Ciò che si può a questo punto studiare è una generalizzazione del concetto di schemi a blocchi per quanto concerne i segnali campionati. Di fatto, introducendo un processo di discretizzazione, l'algebra degli schemi a blocchi subisce variazioni non indifferenti, variazioni che comunque, mediante un “trucco”, possono essere aggirate. Si consideri uno schema di questo genere:

Si ha che:

$$Y_1(s) = G_1(s)R^*(s)$$

In seguito al processo di campionamento:

$$Y_1^*(s) = G_1^*(s)R^*(s)$$

Per quanto riguarda Y_2 si fa qualcosa del genere:

$$Y_2(s) = G_2(s)Y_1^*(s) = G_2(s)G_1^*(s)R^*(s)$$

Da qui, campionando:

$$Y_2^*(s) = G_2^*(s)Y_2^*(s) = G_2^*(s)G_1^*(s)R^*(s)$$

Avendo a che fare con funzioni campionate, è a questo punto possibile passare alle relative trasformate zeta, ottenendo:

$$Y_2(z) = G_1(z)G_2(z)R(z) \implies \frac{Y_2(z)}{R(z)} = G_1(z)G_2(z)$$

Si riassume il risultato, assolutamente non banale, appena riscontrato: dato un sistema in cui tra ogni blocco è presente un campionatore, si può dire che la funzione di trasferimento equivalente, nel dominio della trasformata zeta, sia pari al prodotto delle singole funzioni di trasferimento, esattamente come nel caso di sistemi che lavorano su segnali continui in serie.

Si consideri a questo punto un caso differente, in cui non si ha un campionatore bensì un collegamento in serie diretto tra le due funzioni di trasferimento:

Si calcoli la funzione di trasferimento equivalente. Partendo dall'uscita, si può vedere che:

$$Y_2(s) = G_{12}(s)R^*(s)$$

Finchè si ragiona con segnali continui, si può dire che, dati due blocchi in serie, un blocco equivalente sia dato dal prodotto dei due. Ora, dal momento che si considerano segnali discreti, ciò non è più possibile, dunque si deve considerare una generica funzione $G_{12}(s)$, che tuttavia non è pari al prodotto delle due. In questo caso, si può trovare naturalmente la versione campionata di $Y_2(s)$, come:

$$Y_2^*(s) = G_{12}^*(s)R^*(s)$$

Si può a questo punto passare naturalmente al dominio z :

$$Y_2(z) = G_{12}(z)R(z) \implies \frac{Y_2(z)}{R(z)} = G_{12}(z)$$

Si noti dunque che la trasformata del prodotto è diversa dal prodotto delle trasformate, in questo caso: senza il campionatore, di fatto, la considerazione precedentemente fatta non è ripetibile.

Una rappresentazione dello schema a blocchi equivalente potrebbe essere dunque la seguente:

Dove si ha, come circuito di mantenimento, lo ZOH, con una caratteristica del tipo:

$$G_h(s) = \frac{1 - e^{-s}}{s}$$

Una variante dello schema, presente ad esempio nel caso del levitatore magnetico utilizzato nei laboratori LADISPE del Politecnico di Torino, contiene ulteriori campionatori, dopo G_y :

In questo modo si riesce a sfruttare maggiormente il risultato appena presentato.

6.8 \mathcal{Z} -trasformata di funzioni contenenti ZOH

Si vuole a questo punto passare, a partire da una certa funzione $G(s)$, a una certa $G(z)$, al fine di progettare un controllore di tipo digitale. Si noti che si sta parlando di sistemi, dunque di funzioni: i segnali vengono campionati, i sistemi ovviamente no: essi possono lavorare con segnali campionati o segnali analogici. L'obiettivo finale è, a partire da un sistema analogico, ottenerne uno digitale, che cioè lavori con segnali numerici.

Il percorso che idealmente si potrebbe proporre è: immaginando di partire da $G(s)$, se ne potrebbe fare la antitrasformata nel dominio del tempo, ottenendo una funzione $g(t)$, che verrebbe quindi campionata, ottenendo $g^*(kT)$, con passo di campionamento T , per passare dunque alla trasformata zeta, $G(z)$:

Un problema di questo tipo viene spesso affrontato e studiato, con un approccio di questo genere.

Il problema che viene spesso posto, nella fattispecie (come si vedrà tra breve in un caso specifico), è trovare la \mathcal{Z} -trasformata di funzioni che contengono filtri di tipo ZOH; data $X(s)$ di questo tipo:

$$X(s) = \frac{1 - e^{-st}}{s} G(s)$$

Si vuole determinare la $X(z)$ ad essa equivalente. La cosa interessante che si può dimostrare è un risultato, fondamentale, di questo tipo:

$$X(z) = (1 - z^{-1}) \mathcal{Z} \left\{ \frac{G(s)}{s} \right\}$$

Questo risultato è fondamentale in quanto verrà analizzato sotto un altro punto di vista, tra breve.

6.9 Scelta del tempo di campionamento

Si parla di discretizzazione, di passaggio da un dominio continuo ad uno discreto. Si può immaginare che, al fine di selezionare una certa porzione di informazione, sia necessario introdurre un passo di campionamento, ossia una distanza temporale tra i punti che si intende campionare al fine di caratterizzare il sistema. Esistono diversi criteri atti a selezionare questa distanza temporale.

Finora cosa si è fatto? Ci si è preoccupati di campionare un segnale, senza scegliere il tempo di campionamento. Un buon criterio potrebbe essere quello di avere $\omega_s \geq 2\omega_t$, in modo da soddisfare il teorema del campionamento ed evitare fenomeni di aliasing; se il segnale da campionare non ha una banda limitata, si ha una sovrapposizione di “code” dello spettro del segnale, che modificano di fatto l’informazione rendendola più difficilmente interpretabile. Prima di campionare, in una situazione come questa, è suggeribile introdurre un filtro che faccia in modo da limitare la banda, tagliando le code. Ciò tuttavia non basta: bisogna tenere conto del fatto che, da un punto di vista controllistico, il controllo digitale introduce perdite di fase che dipendono dal periodo di campionamento scelto. Si sa che:

$$G_h(s) = \frac{1 - e^{-sT}}{s}$$

Questa funzione può essere approssimata, per valori di T elevati (dunque per le basse frequenze):

$$G_h(s) \simeq \frac{1}{1 + s\frac{T}{2}}$$

In prima approssimazione l’inserimento dello ZOH è sostanzialmente coincidente con l’inserimento di un polo nel sistema con costante di tempo pari a $\frac{T}{2}$, dove T è il solito tempo di campionamento. Il polo nella funzione di anello fa perdere fase, quindi ci dà fastidio; sappiamo tuttavia in buona approssimazione dove esso sia dislocato, dunque è possibile scegliere un T tale da evitare che esso provochi problemi troppo grossi.

Quanto questo polo è vicino alla pulsazione ω_c , ossia la pulsazione di passaggio per l’asse 0 dB? Beh, si spera e si desidera che esso si trovi a una pulsazione molto più elevata, in modo da rendere ininfluente il polo in un intorno dell’asse a guadagno unitario. Utilizzando ancora l’approssimazione, osservando la fase, si può vedere che:

$$\arg \left\{ \frac{1}{1 + j\omega\frac{T}{2}} \right\} = -\arctan \left(\omega\frac{T}{2} \right)$$

Si introducono a questo punto idee molto approssimate e metodi grafici atti a capire come determinare T . La domanda fondamentale è: quanta fase si è disposti a perdere, a causa dell’introduzione di questo dispositivo? Beh, a seconda di T , si potrebbe essere indicativamente disposti a perdere dai 5° ai 15° massimo, in ω_c , rispetto a prima.

Esiste una tabella in grado di quantificare le perdite:

Considerando un T ridotto si riduce la perdita di fase; d'altra parte tuttavia, scegliere tempi piccoli significa avere un hardware molto veloce, ma dunque molto costoso, sia in termini di convertitore sia in termini di elaboratore numerico. Spesso, nei progetti reali, dunque, si è costretti a scegliere tempi non troppo ridotti.

Ciò che si farà in sostanza è: dato un “buon” progetto si può perdere anche tanta fase, prediligendo il fatto che l'hardware utilizzato sia lento e scadente. In questo modo, si vincola ω_c : essa non potrà essere più di tanto alta.

6.10 Metodi di discretizzazione

Verranno a questo punto presentati alcuni metodi di discretizzazione per sistemi continui, analogici. Questi saranno utilissimi per il seguente scopo: dato il progetto di un sistema analogico, esso si può, tenendo conto di alcuni fattori correttivi che verranno tra breve introdotti, convertire in un progetto analogico, mediante alcuni step abbastanza semplici.

6.10.1 Metodo di invarianza della risposta all'impulso (\mathcal{Z} -trasformata)

Data una generica funzione (si noti che ci si sta concentrando su funzioni, non su segnali), ci si pone il problema di come trasformarle, correggerle nel passaggio dalla variabile complessa s alla variabile complessa z . Questo primo metodo consiste sostanzialmente in ciò: data una $F(s)$, si può trasformare come:

$$F(s) \implies F(z) = \mathcal{Z}\{F(s)\}$$

Questo procedimento naturalmente non è immediato: è necessario effettuare l'antitrasformazione nel tempo, la discretizzazione e il passaggio nel dominio z ; questa notazione non tiene conto di questi fatti, che comunque devono essere ben noti. Questo metodo garantisce il fatto che, se si sollecita il sistema a tempo continuo con un impulso, quindi poi il sistema a tempo discreto con un medesimo, le uscite dei due sistemi negli istanti di campionamento assumeranno gli stessi valori. Ciò introduce una relazione di equivalenza tra i due sistemi, quello a tempo continuo e quello a tempo discreto. Esistono diversi “sensi” di equivalenza, ossia diverse relazioni per cui si possono avere corrispondenze tra i due sistemi. In questo caso l'equivalenza coinvolge la risposta all'impulso: solo per quanto riguarda un'eccitazione

impulsiva (non dunque altri tipi di eccitazioni) le risposte dei due sistemi avranno, per i valori del dominio z , le stesse uscite. Ovviamente con una rampa, un gradino, una parabola, questo tipo di metodo non fornisce alcuna informazione tra le correlazioni dei sistemi.

6.10.2 Metodo di invarianza della risposta al gradino

Si può fare a questo punto un ragionamento quantomeno simile al precedente, imponendo tuttavia l'invarianza di un'altra risposta: se per ora è stata imposta l'invarianza della risposta all'impulso, ora si impone l'invarianza della risposta al gradino, e si ricercheranno risultati in proposito ad essa. Data $F(s)$ iniziale, si intende passare nel dominio z ottenendo la funzione $\bar{F}(z)$, mediante un metodo di discretizzazione che faccia coincidere, negli istanti di campionamento, le due risposte al gradino. Data la $F(s)$ che si intende discretizzare, si otterrà qualcosa di questo tipo:

$$\mathcal{L}^{-1} \left\{ \frac{1}{s} F(s) \right\} \Big|_{t=kT} = \mathcal{Z}^{-1} \left\{ \frac{1}{1-z^{-1}} \bar{F}(z) \right\}$$

Queste due funzioni devono coincidere. Dal momento che si intende ottenere l'espressione nel dominio z , si opera la trasformazione \mathcal{Z} , ottenendo:

$$\frac{1}{1-z^{-1}} \bar{F}(z) = \mathcal{Z} \left\{ \mathcal{L}^{-1} \left\{ \frac{1}{s} F(s) \right\} \right\} \Big|_{t=kT}$$

Dunque:

$$\bar{F}(z) = (1-z^{-1}) \mathcal{Z} \left\{ \mathcal{L}^{-1} \left\{ \frac{1}{s} F(s) \right\} \right\} \Big|_{t=kT}$$

Si faccia a questo punto un'osservazione interessante: questo risultato è già stato visto, nella fattispecie parlando di funzioni di trasferimento contenenti blocchi ZOH. Risulta a questo punto fondamentale la scelta dell'uso di filtri di questo tipo: questo metodo di discretizzazione garantisce la stessa risposta al gradino, dunque anche l'uso del medesimo blocco.

Si noti che il procedimento di discretizzazione si può effettuare in via informatica, mediante il software MATLAB, facendo uso del seguente comando:

c2d

Impostando alcuni parametri quali il passo di campionamento e il metodo utilizzato (in questo caso, metodo 'ZOH'), si riesce a ottenere immediatamente un'espressione di $G(z)$.

6.10.3 Metodo di trasposizione zeri-poli (matched)

Esiste un terzo metodo, piuttosto interessante, che verrà presentato solo come “idea”; nel caso si volesse usarlo, è sufficiente esplicitare in MATLAB il metodo 'matched'.

L'idea che si intende utilizzare è: data una funzione di trasferimento $G(s)$, quello che si può fare è considerare la trasformazione:

$$z = e^{sT}$$

A questo punto, della funzione di partenza, si considerano gli zeri e si trasformano secondo questa; terminato con gli zeri, si trasformano i poli. Si impone, una volta trasformate le pulsazioni continue in pulsazioni discrete, il fatto che i guadagni stazionari delle due funzioni siano gli stessi. Mediante questo metodo si possono mantenere “bene” le caratteristiche in frequenza del progetto, cosa molto importante dal nostro punto di vista (poiché il progetto viene di fatto effettuato in frequenza), dunque questo è il metodo più suggeribile, al momento della discretizzazione del controllore.

6.11 Progetto di controllori digitali per discretizzazione di controllori analogici

Il metodo di progetto di controllori digitali, come già detto, è basato sul progetto di un controllore analogico equivalente, che poi verrà discretizzato. Tendenzialmente, una volta presa l'equazione in s , la si trasforma informaticamente in un'equazione alle differenze finite, per poi implementare il controllore sul calcolatore. Si vedranno rapidamente alcune idee riguardo la conversione da effettuare.

6.11.1 Progetto di $G_c(z)$ per loop shaping di $G_a(s) = G_c(s)AG_pG_tG_y$

Questo primo procedimento è assolutamente banale: si progetta il controllore analogico, quindi lo si discretizza. Come già detto più e più volte, il ZOH comporta una perdita di fase nel sistema, dunque sarà necessario introdurre uno studio della funzione di anello mediante comandi in grado di agire su sistemi discreti, campionati, quali gli equivalenti discreti dei comandi “bode()” e “nichols()”, ossia “dbode()” e “dnichols()”.

6.11.2 Progetto di $G_c(z)$ per loop shaping di $G'_a(s) = G_a(s) \cdot G_h(s)$

Ciò che si potrebbe fare per migliorare il metodo precedente di progetto è tener conto fin da subito del polo introdotto dal ZOH, approssimando la funzione di anello globale con una funzione di anello contenente il polo introdotto dal nuovo sistema:

$$G'_a(s) = G_a(s) \cdot \frac{1}{1 + s\frac{T}{2}}$$

In questo modo si introduce un polo fittizio fin dal progetto analogico, tenendo immediatamente conto di esso, e dunque sovradimensionando in partenza il sistema, in modo da non presentare grossi problemi al momento dell'implementazione digitale. Il risultato finale sarà abbastanza approssimato, ma in molti casi comunque interessante ai fini di ottenere un buon progetto.

6.11.3 Progetto di $G_c(z)$ mediante loop shaping di $G_a(z)$

Un approccio che non verrà descritto ma solo accennato in queste poche righe riguarda la modifica della forma direttamente sulla funzione di anello nel dominio z , $G_a(z)$. Questo metodo un tempo era poco utilizzato poichè difficile da implementare; con i moderni metodi, tuttavia, esso è assolutamente fattibile e sotto molti punti di vista anche interessante, poichè permette di agire direttamente sul sistema discreto, non introducend approssimazioni nel passaggio da sistema continuo a sistema discreto. Si lavora sulla funzione:

$$G_a(z) = G_c(z)AG_tG_y\overline{G}_p(z)$$

Dove $\overline{G}_p(z)$ altri non è che la funzione di trasferimento del sistema discretizzato mediante metodo ZOH.